Emergence as Self-organization and as Generation of Novelty

A Framework for Understanding Emergence in the Context of Interactive Art

Joan Soler-Adillon

TESI DOCTORAL UPF / 2015

DIRECTORS

Dr. Simon Penny (University of California, Irvine)
Dr. Lluís Codina (Universitat Pompeu Fabra)

DEPARTAMENT DE COMUNICACIÓ

Universitat
Pompeu Fabra
Barcelona

A la Noemí, la Joana, la Mariona i el que està en camí.

To Noemí, Joana, Mariona, and the one who is on the way.

Agraïments / Acknowledgements

Als meus pares, pel seu suport constant.

Als meus germans, per ser-hi també sempre.

Als avis, pels qui això arriba tard, però que són sempre recordats en els moments importants.

To my thesis director Simon Penny, without whom this work would have not been possible. Thank you for the help, the patience and the very valuable reviews. Thanks for helping in every step of the way.

Al meu altre director, Lluís Codina, que des del primer dia ha estat un conseller de gran valor, sempre disposat a ajudar amb tot.

To Peter Cariani, with whom I was lucky enough to meet and spend some time in Boston. Thank you for the intellectual (and actual) hospitality.

A l'Ignasi Ribas per obrir-me les portes de la Universitat i donar-me la confiança de poder-hi treballar en un ambient de respecte i vàlua intel·lectual molt difícil de trobar. Gràcies pel suport incondicional i per ser un exemple en tants aspectes.

A la resta de companys de la Pompeu: en Narcís Parés, que com l'Ignasi va confiar amb mi des del principi; en Pere Freixa, germà gran; en Roc Parés, únic; i en Carles Sora, company de trinxera i col·lega de present i de futur.

A en David Casacuberta, que sense saber-ho va posar la primera pedra de tot això.

To my instructors and fellow students in New York University's Interactive Telecommunications Program: to Dan O'Sullivan, who first introduced me to programming; to Danny Rozin, with whom I discovered how to see the world pixel by pixel; to Daniel Shiffman, who was then just starting to teach, and who has been a role model since. And to the great Red Burns, who had started it all and whose warmth I was lucky enough to encounter. May she rest in peace.

Als alumnes de tants i tants cursos, que fan que tot plegat valgui la pena.

Als músics que sense saber-ho m'han acompanyat durant gran part del redactat.

A tots els que van deixar de preguntar-me quan acabaria la tesi.

Abstract

The subject of study of this investigation is the concept of emergence and its viability as a driving force for the creation of interactive artwork. Emergence appears in the literature as related to self-organization and novelty. For many authors it is the result of multiple local interactions among agents within a system, which generate phenomena that could not be understood, nor anticipated, through the analysis of the elements and their behaviors in isolation. For others, emergent phenomena are related to fundamental novelty and, thus, to creativity. These two formulations of emergence are reviewed in this thesis in the context of interactive art and communication. The focus of the study is on the interactive works of the artistic field known as Artificial Life Art. After reviewing the foundational concepts and proposing an approach to both emergence and interactivity, an analytical framework is presented in order to assess the presence of emergence in interactive systems, which in turn shall serve as a first step to designing systems that aim at generating interactive emergent behaviors.

Resum

L’objecte d’estudi d’aquesta tesi és el concepte d’emergència i la seva viabilitat com a força creativa d’obres d’art interactiu. Emergència apareix a la literatura com a concepte relacionat tant amb el concepte d’auto-organització com amb el de novetat. Per a molts autors és el resultat d’una gran multiplicitat d’interaccions locals entre els agents d’un sistema, que generen fenòmens que no podrien entendre’s ni anticipar-se mitjançant l’anàlisi d’aquets agents i dels seus comportaments de manera aïllada. Per altres, els fenòmens emergents estan relacionats amb la idea de novetat fonamental i, per tant, de creativitat. La tesi revisa ambdues formulacions del concepte d’emergència en el context de l’art interactiu i la comunicació. El focus central de l’estudi són les obres d’art interactiu de la disciplina coneguda com Artificial Life Art (art i vida artificial). Després de revisar-ne els conceptes fonamentals i proposar una formulació tant d’emergència com d’interactivitat, es presenta un model d’anàlisi que permet donar compte de la presència de fenòmens emergents dins en els sistemes interactius, i que alhora és un primer pas cap al disseny de sistemes que tenen la intenció de generar comportaments interactius emergents.

TABLE OF CONTENTS

PART 1: SURVEY AND HISTORY

PART 2: EMERGENCE AND ARTIFICIAL LIFE ART

“The possibilities exist for works that perceivably never repeat themselves. Works that respond to their environment not just in a short-term way, but in a long-term way, unpredictably and meaningfully (easier said than done)” Jim Campbell

0. INTRODUCTION

0.1 Subject of Study

The subject of study of this investigation is the concept of emergence and its viability as a driving force for the creation of interactive artwork.

The concept of emergence is used diversely in different disciplines, and within these the approaches can vary greatly among authors. This thesis aims at elaborating a sound conceptual basis in order to discuss emergence in interactive art and digital media. This includes the investigation of a large enough spectrum of approaches to this central concept and to those related, such as that of interactivity. A fundamental goal here is to frame both interactivity and emergence within the context of interactive communication.

Among the surveyed fields of study, a special focus of interest is placed on Artificial Life (ALife) Art, a subdiscipline of Artificial Life that intersects with digital art practices. It is in ALife Art were several art pieces have been built with the idea of emergence and emergent behavior in mind, and were the discourses have often referred to this idea in relation to artistic creativity.

Emergence is often presented with the idea of the whole that is more than the sum of its parts. That is, this whole is presented as being irreducible to the analysis of its constituting elements in isolation. These explanations are usually articulated in terms of different levels of complexity, in which the lower levels (the parts) generate processes which appear at the upper level (the whole) that are not explainable with a classic cause-effect relationship. Typical examples include ant or termite colonies and their social complexity, the human mind understood as a product of the interconnectivity of neurons in the brain, chemical clocks in non-equilibrium thermodynamics, or the complexity generated from the simple rules of cellular automata. Usually, the notion of novelty is also related to emergence in one way or another.

As said, here the interest is on discerning whether or not the notion of emergence can be formulated in a way that makes it into a viable idea to discuss, analyze and create interactive art pieces. In this respect, the approach to the subject of study is in this thesis epistemological and pragmatic. It is epistemological in the sense that the interest does not reside in emergence as a part of reality in the ontological sense, but in elaborating a discourse on how an observer can identify a process as emergent and how this can be scientifically communicated. It is pragmatic because the ultimate goal is to discern strategies for the creation of artificial artistic devices that generate emergent behavior.

0.2 Research Goals

Research goal #1 (RG1): Review the state of the art, covering a body of literature that accounts for the main approaches to the concept of emergence that relate to Artificial Life and, thus, to Artificial Life Art.

As said above, emergence is a topic that appears in a wide variety of contexts: Philosophy, Psychology, Neurology, Chaos Theory, Dynamic Systems Theory or Complexity Sciences among others. Thus, covering a large enough ground while maintaining the points of contact among discourses is on its own already a challenging task.

In addition, the word emergence presents some polysemy issues. Firstly, because emergence also means, simply, appearance or coming to surface: 'the emergence of the Chinese economy' is certainly a correct use of language, but it does not relate to the topic of interest here. In other occasions, emergence is used simply as a synonymous of openness. Secondly, because in some disciplines, such as Cybernetics, terms like self-organization are used essentially interchangeably to what here is referred to as emergence. Therefore, part of the effort is in discerning where emergence is used properly and where it is not, and were other terms are used that are actually synonyms to emergence.

Research goal #2 (RG2): Make a proposal to understand emergence in relation to interactive art that is inclusive and that strives for simplicity.

Defining emergence is not an easy, neither a new task. There are several definitions that serve different purposes. The idea here is not to formulate the definitive definition of emergence, but to propose a formulation that is both simple enough to operate with and that is inclusive in respect to most, if not all, of the relevant literature. The cybernetician Grey Walter stated that models should be simple, transparent and semantically brittle (Walter, 1963). The proposal here will aim at these this simplicity and transparency, as opposed to the baroqueness that is associated with many of the available discourses and definitions of the idea.

Since the context is that of interactive art, this research goal also includes the need for the definition of another complex and polysemyc idea: interactivity. An entire chapter is dedicated to discussing the different approaches to the idea and to elaborate a proposal to understand interactivity not as a functional goal but as an aesthetic one.

Research goal #3 (RG3): Propose and test an analytical framework for interactive art pieces that deal with the concept of emergence.

The third goal is to propose a model of analysis. An analytical framework that can help clarify whether or not interactive art pieces are emergent in any sense: that is, if they result in emergent behavior or if new emergent features or processes appear on the analyzed works. The framework will be applied to three case studies in order to test and refine it. As far as the reviewed literature is concerned, such a method has never been presented.

0.3 Relevance of Research

The ideal of emergent interactive behavior was articulated twenty years ago and it is still an unresolved issue. Arguably, this is in part because of the limitations of the technology at the time and partially because the very idea of emergence has too often remained in a limbo of ambiguous definitions. This thesis aims at contributing to the path of clarifying the concept and to make it useful as a creative tool.

A-Life Art is far from being a mainstream practice, and the idea of linking emergence and interactivity had its momentum in the near past. However, the implications of ALife are more important than one might think at the first glance. Robotics or special effects are among the disciplines that have been strongly influenced by it. For the purposes of interactive art and interactive communication in general, there is still a great deal to learn the bottom-up approach that goes along with the idea of emergence. ALife Art is an excellent ground to study the implications of this approach to interactivity in order to generalize a strategy that aims at creating experiences of interactive communication that are complex in a positive and interesting way.

This approach would be applicable to any domain within interactive communication whenever an artistic or experimental approach to interaction is desirable. Whether it is in the context of a multimedia product, a non-linear narrative experience, a videogame or a work of art, emergence can be a very useful idea to enhance the possibilities of the work. In the domain of interactivity, where contemplation is not anymore the dominating paradigm, the need for interesting and enriching interactive experiences will become more and more evident. Emergent interactive behavior can be a very rich framework to work with in this respect.

Currently, interactivity as it was understood in the digital art of the 1990s is only one approach to the relation between technology and its users[1]. Interaction understood as participation has taken a greater role in the era of social media. However, as big data grows and as connected robotics and wearable accessories become available, how we actually interact with technological devices will become again more and more relevant. It is in this context that understanding and being able to account for emergence (that is, for self-generated novelty and behaviors) in the domain of interactivity will be of most relevance.

---

[1] In the interactive arts discourses, the terminology to refer to that who uses the interactive technology often tries to avoid the Human-Computer Interaction term ‘user,’ to stress the fundamental difference in the creative approach. When referring to interactive installation, ‘visitor’ is a recurring alternative. In addition, several neologisms have appeared over the last fifteen years to address this: terms such as ‘vuser’ (from visitor-user) (Seaman, 2003), ‘spect-actor’ (Frasca, 2001), ‘interactant’ (Baljko and Tenhaaf, 2006) or ‘interactor’ (Hayles, 2004) among them. This thesis uses interactor preferently, but for stylistic purposes visitor and user are also used and considered to be synonyms.

0.4 Research Questions

This section specifies the research questions that constitute the point of departure of this investigation. As said above, digital media is only one among many fields within which the concept of emergence has been investigated. According to this, the first three questions aim at identifying how emergence can be formulated as a useful concept within this particular field, and whether or not the fundamental idea of emergent interactive behavior is coherent according to this. Finally, the last question is concerned with the idea of interactivity, which is yet another concept that is used diversely in different disciplines. Here the question is concerned with how to define this last part of the equation; the relation between the interactive system and the interactor.

Research question #1 (RQ1): Is emergence a useful concept to discuss interactive art?

Emergence is far from being paramount in the discussions on interactive art, let alone on art in general. In fact, the uses of the term are rare outside ALife Art. In addition, only part of the artwork produced within the ALife Art parameters is interactive. Thus, art pieces dealing with interactivity and interaction are just a subset of a subgenre of interactive art. This means that applying the idea of emergence to other domains of interactive communication, and in this case to interactive art in general, is an expansion of this original confinement. Therefore, the question of whether or not this task is viable is justified.

Research question #2 (RQ2): How can we formulate the concept in an unambiguous manner to use it in this context?

In relation to RQ1, how can the concept be formulated unambiguously in order to be used in this larger context? Emergence is often presented though the idea of the whole being more than the sum of the parts, and in relation to surprise and novelty. But often the definitions are not elaborated too much beyond this point, leaving to the reader's intuition what the idea means exactly. Overcoming these ambiguities with a soundly formulated definition and a method of analysis is crucial in the effort undertaken here.

Research question #3 (RQ3): Is the idea of designing emergent behavior coherent?

Can emergence be designed? Emergence is often defined in relation to novelty, surprise or, at least, an important degree of indeterminacy. Thus, an emergent behavior is necessarily linked to these terms. And if it is so, designing it (i.e. pre-specifying the behavior) would invalidate it as such. Thus, it would appear that designing emergence is by definition paradoxical. The question, then, is whether this paradox is only apparent or if it is an insurmountable obstacle.

Research question #4 (RQ4): How can we define interactivity coherently with the idea of emergent interactive behavior?

Interactivity is as much a problematic concept as emergence. It too has a lot of different definitions and approaches, and in order to elaborate a clear discourse it is necessary to state what one means by it exactly. Thus, given RQ3 can be answered positively, a clear definition of the type of interactivity that emergent interactive behavior relates to is also necessary.

0.5 Theoretical Framework

The thesis proposed here is a multidisciplinary investigation. The ramifications of emergence extend from Philosophy to Chaos Theory or Cybernetics. Therefore, part of the effort is to cover a large spectrum of approaches to the idea from different disciplines. The common denominator of the reviewed discourses is the relation to the field of Artificial Life, either directly or as a relevant influence.

Artificial Life, which is central to this study, is strongly linked to Complexity Sciences. This scientific movement developed in the late 1980s and that is deeply rooted in some of the central themes of Cybernetics, a discipline that had flourished and decayed in the preceding decades. Out of the scientific production of ALife grew the artistic movement of ALife Art, at the intersection of Complexity Sciences and digital art. ALife Art borrowed themes and techniques from scientific ALife.

Digital art developed in parallel to the development of the computer. It was a rare practice before the personal computer and increasingly common once the use of computers generalized. Within digital art, interactive art practices place interactivity in the center of interest. The relation of that who uses the technology with the technology itself becomes a matter of artistic interest. Artificial Life Art is found where digital art practices coincide with Artificial Life either thematically or by the use of the same techniques.

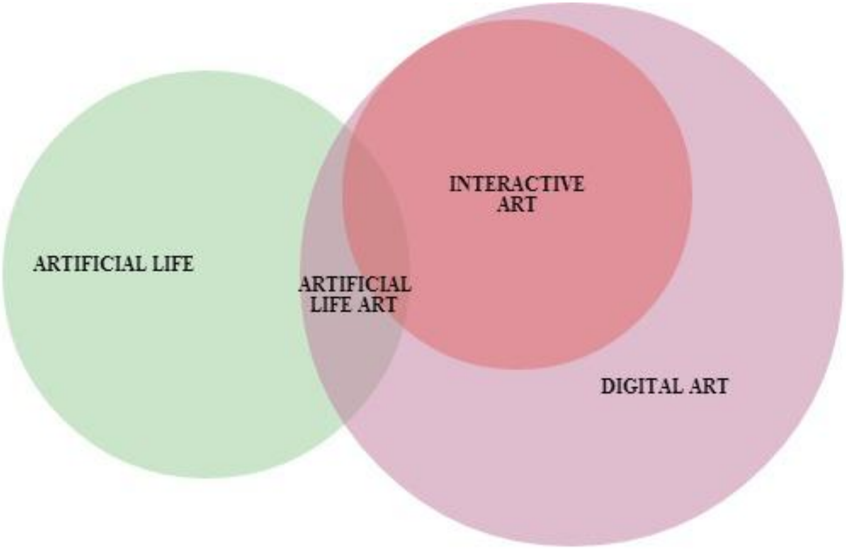


Figure 1.1: ALife Art, a discipline at the intersection between Artificial Life and Digital Art.[2]

---

[2] No author being referenced in the figures (images and diagrams) and tables in the thesis means that these figures' authorship is the same as the thesis.

0.6 Methodology and Sources of Information

Methodologically, this investigation has followed the following procedures:

a. Analysis of scientific and intellectual production

The survey of literature has been based on the analysis of books, journal articles and conference proceedings on a variety of disciplines. The references to review were found though different methods: keyword-driven search queries on academic databases, citations within already known reference, search in specific publications, and recommendations by thesis directors and other consulted experts. The most used keywords used for the searches, besides names of known authors, were: emergence, emergent behavior, self-organization, artificial life, artificial life art, interactive art, digital art, interactivity, interaction, cybernetics, creativity.

The main academic databases used were Scopus, ISI Web of Knowledge, ACM, IEEE Explore, Taylor and Francis Online and Google Scholar. Specific publications consulted were the MIT Press' Artificial Life Journal and the proceeding of the different editions of the International Workshop on the Synthesis and Simulation of Living Systems. Books from a large variety of publishing houses were used. Prominent among them, major academic publishers such as MIT Press or Springer. Specifically, a 2008 MIT Press volume entitled 'Emergence. Contemporary readings in philosophy and science and' and Mitchel Whitelaw's 2006 'Metacreation' were very valuable sources of information and of further references. The online database The Internet Archive was the source for digital copies of some of the oldest books consulted for the thesis.

b. Analysis of artistic production

Along with the scientific literature, the artistic production of the ALife Art movement has also been scrutinized. Some of the most important works of the movement are among the artwork that this thesis analyses. Relevant works were chosen through the reading of Whitelaw's book mentioned above, Steven Wilson's 'Information Arts,' a recollection of digital art pieces published in 2002, and especially through the database of the different editions of VIDA, Art and Artificial Life International Awards. VIDA is an international ALife Art festival organized by Fundación Telefónica in Spain that over the last 16 years has accumulated most of the relevant works in ALife Art as either awards or mentions, all of which are referenced in the festival's website.

Access to detailed documentation on the pieces, however, is not always easy. On occasions book sections or papers have been published, but on others website, photographic or videographic pieces is all the documentation available, which doesn't necessarily lead to a satisfactory knowledge of how the pieces are made.

c. Case studies and key examples

In the first part of the thesis, a series of key examples of what is usually understood as emergence are presented. It is important to note that these are not necessarily relevant examples of what this thesis understands as emergence. In any case, these examples or case studies, once presented, are referenced throughout the work with the understanding that the reader is familiar with them. The key examples are: an organism known as the Slime Mold, the Belousov-Zhabotinsky chemical clock reaction, the group behavior of an ant colony, John Conway's cellular automata Game of Life, Craig Reynold's simulation of a flock of birds and the simulated ecology Tierra, by Thomas Ray. In the second part of the thesis, twenty works of ALife Art are presented and discussed, and used as case studies for the rest of the thesis.

d. Analytical Framework

The last chapter of the thesis presents an analytical framework to discuss the presence of emergent behaviors in interactive art. In turn, this framework is used to analyze three of the art case studies. The framework is based on previous work on interactive systems analysis, and it assumes the definition of emergence that is defended in this work.

e. Interviews with experts, conferences and panels

Finally, informal meetings and interviews, both in person and online, have been conducted throughout the years that this investigation has lasted. Outputs of the research have been presented in international conferences such as the International Symposium on Electronic Art ISEA11 in Istanbul, or the IEEE 2014 Conference on Norbert Wiener in the 21st. Century. In this conference, I chaired the panel Cybernetics, Art and Creativity, in which emergence was a central theme. Three major research figures in Cybernetics and emergence participated in the panel: Paul Pangaro, Peter Cariani and Andrew Pickering.

PART 1: SURVEY AND HISTORY

1. CONCEPTUAL FOUNDATIONS AND CASE STUDIES

This preliminary chapter introduces a series of concepts and case studies with the goal of clarifying some of the most important and recurrent terms that will appear throughout the following chapters. What is presented here covers only a small fraction of the subsequent discussion. Along with many others, the ideas introduced here will be discussed in-depth throughout the thesis, but a general introduction of some selected ideas can be useful to generate a first landscape of some of the recurrent themes that will appear in the discussion. Note that these concepts are here presented as they are generally understood, with introductory annotations to its relation to the main approach of this thesis.

The second part of the chapter introduces a series of case studies. All of them are recurrent examples of emergence in the literature. They will be used in the subsequent chapters as recurrent reference points during the discussion of the many theories and approaches to emergence, with the understanding that they are familiar to the reader. Chapter 6 introduces a series of artistic case studies that complement those presented here.

1.1 Discussion of Terms

1.1.1 Emergence

As discussed in the introduction, emergence is often explained with the idea of a whole being ‘more’ than just the sum of its parts; of being irreducible to the mere addition of its constituting elements. The explanations are usually articulated in terms of different levels: the parts of the system and the systemic level (the whole). What is emergent are properties and behaviors that appear at the group level, which have a degree of complexity that, as it is argued, cannot be accounted for by a simple addition of the properties and behaviors of the parts.

The typical examples of emergence are of systems that exhibit complex behaviors from a relatively small set of simple rules. Such examples are, for instance, ant colonies or the human mind, as mentioned, and also traffic jam patterns or neighborhood configurations in cities, computer simulations of crowd behavior and particle systems that move in group exhibiting orderly patterns. When related to novelty, emergence is often used when referring to learning systems or adaptive devices.

The idea questions the traditional reductionism of science, since it implies that not everything is explained by studying smaller and smaller parts of whatever system is under analysis. Whenever emergence is present, reductionism is brought into question. A classical generic example would be to question, in the succession of orders of knowledge physics-chemistry-biology-psychology, if each one is fully reducible to the previous or if, instead, emergence occurs when a level of complexity increases. It also refers to the appearance of order in chaotic and far-from-equilibrium thermodynamics, and it is arguably equivalent to the cybernetic notion of self-organization.

This thesis defends that, in fact, emergence refers to two different concepts: self-organization and generation and appearance of novelty. These two aspects of emergence can appear simultaneously or independently, as none of them depends on the other. As discussed throughout the following chapters, this characterization allows for the accommodation of most of the discourses of emergence in an integrated discourse, while discussing the differences in the diverse approaches.

1.1.2 Reductionism

Classical and modern sciences operate mainly within the reductionist paradigm. It is, in a nutshell, the idea that any system can be understood by breaking it apart and analyzing all of its parts. Reductionism holds that every system’s behavior and component is completely explainable, in terms of structure and behavior, by analyzing its constituting elements. Anything can be accounted for by going into a more elementary level and explaining whatever is found there (e.g. bodily functions are explained by systems made out of organs, which in term are made of cells, and so on).

It also implies the idea that some scientific disciplines are more basic than others and that, consequently, those at superior levels of complexity are, at least theoretically, entirely explainable by the principles and theorems of the more basic ones. Reduction is, in this sense, “the explanation of a theory or a set of experimental laws established in one area of inquiry, by a theory usually thought not invariably formulated for some other domain” (Nagel, 1961: 338). Typically, physics is the basic science and chemistry, biology and psychology constitute the escalation in complexity. Thus, according to the reductionist approach any psychological phenomenon, for instance, is theoretically reducible to a biological theory, which can be broken down in terms of chemistry and, ultimately, of physics.

Therefore, the reductionist approach is always to break down a system and analyze its parts in isolation. When these are understood, the system can be understood too. Any gap in this chain of explanations, any black box that is inaccessible to knowledge, is only the result of temporary ignorance. Eventually, advances in human knowledge will necessarily –at least theoretically– be able to fill in the gap or open the black box for examination and explanation.

It is in contrast to this method that many of the approaches analyzed in this thesis are non-reductionist. They relate, instead, to a more holistic approach which considers that systems can only be understood when also the system as a whole and the interactions among its components are taken into account. If emergence is colloquially known as the idea that a whole is more the sum of its parts, reductionism would be the position that holds that the system is, in fact, nothing but the sum of its parts. This does not mean that emergentism is necessarily anti-reductionist. When authors of several disciplines defend the idea of emergence, there is usually no intention in denying reductionism, much less attacking it as a valid scientific idea. What they say it that, under certain circumstances, reductionism is either insufficient or, simply, inappropriate.

1.1.3 Complexity Sciences and Artificial Life

Complexity is a name given to a series of disciplines that, for the most part, acquired an important status in the last two decades of the twentieth century. Dynamical Systems Theory, Neural Networks or Artificial Life fall within the Complexity Sciences category and in many aspects Cybernetics can be considered a predecessor of it. All of them are disciplines that pursue an approach which differs to the classic reductionism of science, as they often seek to understand processes realized by large groups of agents. As such, it is a natural ground for the discussion of emergent phenomena.

The variety of disciplines that Complexity Sciences encompass makes it difficult to delimitate. Philosopher John Protevi proposes to do so in opposition to Chaos: “complexity theory is not chaos theory. Chaos theory treats the growth of unpredictable behavior from simple rules in deterministic nonlinear dynamical systems, while complexity theory treats the emergence of relatively simple functional structures from complex interchanges of the component parts of a system. Chaos theory moves from simple to complex while complexity theory moves from complex to simple.” (Protevi, 2006).

Artificial Life is a discipline that, with Cybernetics as a fundamental antecessor, arose during the decade of the 1980s around the figure of Christopher Langton in the Santa Fe Institute of New Mexico. As Langton presented it, “Artificial Life is the study of man-made systems that exhibit behaviors characteristic of natural living systems. It complements the traditional biological sciences concerned with the analysis of living organisms by attempting to synthesize life-like behaviors within computers and other artificial media. By extending the empirical foundation upon which biology is based beyond the carbon-chain that has evolved on Earth, Artificial Life can contribute to theoretical biology by locating ‘life-as-we-know-it’ within the larger picture of ‘life-as-it-could-be’” (Langton, 1988). According to him, Artificial Life was an extension of biology’s study of life (which had to be restricted to “the study of life based on carbon-chain chemistry”) or the material basis of life. The new discipline, instead, was aimed at the formal basis of life.

It is not a coincidence that this new discipline arose at a time were computer simulations were beginning to become available for most researchers. The computer was in fact the medium in which these researchers were able to create, breed and experiment with their artificial creatures. Genetic algorithms, neural networks, bottom-up robotics and many approaches that took biological systems as a model were developed within this context. In parallel, a several artists developed these same techniques into their own realm, thus creating the discipline known as Artificial Life Art. A more detailed account of Artificial Life is found in Chapter 3. Artificial Life Art will be analyzed in detail in chapters 6, 7 and 8.

1.1.4 Self-Organization

Self-organization is the process by which a group of agents (particles, molecules, animals…) generates, through local interactions, a process by which organization –generally a pattern of behavior– appears at the group level. Despite what could be interpreted if the term were to be understood literally, self-organization describes a process where there is no real self involved (Kelso, 1994). It is not that the agents in a group organize themselves without external intervention but, rather, that organization at the group level appears spontaneously though their local interactions. What appears is a pattern of group behavior that is not coordinated by any singular or specific group of agents.

This thesis defends that self-organization is the most common formulation of emergence, the one that relates to the idea than the whole is more than the sum of the parts. This idea is articulated in the sense that the self-organized phenomena are in one way or another beyond what one might usually anticipate from the analysis of the elements or agents which, thought interactions, generate it. In fact, many of the examples of emergence discussed in this dissertation are actually cases of self-organization. Arguably, emergence and Self-organization are interchangeable in some disciplines. Much often, self-organization is used to refer to the same phenomena in Cybernetics or Complex Systems Theory than emergence in Artificial Life. This is one of the main reasons for the concept of emergence as self-organization defended here.

Self-organization refers also to the idea of emergence as ‘order out of chaos,’ an expression popularized by Ilya Prigogine, when he used it as a title for a 1984 book in which he presented, to a non specialist audience, his work on far-from-equilibrium thermodynamics for which he had been awarded a Novel Prize in 1977. The idea refers to the spontaneous appearance of order in chaotic and stochastic systems. A group of particles, for instance, which are apparently randomly interacting with each other, suddenly starts exhibiting patterns of system behavior, i.e. order, when certain conditions are given. A key concept here is that of the role of the observer. Patterns appear in respect to an observer who is analyzing the system, and whom, more often than not, has a different point of view than that of the agents. It is precisely this point of view what allows him or her to consider the events occurring at the group level.

The idea of self-organization is easy to understand if we compare three different human group activities: football (soccer), a social cooperative organization, and a traffic jam. In football, there are certainly patterns involved. Individual players interact with each other and alternatively the team defends and attacks, for instance. But there is a hierarchical structure with a coach telling everyone were to be and how to play. That is, doing the organizing: the football coach elaborates a game-plan and communicates to each one of the players what his or her role should be. Although there’s obviously decision making executed locally –each player decides what to do at any particular time– the coach has both the ability and the hierarchical empowerment to make decisions that affect the group as a whole. In addition, each of the players (the agents) is aware of the group level ultimate goal and strategy, which is also incoherent with the usual idea of self-organization. Typically, agents interact locally and have access to solely local information. That is, they do not have informational access to the group level, as it does happen in football.

In a cooperative social structure, individuals explicitly try to avoid hierarchy, so they coordinate the decision making in order to create a structured group behavior. Therefore, here the organizing is distributed, ideally among all members of the community. But there is still a clear intention to create a particular kind of group behavior, and at least some of the members try to have a global view over the community in order to do so. The members of a food coop, for instance, have explicit goals that they decided as a group (e.g. buy to local producers, promote bio products, keep prices reasonable, etc.) and they organize in order to achieve them. So organization occurs not spontaneously, but through the explicit guidelines elaborated by the coop members.

In contrast, the behavior of the cars involved in a traffic jam, which also produces patterns, is not directed in any of these ways and is, thus, a case of self-organization. In traffic, each individual driver has his or her own local goals and interacts solely with the neighboring agents (the cars nearby at each moment) and with the local environment (weather, traffic regulations, etc.). The patterns of behavior that, on a traffic jam, one could watch from a helicopter are not predefined or driven by any particular or group of cars intentionally. They spontaneously form as a result of the sum of all the individual interactions, none of which can even perceive the pattern that is formed, not is driving towards the achievement of the group goal. Rather the contrary is true in this case: each car's individual goal towards the group would be to actually eliminate the traffic jam patterns that it is contributing to create, and change them for a fluid circulation of the vehicles.

1.1.5 Supervenience and Downward Causation

In the context of emergence as self-organization, supervienience is a relation dependence of the emergent phenomena to a causal base. The idea is that whenever emergent properties occur, even though they are not reducible to lower level processes, they must have a series of basal conditions (they don't just randomly appear). The causal link is statistical in the sense that whenever these basal conditions are given again, there are chances for the emergent phenomena to reoccur.

The problem of Downward Causation is central for many authors as it is the battleground for many discussions on whether or not emergent phenomena truly exist ontologically, while others dismiss it as a false or ill-conceived problem. Downward Causation, again in the context of self-organization, is a relation of influence from the emergent phenomena to its causal base: i.e. that the behaviors of the elements or agents that constitute the causal base are affected by the emergent phenomenon they create. The philosophical problem of downward causation is that, on the one hand, this can end up in circular argument, since if downward causation is exerting and influence on the causal base, it is in fact creating itself. On the other hand, if downward causation does not exist, then emergence is in danger of becoming a mere epiphenomenon, since it never affects back the agents that generated it. The advocates and critics of philosophical downward causation are discussed in chapter 2.

1.1.6 Emergence-Relative-to-a-Model

The notion of emergence-relative-to-a-model was introduced by Peter Cariani in the late 1980s and has been developed by the author throughout the years. The origins of the idea were to propose a framework to understand emergence in a way that can be clearly accounted for and, thus, scientifically communicated. It appeared as a response to the discourses that, arguably, reside in a certain degree of vagueness in their definitions, leaving too much to the intuition of the reader in order to understand what is emergent and what is not exactly.

Cariani is concerned with how new functions can appear in systems or devices that perceive and act on their environment. This newness can only be accounted for scientifically if, first, the observer of the system defines the states and state-transitions of the system under observation by creating a model of it. Once this is done, these observations are used to make predictions on the futures states of the system. In this context, emergence occurs whenever unanticipated behaviors, states or functions appear. Within this framework, Cariani identifies two ways in which emergence can occur: combinatoric and creative emergence. In the context of the study of cybernetic systems, the first consists in the appearance of new system functions through new combinations of what the system already operates with, while the second consists in the appearance of new functions trough the introduction of new elements in its calculations.

Emergence-relative-to-a-model is the basis of what is understood in this dissertation as emergence as generation of novelty. While this idea is pervasive along the rest of the thesis, chapter 5 is dedicated exclusively to analyze Cariani’s proposal and influences.

1.2 Case Studies

In this section the scientific case studies of emergence are presented. These are cases typically related to the concept of emergence, which appear as examples in the literature that discussed the term. Like the preceding ideas, they are presented here as a point of reference, and will be used as examples in the rest of the text with the understanding that the reader is familiar to them. When necessary, further analysis on some of them will be elaborated. Chapter 6 introduces another group of case studies, in this case directly related to Artificial Life Art, that are discussed throughout the second part of the thesis.

1.2.1 Slime Mold

The Slime Mold is a classical example of emergent behavior in biology. In fact, slime mold is the name of a group of several species. But in the literature which is of interest here the term refers to one of these species pertaining to the cellular slime mold group: Dictyostelium discoideum. This is an amoeba that exhibits two very different behaviors depending on its environment. When it can find food supplies, it lives as a single-celled organism. But when this is not the case, the individual cells group together and become part of a multi-cellular organism with specialized functions. Then, once the environment allows it, the complex organism disaggregates again and each of its cells comes back to its independent single-celled form.

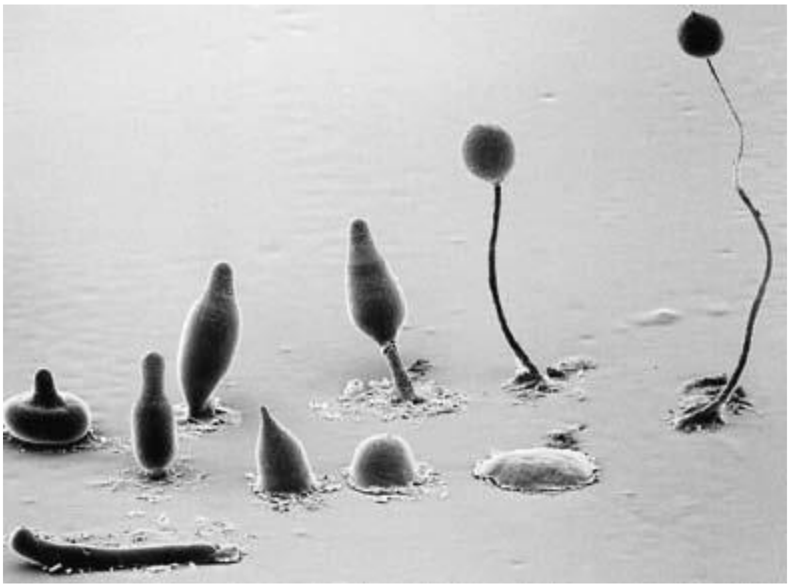


Figure 1.1: The Slime Mold (Keller, 2007: 297).

The first step of the process, aggregation, was puzzling to the biologists. Prior to it, there is no apparent differentiation among cells. Once the multicellular organism has formed, however, the cell differentiation can be accounted for. The usual hypothesis for the aggregation process was that there existed some sort of founder cells which would be in charge of the organizing. But in the late sixties, Evelyn Fox Keller and Lee A. Segel demonstrated an alternative hypothesis: that the aggregation is in fact a process which is done without any central controller at all, and that there are no specialized cells directing the mutation. All of the dictyostelium discoideum cells have the exact same composition, and it is through local interactions that they can successfully group into the slug or fruity body form. In Keller's words, they demonstrated that "organization could emerge from the dynamics for the population as a whole" (Keller, 2007: 301)

Their approach was largely ignored by the rest of biologists for over a decade, in which they insisted, despite the lack of evidence, in the existence of pacemaker cells: "the biologists with whom I talked saw no virtue whatsoever in our account; in fact, they clearly preferred the hypothesis of founder cells, despite the absence of supporting evidence and despite the absence of any explanation of how such specialized cells might themselves have arisen" (Keller, 2003). It was not until the early eighties that the idea of aggregation among homogenous cells began to be widely accepted, as they started to overcome 'the centralized mindset' in understanding how the dictyostelium discoideum behaves (Resnick, 1995).

The slime mold aggregation is a clear example of self-organization and, thus, of this kind of emergence. It is performed by a group of, in this case, very simple agents interacting among each other with no central control or hierarchical structure whatsoever. These local interactions generate a phenomenon at what can be identified as a superior (group) level: that of the slug of fruity body. And this phenomenon, which depends on the local interactions, in turn affects the agents as it gives to some of them a specialized role.

1.2.2 The Belousov-Zhabotinsky Reaction

Chemical clocks are typical examples of order generated in a far-from-equilibrium state. The term refers to a system which is not in thermodynamic equilibrium, nor in the process of being in such a state. The aggregates known as chemical clocks are prominent examples of such cases. They consist in oscillatory chemical reactions, processes which sustain themselves over time through continuous changes.

The most famous example of a chemical clock was discovered in Russia separately in the 1950s by Boris P. Belousov and subsequently studied in the following decade by Anatoly Zhabotinsky. The Belousov-Zhabotinsky reaction is a reaction involving "the oxidation of an organic substance catalyzed by a metal ion in a watery solution" (Kwa, 2002: 41). Usually, such reactions are distributed homogeneously in time and in space. But in this case, instead, "sudden bursts of chemical activity begin at a random place in the petri dish, and they give rise to propagating colored rings across the surface. … The reaction pulsates like a pendulum, a harmonic oscillator, and therefore is compared to a 'clock.'"

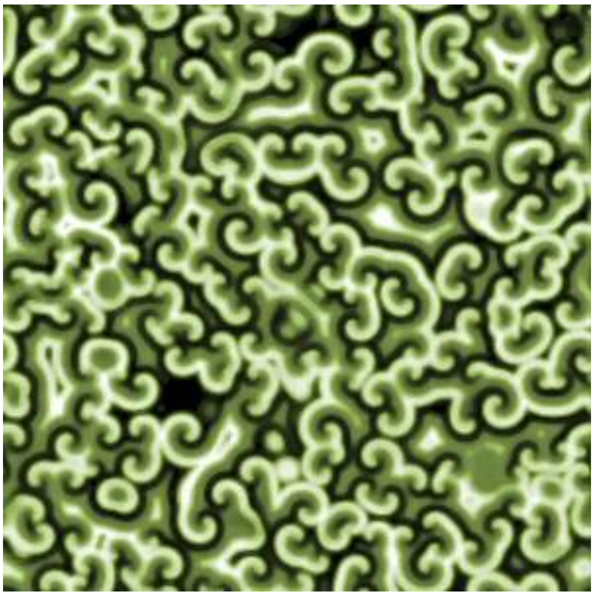


Figure 1.2: A simulation of the Belousov-Zhabotinsky reaction. Author: Chiswick Chap (source: http://upload.wikimedia.org/wikipedia/commons/3/35/Belousov-Zhabotinsky_Reaction_Simulation_Snapshot.jpg).

In 1968, Illya Prigonie proposed the abstract model known as the Brussellator[3], a similar reaction which, although physically unrealistic, was very useful because it could be relatively easily modeled in a computer. On changing its parameters, the Brussellator can go from a steady to an oscillatory state, or even cycle through different oscillations and become "what the meteorologist E. N. Lorenz would call a 'strange attractor,' performing first one oscillation then another" (Kwa, 2002: 41). Chemical clocks are a good example of how systems can exist in which order appears spontaneously, and as such have been consistently used as examples of emergence as self-organization.

1.2.3 The Ant Colony Behavior

The behavior of ants varies among its many species, but there is a common trait that all of them, along with other social insects, share: there is a considerable complexity in the behavior of the ants as a group (the ants colony) which cannot be deduced by the knowledge obtained from studying the specimens in isolation. Ant colonies are complex systems that live much longer than the ants that form them. While individual ants live

---

[3] Brussellator is a contraction of Brussells, for the Free Univerity of Brussells were Prigogine's research was taking place, and oscillator.

for about a year, colonies typically exist for 15 to 20 years. The only exception to this is the queen ant, which will in most species be producing the offspring for all the rest of the ants (Gordon, 2003). But despite a misleading name, the queen ant does not have any hierarchical power on the colony. It is dedicated exclusively to the task of reproduction. Instead, the ant colony is actually, according to what research has been able to show, a clear example of self-organization and, thus, of emergence.

Figure 1.3: Ant Receives Honeydew from Aphid. Author: Dawidi (source: http://en.wikipedia.org/wiki/File:Ant_Receives_Honeydew_from_Aphid.jpg).

Ants respond individually to environment and to interactions with other ants mostly though olfaction (using their antennae). But ants also have been shown, depending on the species, to sense vibration and light (mostly some crude differentiation among dark and bright) (Gordon, 2010). There is no evidence of anything that remotely resembles a central command. Rather, the organization of ant colonies seems to be created solely though these local interactions of individual ants with each other and with their environment: “ants are constantly reacting to each other. Interactions between ants are usually chemical or tactile: an ant responds to the odor of the ants it meets or to a chemical the other ant has emitted, or it responds do the impact or vibration caused by another ant” (Gordon, 2010: 39).

It is through these multiple interactions that ants decide how to perform their tasks, and even in some cases which tasks to perform or whether or not to perform them. Stanford researcher Deborah M. Gordon has used the term ‘task allocation’ to describe this process, in which the ant makes decisions based on its local knowledge, mostly depending on the pattern of interactions with other ants. Though these direct or indirect interactions (like in the case of the very common and well-known pheromone trails) individual ants make decisions and as this repeats for every individual, the behavior of the colony as a group, with all of its separated tasks, is formed.

That said, the exact nature and motive of ant's individual interactions is largely unknown: "an ant's body has about 15 glands (depending on the species), each of which secretes a different substance. We do not know the function of most of these chemicals, nor how much they are used in combination" (Gordon, 2010: 40). To a large extent, this is because ants interact with very small amounts of chemicals, which are not possible to measure with current technology. According to Gordon, this "inhibits the researchers' ability to study the ants' chemical signals in context." As she admits, this makes some experiments, consisting in putting down an extract of a certain gland and watching the ant's responses, rather crude and potentially misleading.

Simulation of ant and termite colony behavior has been performed in robotic research. These simulations are based on a process known as stigmergy, which refers to the indirect communication among termites or ants though the sensing and modification of the local environment. Examples of such experiments can be found in Deneubourg et al. (1990) or Beckers, Holland and Deneubourg (2000) for details.

1.2.4 Conway's Game of Life

Game of Life is the most popular instance of a cellular automaton. It was proposed on 1970 by mathematician John Conway, but it became much more popular some years later, once computer simulations could be easily implemented. A cellular automaton is an abstract model of an informational system. In it, a series of neighboring cells, usually in a one or two-dimensional grid, interact with each other thought a set of simple local rules. Each cell determines its behavior (often a change in its binary on/off state) through analyzing the neighboring cells according to a predefining set of conditions. At every iteration of the system, cells decide whether to change or stay still according to this local information.

Cellular automata are paradigmatic examples of how very simple rules can generate complex structures. In the case of Game of Life, the model for the simulation consists of a two-dimensional grid of cells (which can be of any size) with binary cells that can either be alive of dead. At each iteration of this cellular automata system, the cells on the rectangular grid change from an alive (on) state to a dead (off) state depending on these very simple rules:

- If a cell is alive, will remain alive if it has two or three neighbors, but it will die otherwise.
- If a cell is dead it will come to life if it has exactly three neighbors. Otherwise it remains dead.

Throughout several years of research and simulation[4], Game of Life fascinated not only mathematicians, but also and very especially the Artificial Life community: cellular automata are canonical examples of emergence within the Artificial Life community, since "for some sets of rules, these worlds exhibit surprising complex and propagating behaviors" (Penny, 2010). The ALife researchers saw Game of Life a perfect example of emergence, since these very simple rules were able to create some very surprising and considerably complex behaviors in the system. Some of them were structures that appeared and disappeared, but some even sustained themselves or created others, such as the case of gliders and glider guns.

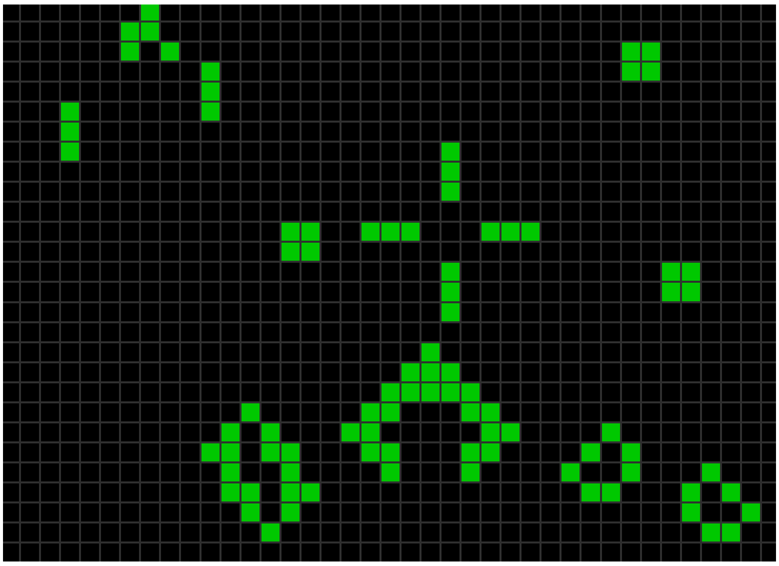


Figure 1.4: A screenshot of Game of Life.

If we understand cellular automata as a heavily abstracted simplification of a system of agents making local decisions, they can be seen as an example of self-organization emergence. What creates the complex and propagating structures is indeed the coordination of these local decisions of the cells interacting with each other. In some cases, too, they can be an example of emergence as generation of novelty. This would be the case of a researcher who, well informed of the rules of the system and aware of what it can do in a simulation, discovers a new structure in it.

---

[4] There are hundreds of implementations of this cellular automata system, both off and online. Some of them are interactive in the sense that they allow the user to pause the simulation and change the state of particular cells. This is the case of an example I created for the software Processing, which since 2013 is included in its examples and is also online at http://processing.org/examples/gameoflife.html.

1.2.5 Craig Reynold’s Flocking Behavior

The word ‘boids’ was popularized (mostly among the A-Life community) by New York University's Craig Reynolds in 1987. It is a contraction of bird-oid, referring to something that resembles a bird or a group of birds. The boids are agents in a system that, while operating strictly through local rules (i.e. referred to individual agents only), are able to create group behaviors that are a clear example of self-organization. In the original version of the boids, the simulation is based in three simple rules that affect each of the agents in their movement through a virtual space:

- **Separation**. When it moves, each agent always tries to maintain a certain distance respect each one of the other group members that are closest to it.
- **Alignment**. In its movement, it will tend to orient itself respect the average orientation of the nearby agents.
- **Cohesion**. At the same time, the movement will also tend to approach the average position of the local group mates.

Each agent does all of this solely in respect to the group mates that are close to it, according to a predefined distance. It is, therefore, an action based only in local information. In any of the simulations the global behavior of the group is amazingly coordinated if one takes into account the agent's such simple (and such local) behavioral rules. When collision detection is added, the simulation of a flock of birds or of fish schooling is nearly perfect.

Figure 1.5: A screenshot of *Stanley and Stella in: Breaking the Ice*¸ 1987 (c) Craig Reynolds.

The boids were not only soon taken as a clear case of emergence, but can even be considered an icon of the a-life art discipline (Penny, 2009). Indeed, this paradigmatic example shows, again, very clearly what is usually understood as emergence: a complex group behavior is achieved through solely the application of simple local rules. It is once more an example that falls into the category of emergence as self-organization.

1.2.6 Tom Ray's Tierra

In late 1989, Tom Ray, an evolutionary biologist, visited the Santa Fe Institute to meet Christopher Langton and James Farmer to discuss the idea of what would eventually be known as Tierra (Levy, 1992).Tom Ray's idea was to build a computer program that would emulate but also accelerate the process of evolution, as compared to the actual biological evolution, with which slowness he was frustrated (Feferman, 1995). This computer program was modeled after the biological systems he was familiar with (Whitelaw, 2006) and simulated a virtual and abstract ecosystem in which creatures reproduced, inherited their parent's characteristics, and suffered mutations in certain cases. In fact, it can be assessed that with Tierra Tom Ray reinvented genetic algorithms, since he implemented its basic ideas in his own way.

Tierra starts up with a series of digital organisms that compete for resources in the computer: computer time as the source of energy, and computer memory as space. These digital organisms' objective is to reproduce. To do so, they have three genes: one to measure themselves, one to allocate space in memory for reproduction, and one that contains its genome. Additionally, each organism has a processor which, when it gets computing time, triggers the process of self-reproduction, thus attempting to realize their goal of replicating themselves. Eventually, however, during the process of reproduction mutation will happen. When this is the case, the genome is misinterpreted in the reproduction and instead of in duplication the process results in the creation of different creatures that can, for instance, become parasites of others. Finally, every once in a while some of the creatures will die, allowing for successive generations to appear.

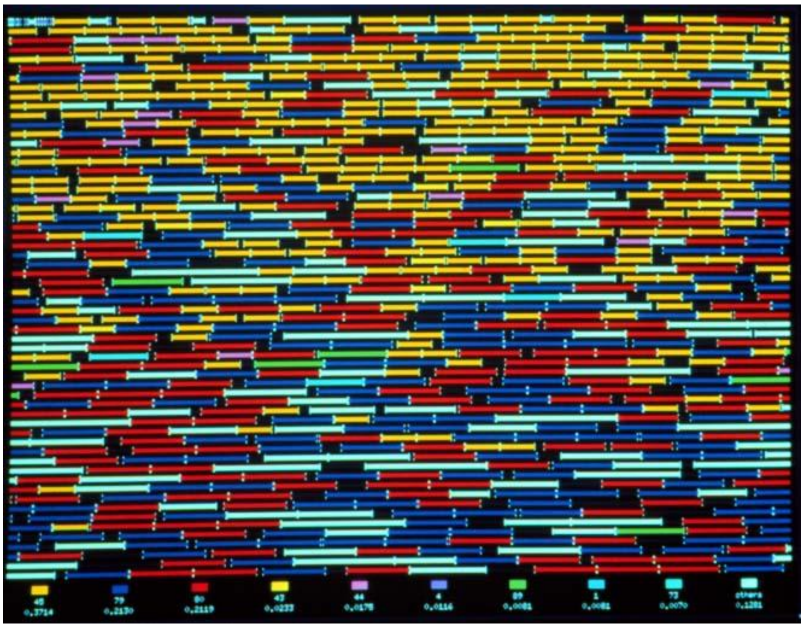


Figure 1.6: Immune hosts are increasing in frequency, separating the parasites into the top of memory (source: http://life.ou.edu/pic/almondc3.jpg).

Ray started up with this simple setup, expecting that he would still have to add a several more features to be able to observe interesting results (Feferman, 1995), but the fact is that very soon he realized that some complex phenomena occurred sooner than he had expected, and differently: hyperparasites formed when a lot of parasites were in the system, and even cooperation among creatures appeared, despite the fact that the initial design was one of competition for computer resources. Specifics aside, Tierra is, again, an example in which, similarly to Game of Life, simple rules generate more complex behaviors, which can be categorized as an abstract form of self-organization emergence, but also as creation of novelty when these behaviors are observed for the first time. Section 6.4 elaborates further on Tierra and its relevance to the study of evolution and to Artificial Life.

1.2.7 Counterexample: The Butterfly Effect

‘Does the flap of a butterfly’s wings in Brazil set off a tornado in Texas?’ This was the title of Edward Lorentz’s talk at the 139th meeting of the American Association for the Advancement of Science, which very graphically describes the principle of Sensitive Dependence on Initial Conditions, also known –precisely by this definition– as the Butterfly Effect. A more formal version of the question, given by Lorenz himself in the same talk, would be: “Is the behavior of the atmosphere unstable with respect to perturbation of small amplitude?” (Lorenz, 1972). The main idea behind this is that, in some systems, there is an inherent degree of uncertainty that cannot be avoided: “simple deterministic systems with only a few elements can generate random behavior. The randomness is fundamental; gathering more information does not make it go away.

Randomness generated in this way has come to be called chaos” (Crutchfield et al, 1986).

Lorentz was studying the weather when he realized that very little variances in the measures utilized to model future events drove the models to drastically different outcomes. He found out that rounding his numbers to three decimal places, he would obtain results that were unpredictability different of those of using the same numbers with six decimal places. These counterintuitive results were the basis for the Sensitive Dependence on Initial Conditions theory (Penny, 2009). The general idea is that a small event can trigger, through a long series of causalities, another event that is perceived as being completely unrelated. That is, that there are certain systems which have an intrinsic degree of indeterminacy, which makes long term predictions impossible. Added to this, any error or lack of accuracy in the initial measurements or calculations can add up to a very distorted prediction.

Sensitive Dependence on Initial Conditions, however, and despite this indeterminacy, is not a case of emergence; the butterfly effect is in fact an epiphenomenon. In respect to self-organization, it is so because there is not necessarily a connection between what results from the phenomenon and what caused it, in the sense that the tornado in Texas would not affect in any perceivable way the behavior of the butterfly whose wing flapped at the beginning of the process. Simple aggregation of facts cannot account for self-organization. But it is also epiphenomenal in respect to emergence as generation of novelty. This is so because, once observed the phenomenon it created (the tornado in Texas), there is no base whatsoever to think that the event that triggered it (the flap of the butterfly wings) is likely to create it again (not any more than any other phenomena).

1.3 Concluding Note

As said above, this chapter offers only a small selection of concepts and examples that are presented here as an introduction to the discussion on emergence. Along the following chapters each of the introduced concepts appears and is analyzed in several contexts and though the views of different authors. The presented case studies are used as recurrent examples, and eventually the information on some of them will also be amplified.

2. EMERGENCE IN PHILOSOPHY

The first uses of the term emergence in philosophy date from the late nineteenth century, when George Henry Lewes coined it in 1875. Lewes used the concept in a context previously inaugurated by John Stuart Mill in discussing different types of causation. The previous philosophical context had allowed little room, if any, for such an idea. Emergence didn't have an important role in classic Greek philosophy, although some formulations have been attributed to Greek Philosophers such as Heraclitus, Anaxagoras, Aristotle, or even to Galen. Arguably, it is a concept that would find a better fit among oriental traditions, such as Taoism or Buddhism, but it is beyond the scope of this PhD proposal to pursue this.

In the western tradition, this conceptual inadequacy was perhaps most clearly the case within the modern Cartesian paradigm, and even within its empirical counterpart, since both rationalism and empiricism were strongly committed to reductionism. In this context emergence was not a welcome concept, as it is by definition inconsistent with a science that aims to reduce all phenomena to simple facts and laws. Modern philosophy's reductionism was axiomatic. All theory could be theoretically traced back to a series of axioms and propositions, from which all theorems and particular cases would be explained. In this respect, Euclid's Elements was a model of how sciences should be done, and the clock, a machine made up of perfectly combined and synchronized gears and mechanisms, became a perfect metaphor of how the world worked.

After Mill and Lewes, the concept of emergence took for the first time a central role in some of the discourses debating evolution at the beginning of the twentieth century. Prominently, and with direct connections to Mill's theories, among the members of the movement known as British Emergentism[5], which defended that reality is organized in different levels of complexity, and that superior levels are not reducible to inferior ones. That is, they are not explainable as resultants of the levels below but, rather, they are emergent in the sense that Mill and Lewes formulated the idea. British Emergentism flourished before the turn of the twentieth century and vanished with the event of quantum mechanics. As will be explained below, this was a little more than a coincidence in time, as the new theories partially influenced their decay. In parallel, vitalism, through ideas such as Bergson's 'Élan Vital,' would hold that there is something fundamentally different between living and non-living systems: fundamental forces or impulses that drive the appearance and the evolution of beings. Although the vitalists did not use the term emergence, they understood these forces to be emergent in the sense that these forces, too, were not reducible to the lower levels of reality

---

[5] The term was inaugurated by Brian P. McLaughlin in (McLaughlin, 2008 [1992]).

sustaining them (i.e., matter). The interest in vitalism also diminished as scientific advances were made.

It wasn’t until a few decades later when the new context of Cybernetics and, later, of complexity sciences allowed the concept of emergence to regain importance in the philosophical debate, mostly, in discourses related to the philosophy of science, artificial life and to the physicalist debate regarding reductionism. Within this line of thought, it is prominent the work of Jaegwon Kim, which builds a strong philosophical critique on some of the notions around emergence.

The contemporary discussions are in general articulated around a series of central concepts. Prominently among them are supervenience and downward causation. As discussed in the previous chapter, supervenience is the idea that any phenomena that is to be emergent has to depend on a causal base. The idea is that whenever emergent properties appear, even though they are not reducible to lower level processes, they must have a series of basal conditions (they don’t just randomly appear). Whenever these basal conditions are given again, there are chances for the emergent phenomena to reoccur. Downward causation is a relation of influence from the emergent phenomena to its causal base. It is what differentiates it from an epiphenomenon that randomly appears in a system. To be a truly relevant phenomenon, what is emergently generated must exert some kind of influence to the same elements that, with their individual behaviors, generated it. Resolving the potential circularity problem of combining supervenience and downward causation is one of the key issues in some of the philosophical discussions around emergence.

These philosophical approaches are, for the most part, ontological. That is, they are concerned with determining whether or not emergent phenomena actually exist, i.e. if they are part of reality or just an epistemological construct at best. The debate on downward causation and supervenience is usually rooted around the theoretical efforts on defending or refuting emergent phenomena as truly existent. In contrast, as it was explained in the introduction, this thesis pursues an epistemological approach that is concerned with whether or not an observer can identify a process as emergent (either as self-organization or as generator of novelty).

This chapter presents different historical views on emergence, without claiming to exhaust all the philosophical debates and points of view, and without necessarily relating all of them to the approach that will be defended in the second part of this work. Also, it is due to the differences in context and approach among some of the presented views that the exemplification with case studies –which were chosen with the contemporary discourses in mind– will only be introduced in the second part of the chapter, once the analysis of Jaegwon Kim’s work allows them to fit without forcing the analogies.

2.1 Greek Philosophy

2.1.1 Presocratic Philosophers

In general, the Presocratic philosophers were reductionists in the sense that their main goal was to explain the multiplicity of the world with a set of basic principles or substances. The questions that drove these first philosophers were such as: Can we reduce the apparent confusion of the world to a simple and unitary principle? If so, what would that be and how did multiplicity appear from it? (Bernabé, 1988) Despite this common approach, however, the answers they offered were radically different among them.

In some of the discourses of these early philosophers, latter authors have found ideas reminiscent of emergence. In the Artificial Life literature Heraclitus and Anaxagoras are the first Greek philosophers to be referred to as early predecessors of the concept. Andrew Assad and Norman H. Packard suggest that the origins of emergence might be embedded in the theories of these two Presocratic philosophers: “The idea of emergence could perhaps be traced to Heraclitus and his theory of flux, and Anaxagoras, with his theory of ‘perichoresis’, which held that all discernable structure in the world is a result of a dynamical unmixing process (his version of emergence) that began with a homogeneous chaos” (Assad and Packard, 2008). No further elaboration is offered by the authors after pointing out the idea. However, a close look at the theories of the cited Presocratic, the analogy that they propose seems to be rather weak.

There are two aspects of Heraclitus’ theory of flux which might be linked to modern approaches to emergence. The first is the idea that everything is in constant change. This idea suggests the importance of time in the creation processes, an importance that would be recalled much later in the theorizations of irreversible processes in non-equilibrium thermodynamics. As discussed in section 2.3 below, Henri Bergson articulates a concept of time that is very similar to that of Heraclitus. Both for the Greek and for the French philosophers the importance of time as a process is greater than that of matter. This means that for Bergson, as for Heraclitus, becoming is higher in importance than being. However, while the idea of process is certainly important in theorizing emergence, the fact that process is an important concept for Heraclitus is not a sufficient condition to state that this author’s theories anticipate it. Emergence is an idea that exists in opposition to a certain understanding of causality which had not been elaborated when Heraclitus articulated his discourse. Heraclitus’ idea of change does not challenge the traditional philosophical causation of his time, because such a notion did not yet exist.

In fact, it could be argued that to project emergentism to the Presocratics is an anachronism altogether. As said above, emergentism in philosophy is always a discourse that is elaborated in contrast or dialog to some sort of reductionism. But the Presocratics were actually the first thinkers to propose a kind of rational explanation that was based in this reductionism. Therefore, they could not write in contrast to it. If anything, when Heraclitus, Anaxagoras and other of the earliest philosophers were inaugurating the tradition of philosophy they were contrasting themselves to the tradition of explaining reality through poetry and religion –hence the textbook phrase that explains the beginning of philosophy as the transition 'from myth to logos'– so it is coherent to think that they were rather trying to avoid formulations in which links from one level of reality to another were not clear in terms of causality.

The second possible link between Heraclitus and emergentism is the fact that underlying the constant flux of events there is a series of basic elements which are in combination to create phenomena at the levels of perception. The river remains the river, but what forms it constantly changes. This idea seems to suggest that nature is stratified in different levels of reality, which is one of the main assumptions of most of the philosophical emergentist discourses. But as it was explained above, this stratification of reality would be in this case not an intellectual instrument to present emergentist ideas but to pursue the explanation of reality's diversity in terms of a primary element: fire in Heraclitus.

Finally, Anaxagoras' perichoresis[6] is a slightly more sophisticated concept which elaborates onto Parmenides' idea of Being. Perichoresis is, in short, a theory which states that everything is contained in everything else. Matter is at the same time one only thing and it is infinitely divisible. According to this, nothing ever perishes or comes into being; there's only separation and combination (Anaxagoras, in: Bernabé, 1988). According to Barnes, perichoresis is also a synonymous of revolution, a form of locomotion that makes some things 'come together' (Barnes, 1979: 38). So the term would involve according to him both the fact that everything contains everything else and the movement that makes the changes happen. This is in fact consistent with a theory that states that everything is permanently and since the beginning of time contained into something else (in this primary matter in the case of Anaxagoras). From such an idea itfollows logically that, if any change is ever to occur, the possibility of this change must already be contained in this all-encompassing container.

This notion of change as a combination of what already preexists does resonate with some of the notions of emergence such as Cariani's Combinatoric Emergence[7], in which new processes are created (set in motion) involving only what is already preexisting in a system. However, when Cariani articulates the notion of Combinatoric Emergence, he is thinking in combination as a source of novelty. In contrast, Anaxagoras' perichoresis is contradictory with the fact that emergence is concerned with novelty in any relevant

---

[6] Perichoresis is better known for its use in Christian Theology, along with synonymous notions such as coinherence, in order to explain the unity of the three elements that constitute the Holy Trinity (Sierra Díaz, 2003).

[7] See Chapter 5.

manner. If everything exists in everything else, nothing ever comes into being. Thus true change, and let alone true novelty, cannot find its place.

It is likely that the resemblance to emergence that Assad and Packard found in Anaxagoras' thought was rooted in the fact that perichoresis implies that everything is self-contained, the force for change included. This can be reminiscent of emergence as self-organization in the sense that in neither case any external force must intervene for a phenomenon to occur. Nevertheless, as stated with respect to Heraclitus and reductionism, the philosophical context in which Anaxagoras writes is far from dealing with the concerns of complexity theory. The presocratics were concerned with the perceivable differentiation of matter and forms and how these could be understood intellectually and scientifically.

2.1.2 Aristotle and Galen

Aristotle is also mentioned occasionally as a possible predecessor of emergence (Corning, 2002; Alsina 2009). According to Corning, Aristotle formulated essentially the same argument that John Stuart Mill would make two thousand years later. In the Metaphysics passage cited by him, Aristotle stated that a whole is something different than a simple addition of parts: “The totality is not, as it were, a mere heap, but the whole is something beside the parts” (Aristotle, 1994: Book H, 1045:8-10). It is certainly a formulation that connects directly with the more popular idea of emergence, that of the whole being more than the sum of its parts. It connects too with the main idea of Anderson's seminal 1972 paper entitled *More Is Different* (Anderson, 1972). But Aristotle resolves the problem in a different way than Anderson and his contemporaries. As the Metaphysics Book H advances, he uses the classic Aristotelian distinctions of form and matter and of potentiality and actuality. With these, he argues, “the difficulty disappears” (Aristotle, 1994: Book H, 1045:25). Thus, again, the analogy does not hold a detailed scrutiny beyond the pointing out of the intuitive idea behind emergence: the inspiring whole being more than the sum of the parts.

Much more elaborated is Victor Caston's claim that a latter Greek figure, Galen, was himself an emergentist (Caston 1997). This Greek philosopher and physician anticipates some of Mill's ideas, focusing, as Caston notes, on the case of sensation. The basic idea is that perception is emergent respect the organs that generate it. According to Galen there are two types of composed entities. First, there are those that are simple aggregates of simpler elements (the resultants in Mill's and Lewes' terms). Then, there are those complex wholes were a novel characteristic appears which is of a different type of all of which they are compounded of (the emergent in Mill's and Lewes' terms): "to speak generally, the type to which sensation belongs ‘differs from all the other properties’ that belong to bodies" (Galen, cited in: Caston, 1997). Thus, according to Galen, and although he does not utilize this terminology, it follows that, since none of the elements underlying perception has the ability to perceive for itself, sensation is emergent.

Caston notes that Galen anticipates emergentism in a much more detailed way than in the cases of Heraclitus, Anaximander or Aristotle: “[He offers] a clear notion of mental powers that do not belong to the elements, but arise only at certain levels of complexity. These powers depend nomically on the constituents in the blend –they supervene on them– but they are not simple aggregates of the powers at this lower level” (Caston, 1997). Remarkably, Galen’s choice of examples in elaborating his theory –sensation– turned out to be a much better one than that of Mill and the other British Emergentists, at least regarding some approaches to contemporary philosophical theories. While advances in Physics and Biology in the early twentieth century invalidated the examples of the British theoreticians, Galen’s examples are still valid in some approaches to contemporary philosophy, like in the case of Jaegwon Kim (review below) who states that, if anything is to be emergent, qualia would be the best candidate[8] (Kim, 1999).

Therefore, while in the first examples (Heraclitus, Anaxagoras and Aristotle) aspects of the contemporary formulations of emergence are elaborated, arguably in none of them a link to contemporary formulations can be clearly established. According to Caston, Galen does articulate a discourse that anticipates a coherent theory of emergent phenomena, although as said he does not use the modern or any other equivalent terminology to describe them.

---

[8] Thus, this argument is correct provided one subscribes to a theory that accepts qualia as existent. For certain theories of cognition, such as enactivism, qualia do not exist.

2.2 British Emergentism

As said above, the debate around emergence didn’t really encounter a prolific context until Mill and his followers created the discourse that has become known as British Emergentism, after Brian P. McLaughlin coined the term in 1992. It is a philosophical and scientific movement inaugurated by John Stuart Mill in the 1840s. As a clearly distinct current, British Emergentism lasted until the 1920s, when some of its central arguments were rendered invalid by the new scientific theories of the twentieth century (McLaughlin, 2008). One of the reasons for this decline, according to philosopher Jaegwon Kim, was that most of the examples of emergence that the British Emergentists chose in order to support their theories, mostly chemical and biological properties, were annulled as such by the scientific explanations –articulated in reductionist terms– of the twentieth century solid-state physics and molecular biology (Kim, 1999).

The emergence of British Emergentism is, like in other discourses, an idea used to explain the relation between phenomena that depend on each other, but for which the causal links cannot be fully accounted for. These explanations are often strongly rooted in an interpretation of reality as being composed of different layers of complexity. These levels, and most of the relations among them, are perfectly coherent internally in terms of causation and are explained by different scientific disciplines, namely: physics, chemistry, biology and psychology. Their idea of emergence refers more often than not to the impossibility of some of the events happening at one of the layers to be explained entirely in terms of –i.e. reduced to– the lower level discipline.

The main works of the movement can be traced from Mill’s inaugural *A System of Logic* (1843) to Alexander Broad’s *The Mind and Its Place in Nature* (1925), but even contemporarily a number of authors are still connected to it, such as neurophysiologist Roger Sperry (McLaughlin, 2008) or philosopher Jaegwon Kim (Campbell and Bickhard, 2001). Mainly, these relations to contemporary academics are established in the context of nonreductive physicalism. The term physicalism is in this context of contemporary philosophy essentially interchangeable with materialism.

2.2.1John Stuart Mill

In *A System of Logic*, John Stuart Mill, in discussing different kinds of causation (and how causes accumulate), distinguishes among two kinds of laws and effects, which would relate to mechanical and chemical processes respectively: those resultant from mere aggregation of causes (e.g. the sum of two force vectors producing a resulting aggregate vector) and what he labeled heteropathic effects. In what would be the equivalent of the latter distinction between resultant (aggregate) and emergent properties, Mill establishes two different ways in which causes can accumulate in order to create new effects. First, there’s simple aggregation. In what he calls the ‘mechanical mode', different forces, calculated as vectors, add up to produce a single resultant force. This is, Mill says, the basic dynamics principle of Composition of Forces. And in analogy to it, he proposes the principle of Composition of Causes "which is exemplified in all cases in which the joint effect of several causes is identical with the sum of their separate effects" (Mill, 1882: 459). That is, whenever a complex effect is fully understood by examining (and aggregating) any number of causes underlying it, we can consider that the principle applies.

In contrast, there are the cases in which the Composition of Causes is not that easy to trace. In Mill's words, there is a 'breach', a gap in the continuity of causes: "there are laws which, like those of chemistry and physiology, owe their existence to a breach of the principle of Composition of Causes" (Mill, 1882: 462). These are the heteropathic laws and effects (which would later come to be called emergence).

These heteropathic laws do not entirely supersede the laws at lower levels. The new laws and effects are added to the previous ones, creating a new layer of complexity, while laws at lower levels continue to operate and can even interact with the newly created ones: "the component parts of a vegetable or animal substance do not lose their mechanical and chemical properties as separate agents, when, by a peculiar mode of juxtaposition, they, as an aggregate whole, acquire physiological or vital properties in addition. Those bodies continue, as before, to obey mechanical and chemical laws, in so far as the operation of those laws is not counteracted by the new laws which govern them as organized beings. … The new laws, while they supersede one portion of the previous laws, may co-exist with another portion, and may even compound the effect of those previous laws with their own" (Mill, 1882: 462).

As mentioned above, despite being the first modern author to theorize about it, Mill never used the term 'emergence' in his writings. It was George Henry Lewes, in 1875, who coined it in discussing Mill's ideas. Lewes distinguished between resultant effects, those that would be explained by the principle of Composition of Causes, and emergent effects, which would be the equivalent to Mill's heteropathic effects (McLaughlin, 2008).

2.2.2 Alexander, Morgan and Broad

During the 1920s, three more authors took on Mill's ideas and Lewes' terminology and developed some original ideas, in the context of evolutionary emergentism: Samuel Alexander, C. Lloyd Morgan and C. D. Broad. Alexander and Morgan's notions of emergence are closely related (McLaughlin, 2008). The first introduced the notion of a separation of levels of explanation in terms of complex configuration of elements. According to him, the superior levels that these complexes represent possess characteristics that cannot be understood under Mill's principle of the composition of causes, in regards to the causality of the elements in the lower levels. These are emergent characteristics that, according to Alexander, are truly unexplainable in deductive terms. There is no connection to be found in this respect between what causes these phenomena and the phenomena themselves. These can only be empirically observed and have to be accepted as unexplainable facts 'with natural piety' by the scientific mind: "The existence of emergent qualities thus described is something to be noted, as some would say, under the compulsion of brute empirical fact, or, as I should prefer to say in less harsh terms, to be accepted with the 'natural piety' of the investigator. It admits no explanation" (Morgan, cited in: McLaughlin, 2008: 31).

Morgan wrote in the context of emergent evolution (this is indeed the title of his main book) in contraposition with both vitalism and mechanism. The mechanist view was that everything is reducible to physics. That is, all sciences are particular cases of the physics as the ultimate science. Reality is defined, ultimately, through the laws of physics, and that includes living beings and systems. Thus, all the levels that, in respect to physics, for the British Emergentists were superior and not entirely explainable to the laws of the inferior ones, were for mechanists unproblematically reducible to the fundamental discipline of physics. On the other extreme in respect to mechanism was vitalism, with its claim that there are a series of substances, forces or entities which are superior to the laws of physics (and thus irreducible to them), such as the 'Élan Vital'[9]. Among theses, the emergentist view was an in-between position of sorts. It did not deny the existence of an ultimate physics substrate of reality (the elementary particles). Rather the contrary, these elementary particles were often an important element in basing these theories. But it too accepted that matter could be organized in levels of complexity that had properties inherently specific to them, and thus irreducible to the inferior levels.

Strictly in the context of evolutionary theory, emergent evolution was a means of explaining the continuity gaps in the history of evolution that the Darwinian theory found very difficult to explain[10] (Alsina, 2009). McLaughlin states that "Morgan's chief concern is to argue that, through a process of evolution, new, unpredictable complex phenomena emerge" (McLaughlin, 2008). According to him, this view is explained in opposition to the mechanistic view, which operates in terms of resultants only, ignoring "the something more that must be accepted as emergent" (Morgan, cited in: McLaughlin, 2008: 33).

Within this discussion, we find, as noted by McLaughlin, perhaps the most remarkable trait in Morgan's literature on emergence from a modern point of view: the notion of

---

[9] See section 2.3 below.

[10] Some of these gaps have been explained by contemporary science. E.g. new research in Epigenetics is revealing transmission of inherited qualities by environmental causes, impliying that there exist mutations that are inherited. Also, the concept of punctuated evolution (see chapter 6, section 4) has helped to account for some of these possible evolutionary inconsistences.

Downward Causation. Borrowing Alexander's request for natural piety, Morgan inaugurates an idea that would become a classic focus of discussion in the discourses on emergence. He explains how life, which is, according to him, emergent in respect to the living organism, affects in turn this very same organism that generates it: "I accept with natural piety the evidence that there is more in the events that occur in the living organism than can adequately be interpreted in terms of physics and chemistry, though physico-chemical events are always involved. Changes occur in the organism when vital relatedness is present the like of which do not occur when life is absent. This relatedness is therefore effective" (Morgan, 1927: 20-21).

Finally, Broad represents, on McLaughlin's account, the best attempt to systematize the British Emergentist's view. This view is presented mainly in opposition to what he labels 'mechanism', which, similarly to Morgan, would be the reductionist approach that states that everything would be reducible to a certain number of fundamental particles, properties and forces. In contrast to this view, Broad elaborates a theory of emergentism that, although accepting some of the axioms of mechanism such as the existence of elementary particles, radically differs from it in accepting that "wholes can possess force-generating properties of a sort not possessed by any of its parts" (McLaughlin, 2008: 41).

Broad understands reality as a hierarchy that can be expressed with the different orders of knowledge, starting with physics in the lowest level, and ascending to chemistry, biology and psychology. He accepts reductionism and causation inside each level of science. These levels, then, have intra-ordinal laws which work in the classical sense. But when it comes to examining the different levels not-in-isolation, one needs what he calls trans-ordinal laws; a type of laws that connect elements in adjacent orders and which 'can' be emergent: "A and B would be adjacent, and in ascending order, if every aggregate of order B is composed of aggregates of order A, and if it has certain properties which no aggregate of order A possess and which cannot be deduced from the A-properties and the structure of the B-complex by any law of composition which has manifested itself at lower-levels" (Broad, cited in: McLaughlin, 2008: 42).

The British Emergentism discourse, as said above, declined with the advent of quantum mechanics and other advances of early twentieth century science, which demonstrated that some of its fundamental ideas were invalid or very implausible. According to McLaughlin, this means that what led the to movement's fall weren't philosophical but scientific advances. In any case, they did drive these theories to a dead end nonetheless. One of his examples suffices to understand this point: "quantum mechanical explanations of chemical bonding suffice to refute central aspects of Broad's Chemical Emergentism: Chemical bonding can be explained by properties of electrons, and there are no fundamental chemical forces" (McLaughlin, 2008: 49].

Jaegwon Kim, elaborating on McLaughlin's point, states that the unfortunate choice of the British Emergentist's examples was understandable regarding the state of the sciences at their time. This, he insists, does not invalidate their philosophical interest: “the interest of the ideas underlying the emergentist’s distinction between the two kinds of properties [i.e. resultant and emergent] need not be diminished by the choice of wrong examples” (Kim, 1999). Along the same lines, Campbell and Bickhard have stated that British emergentism “fails to be naturalistic, not to mention that it is ad hoc and unreal,” but nonetheless “other models of genuine emergence are possible” (Campbell and Bickhard, 2001).

2.3 Henri Bergson

In 1907, the French philosopher Henri Bergson published *L'Évolution créatrice* (Creative Evolution), a book in which he articulates his vision on evolution as opposed to the mechanist view. In it, he also develops his discourse on the concept of time and its role in the processes of biology and psychology. Like the emergentists, Bergson wrote in contraposition to the dominant reductionist paradigm in science and philosophy, which advocated for mechanism and materialism. In a context in which concepts such as growth, change or creativity were an uneasy fit at best, "Bergson's commitment to the reality of time as a source of creative change enabled him to clarify many problems in psychology and biology that appeared contradictory from existing scientific and philosophical perspectives, and to provide a rigorous account of a creative evolution, the creative mind, and the nature of their relation" (Vaughan, 2007).

Although his much cited concept of 'Élan Vital' might suggests that he was a vitalist, his position is probably more accurately defined as being in between vitalism and emergentism. Bergson placed himself close to teleology and vitalism regarding evolution, in contraposition to mechanistic biology (Vaughan, 2007). The emergentism as a movement in evolutionary theory had not appeared by then as a third option. Nonetheless, it is doubtful that he would have found his theory to be particularly close to that of the British Emergentists discussed above. In this respect, it is noteworthy than Bergson does not once use the term 'emergence' in *Creative Evolution,* nor does he use an equivalent term as does Mill. But many of the ways in which he describes the process of evolution, and how he presents it as the fundamental process in which creativity and newness appears, do resonate with the concept of emergence as generation of novelty.

In Bergson, time is the essential notion in this process of creation. Reality is not about isolated facts in temporal terms. Borrowing John Protevi's terms[11], it is about diachronic, not synchronic processes: "The more we study the nature of time, the more we shall comprehend that duration means invention, the creation of forms, the continual elaboration of the absolutely new" (Bergson, 1911: 11).

Bergson argues that the mechanist reductionist view is wrong in considering that the complexity of life can be understood by analyzing smaller and smaller parts of beings and organisms. There are different orders of elements, and simple processes at superior levels can arise from very complex assemblages of elements at inferior ones, in a process, he argues, that is wrongfully regarded in the cited disciplines from an anthropomorphic view. According to Bergson, mechanism mistakenly understands organization as manufacturing. Manufacturing, he says, goes from the many (the parts) to the one (the whole). The parts are assembled so that a certain whole is obtained,

---

[11] See section 2.6 below.

which in turn will generate a certain function. In contraposition to that, Bergson proposes the idea of organizing, which is the contrary of that. Organizing goes from the center to the periphery: “It begins in a point that is almost a mathematical point, and spreads around this point by concentric waves which go on enlarging” (Bergson, 1911: 92).

This is essentially the same claim that Richard Dawkins would make in his popular 1986 book *The Blind Watchmaker*: Natural selection is the organizational force, there is no manufacturing, no central mind, controlling the process of creation of new species and new functions within them: “Natural selection, the blind, unconscious, automatic process which Darwin discovered, and which we now know is the explanation for the existence and apparent purposeful form of all life, has no purpose in mind. It has no mind and no mind's eye. It does not plan for the future. It has no vision, no foresight, no sight at all. If it can be said to play the role of watchmaker in nature, it is the ‘blind’ watchmaker” (Dawkins, 2006: 5). Despite the coincidence, however, it would be wrong to present here Dawkins as an apologist of Bergson’s ideas. They differ in many other issues, but these discussions are beyond the scope of the work presented here.

Nevertheless, Dawkins is indeed very relevant regarding Artificial Life, which will be examined in detail in the subsequent chapters. His 1986 book (and the software that accompanied the first editions) was strongly influential among the first members of the Artificial Life Art community, which articulated a discourse in defense of bottom-up processes in computation as opposite of the wider used top-down approaches (such as those used in Artificial Intelligence). It is an idea that contains a clear parallelism with the Bergsonian opposition between organizing and manufacturing.

Turning back to Bergson, we find that organization is the process from which beings and their functions are created. Contrary to the reductionist approach of mechanism, Bergson claims that “life does not proceed by the association and addition of elements, but by dissociation and division” (Bergson, 1911: 89). Here is where the ‘Élan Vital’ (translated as ‘original impetus’ in the English edition, although the original French is almost always used) finds its main context: it is the force behind this organization: the force behind all the variations that causes the creation of new species and their functions. The counterforce of this impulse is matter, which tends to stay the same. Thus, the ‘Élan Vital’ is the only possible source of novelty in this bergsonian context. It is what “strives to introduce into [matter] the largest possible amount of indetermination and liberty” (Bergson, 1911: 251)

As this process of creation unfolds, two important notions acquire importance: time and contingency. The first is, as said above, a central concept in the theory, but the latter is also an important concept in his understanding of evolution. When these indeterminations are introduced into matter, the process of adaptation occurs. And this process, says Bergson, is largely contingent. “Two things only are necessary: (1) a gradual accumulation of energy; (2) an elastic canalization of this energy in variable and indeterminable directions, at the end of which are free acts" (Bergson, 1911: 255).

As said, along with contingency, time plays the role of allowing these processes to occur. Bergson explains his vision in contraposition with what he calls the 'cinematograph method', which splits reality in single moments to be analyzed[12]. This method, which consists in substituting the objects themselves by signs to be handled, belongs both to ancient and modern sciences: "ancient science thinks it knows its object sufficiently when it has noted of it some privileged moments whereas modern science considers the object at any moment whatever" (Bergson, 1911: 330). In contrast, Bergson values the becoming more than the being at a certain point in time. In terms that resonate with Heraclitus, he understands that it is the process what should be the main concern of knowledge: "he who installs himself in becoming sees in duration the very life of things, the fundamental reality" (Bergson, 1911: 317). It is in this context that we can understand his strong claim: "Time is invention or it is nothing at all" (Bergson, 1911: 341). Time is what channels change, and this is the main way in which it shall be understood.

In conclusion, there are three concepts regarding which the links between Bergson's view on emergence are clearly traced: creativity, contingency and time. As stated in the introduction, emergence is often related to fundamental novelty and, thus, to creativity. It is a concept that aims to explain how essentially new properties, processes, etc. can appear in a context in which they are not present. And in any of the many accounts of emergence, these processes happen trough a certain period of time (that is, they cannot be scrutinized with the 'cinematograph method') and carry with them a certain degree of contingency. The processes and conditions that generate emergent phenomena always carry, in one way or another, a certain degree of indeterminacy.

Bergson would not be properly labeled an emergentist. Neither is he a clear influence to later philosophers who have elaborated on the concepts, such as the ones following in this section. But there is no doubt that his general influence as a major figure of the twentieth century philosophy can't be disregarded in the context of the study of a concept with such connections to his work.

---

[12] This bergsonian idea was captured by Flann Obrien in one of the writings of the fictitious savant de Selby (Simon Penny, personal communication, 2013-12-10): "Human existence de Selby has defined as 'a succession of static experiences each infinitely brief', a conception which he is thought to have arrived at from examining some old cinematograph films. … Apparently he had examined them patiently picture by picture and imagined that they would be screened in the same way, failing at that time to grasp the principle of the cinematograph" (O'brien, 1967).

2.4 Jaegwon Kim

The philosopher Jaegwon Kim has written some of the most influential literature on the topic of emergence. He does not advocate for it, rather, his examination of the main issues of the emergentist discourse leave little room to the possibility of any emergent phenomena. In respect to the authors reviewed up to this point, Kim widens the context of the discussion. While not abandoning the more general philosophy of science point of view, he also –and mainly– focuses on emergence as related to mental phenomena. According to Campbell and Bickhard (2001), Kim writes in the context of physicalism, a doctrine which states that all mental phenomena have a correlation with physical states. Whether or not these states can be reduced to their physical equivalents is the debate of non-reductive vs. reductive physicalism, in which Kim would fall closer to the latter.

The conclusion of Kim (2006) is a good summary of his position on Emergence in his most influential writings: "The idea of emergence is an attractive, and initially appealing, one in many ways, and it is not difficult to understand its popularity. But it is not easy to make the idea precise and give it substantive content. Two important unresolved items remain on the emergentists' agenda. The first is to give emergence a robust positive characterization that goes beyond supervenience and irreducibility. The second is to come face to face with the problem of downward causation. Somehow the emergentist must devise an intelligible and consistent account of how emergent properties can have distinctive causal powers of their own —in particular, powers to influence events and processes at the basal level."

In discussing emergence, Kim identifies two main ideas to be thoroughly analyzed: supervenience and downward causation. The first is derived from the notion that "emergent properties are 'novel' and 'unpredictable' from knowledge of their lower-level bases, and they are not 'explainable' or 'mechanistically reducible' in terms of their underlying properties" (Kim, 1999). The second, downward causation, is according to Kim the most prominent emergentist doctrine regarding the causal powers of these emergent phenomena. It is, he says, what can make a difference between emergent phenomena being either mere epiphenomena or relevant in causal terms.

Like most of the philosophical analyses, Kim's is elaborated from the ontological point of view. That is, he is concerned with whether or not emergent phenomena are a part of reality as he understands it. As a physicalist, a position which implies that everything must be reducible to physical phenomena, it is natural to find emergence at least controversial. This does not mean that physicalism and emergence are contradictory. Emergence is a concept that lives in opposition to reductionism, but not to causality in general. And physicalism, as reviewed below, can either be reductive or non-reductive. Therefore, it is only in reductive physicalism were emergence is unlikely to be a welcome idea. In the non-reductive version, the issue is more open to discussion.

However, if the position is of trying to explain emergence as an ontological reality, the task is arduous nonetheless.

In general, ontological approaches to emergence like those of some of the reviewed authors in this chapter make any discussion with concepts of emergence understood epistemologically a little difficult to articulate. However, while keeping these fundamental differences in mind, it is worth trying to apply his analysis to the ideas of emergence as self-organization and emergence as generation of novelty and some of the case studies.

2.4.1Supervenience and Irreducibility

For Kim, supervenience, which is synonymous with upward determination, is an essential component of emergence. Upward determination is essentially the contrary to downward causation. I.e. the causal influence from the base to the phenomena on the superior level. It is, for Kim, what allows emergence not to be just random phenomena. He defines supervenience as follows: "If property M emerges from properties N1, … Nn, then M supervenes on N1, … N. That is to say, systems that are alike in respect to basal conditions N1, … Nn must be alike in respect of their emergent properties" (Kim, 2006).

The idea is that whenever emergent properties occur, even though they are not reducible to lower level processes, they must have a series of basal conditions (they don't just randomly appear). Thus, the causal link is statistical in the sense that, whenever these basal conditions are given again, there are chances for the emergent phenomena to reoccur: "I believe there may well be a viable concept of 'statistical' or 'stochastic' emergence, which assigns a stable objective chance of the emergence of a property given that an appropriate basal condition is present" (Kim, 2006). In self-organization emergence this emergent phenomena would be the pattern of behavior that appears (the flocking behavior of the boids, for instance), while the basal conditions would be the local interactions among individual agents (the behavior of each one of the birds in respect to her flock mates). This is the basis for what Kim labels 'inductive predictability': the phenomenon has been observed once, and therefore it can be expected to occur again if the same conditions are given.

Predictability is presented as a key issue after some elaboration on the classical distinction between resultant effects and emergence, which can be explained in terms of additivity and subtractability, according to him. For Kim, this classic distinction is not crucial. Rather, "predictability is the key idea here" (Kim, 1999). Inductive predictability, he says, is what the emergentists can accept, since it is not contradictory to their anti-reductionist positions. But Kim notes that there is another type of predictability that is denied by them: 'theoretical predictability'. This would be the case in which the phenomenon can be deduced through an analysis of the basal conditions before it appears, which is a clear contradiction to the most basic definitions of emergence. Similarly, Mark Bedau, as will be discussed below, explains how the actual simulation of a system set in motion is crucial in order to define emergence. Before simulation, non-emergent events can be predicted through theoretical predictability, whereas it is the simulation –or the passing of time if it is not a formal system– that allows for emergent phenomena to be observed. Once shown, one knows that these emergent phenomena can be expected to appear again if the simulation is run again. This is essentially Kim's inductive predictability.

This idea is problematic when emergence is considered only as generation of novelty and from an epistemological perspective, like in the case of Cariani's theory, discussed in chapter 5. According to the emergence-relative-to-a-model view, once the emergent phenomenon appears it must be incorporated into the model from which the system is being observed and, therefore, the next time it appears will not be labeled as emergent anymore. However, in the ontological approaches the problem of novelty is looked at differently. Kim points out that the basic definitions that he analyses, in one way or another, stress the 'newness' of the emergent phenomena, a newness to which he attributes both an epistemological and a metaphysical sense: "I believe that 'new' as used by the emergentists has two dimensions: an emergent property is new because it is unpredictable, and this is its epistemological sense; and, second, it has a metaphysical sense, namely that an emergent property brings with it new causal powers, powers that did not exist before its emergence" (Kim, 1999).

The metaphysical sense will be examined in the following subsection on Downward Causation. The epistemological sense is that of the more typical idea of emergence: the something more that appears unexpectedly or, at least, surprisingly. Kim explains it with the ideas of supervenience, inductive predictability (both discussed above), and irreducibility. In order to explain the latter, Kim establishes three steps that must exist for a property to be reduced to its base. First, for the property E to be reduced to the base B, E must be functionalized, i.e. it must be given a functional interpretation "by construing it in terms of the causal work it is to perform" upon the properties of the causal base (Kim, 1999). Thus, E must be explainable in terms of its causal influence to its basis: "Functionalization of a property is both necessary and sufficient of reduction." This explains also why non-reducible properties are "neither predictable nor explainable on the basis of the underlying processes." The question, for Kim, becomes whether or not there are any properties of this sort. The second and third conditions are that the causes of E are clearly identified in B and theoretically explained. In Kim's words, that both the realizers of E and a theory to explain how they perform the causal task that is constitutes it are found at the level of B.

All these arguments on supervenience and reductionism allow Kim to explain how emergence is presented by its advocates: "emergent properties are 'not predictable' from their basal conditions, that they are 'not explainable' in terms of them, and that they are 'not reducible' to them." His conclusion is that it is doubtful that emergent properties exist at all, and if they do, it is in very particular cases: “If the emergentists were right about anything, they were probably right about the phenomenal properties of conscious experience: these properties appear not to be theoretically predictable on the basis of a complete knowledge of the neurophysiology of the brain.” More specifically, he later refers to qualia as the best candidates: “If anything is going to be emergent, the phenomenal properties of consciousness, or ‘qualia’ are the most promising candidates.” And admitting a bias towards the pro vs. the con: “Qualia are intrinsic properties if anything is, and to functionalize them is to eliminate them as intrinsic properties” (Kim, 1999).

2.4.2 Downward Causation

Prominent among Kim’s list of the central doctrines of emergentism, downward causation is the one to which he dedicates the most attention. Within emergentism, downward causation is the idea that explains how the emergent properties can have causal powers of their own, which is in fact, according to Kim, what differentiates them from mere epiphenomena (figures 2.1 and 2.2): “The principle of downward causation directly implies that if emergent properties have no downward causal powers, they can have no causal powers at all, and this means that emergent phenomena would just turn out to be epiphenomena” (Kim, 1999).

Downward causation is for Kim the most important issue in the emergentist discussion, a key idea that needs to be explained in order for emergentism in its different forms to be a valuable scientific theory: “It will become clear, if I am right, that non-reductive physicalism is a form of emergentism, and that both positions stand or fall with downward causation” (Kim, 1992). The idea of downward causation, according to Kim, implies a layered model of reality. That is, that reality is organized in different levels and that one can refer to each of these layers as high or low depending on were the focus is placed at each time. Within this framework, the lower levels are where the basal conditions for both resultant and emergent phenomena are located. The resultant and emergent phenomena are found, then, at the level above, the higher level in respect to the one were local interactions are analyzed. When emergent phenomena are identified according to this approach, downward causation is the exertion of some kind of causal effect to the constituting elements of the basal level by them.

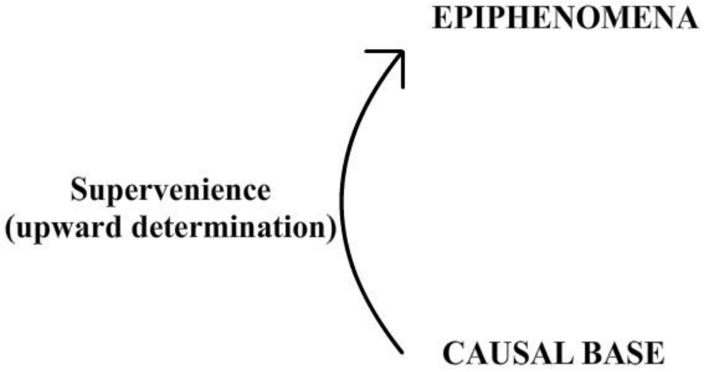


Figure 2.1: According to Kim, if there is no downward causation, supervenience results in mere epiphenomena.

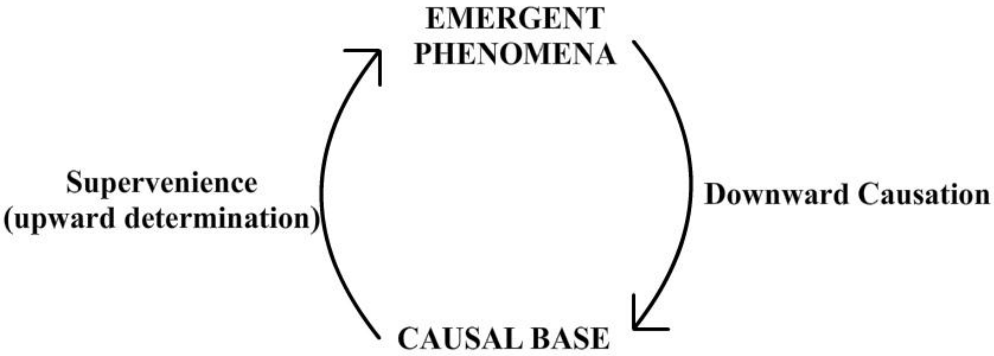


Figure 2.2: Supervenience and Downward causation affect, respectively, the emergent phenomena and the causal base.

Kim notes that there is a very simple and unproblematic formulation of downward causation. It is the idea that microlevel properties in complex systems can change in virtue of their marcrolevel properties. Simply put, that the parts will for instance move when the whole moves. If a person goes to the library, all of its constituents go to the library with him. This is downward causation. But of course this is not a notion worthy of much philosophical discussion. The big questions arise when this is combined with upward determination regarding the same property that affected the lower levels.

This meaningfully influential phenomenon is what Kim calls reflexive downward causation, a concept in which he identifies some problematic circularity of arguments: "Some activity or event involving a whole 'W' is a cause of, or has a causal influence on, the events involving its own microconstituents" (Kim, 1999). Thus, the problem here is that, if the emergent phenomenon exerts a causal influence on the elements at the causal base, then it could be its own cause of existence. The argumentation on this is outlined below.

The circularity problem can exemplify this in terms of the ant colony. The overall organization of the ants is caused by the local interactions of the individual animals. However, if downward causation happens, then the overall organization affects decisively the behavior and decision of the individual ants, which in turn affects the overall behavior, and so on. If the concern here is ontological causality, there is a problem in determining which of the two –individual action or group behavior– is primarily causing the other. Similarly, if we take the bee dance (or waggle dance) as an example, the same problem arises: a forager bee interacts locally with her hive mates in a series of movements that informs them about the proximity of water or food sources or even new housing locations. These local interactions will affect the locations were the rest of the bees will move to in order to gather aliment (supervenience). In turn, the overall behavior of the hive, where it eats, where it lives, affects were the individual bees are and what they do at any particular time. Thus, the elements for a potential circular argument appear again.

When proposing reflexive downward causation, Kim argues against one of the prominent advocates of contemporary emergentism, the neuropsychologist Roger Sperry, who used the expression synchronic reflexive downward causation to talk about this. According to Kim, Sperry’s arguments fall into the circularity problem just mentioned. The problem in Sperry’s formulation, according to Kim, is that in combining downward causation and upward determination synchronically, the causal base of the emergent property must be affected by this property that it creates at the time it is creating it. That is, the effects of the emergent property must exist at the same time that the property is being created. To overcome this circularity problem, Kim proposes ‘diachronic reflexive downward causation’, in which he introduces a time delay between the emergent property appearing and its influencing onto the lower levels. Here, the emergent property ‘M’ of the whole ‘W’ causes the micro constituent ‘a’ to have a property ‘p’ in a latter time than that which instantiated ‘M’. There is a “time delay … [that] removes potential circularity” (Kim, 1999).

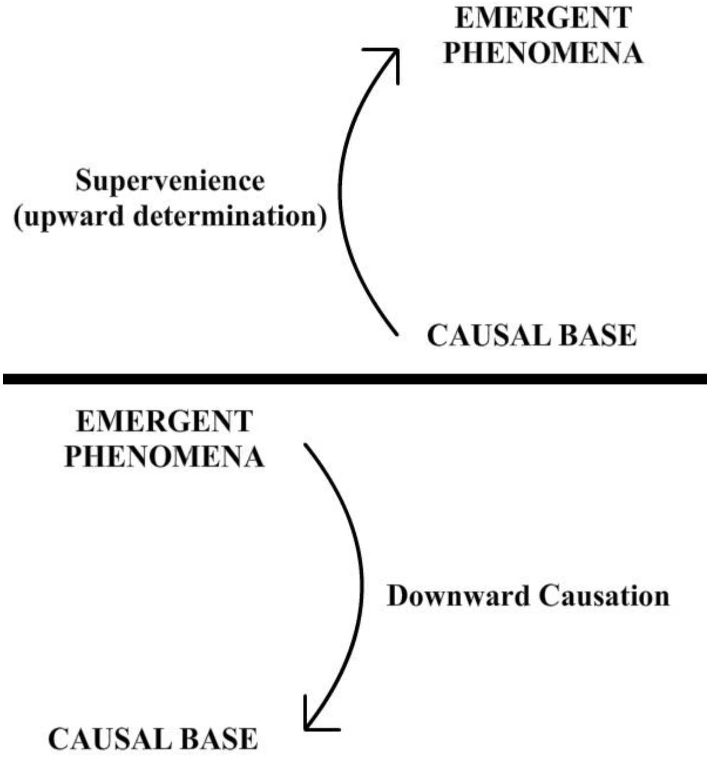


Figure 2.3: Diachronic reflexive downward causation: supervenience and downward causation don’t happen at the same time (synchronically). There is a time delay between the first and the second.

In conclusion, Kim describes two kinds of downward causation, one of which is problematic, while the other is rather uninteresting: “We must conclude then that of the two types of reflexive downward causation, the diachronic variety poses no special problems but perhaps for that reason rather unremarkable as a type of causation, but that the synchronic kind is problematic and it is doubtful that it can be given a coherent sense” (Kim, 1999). In fact, Kim says, some emergentists would argue that this notion of synchronic downward causation is not related to emergence at all, leaving only validity to the diachronic. But even this second one he finds problematic. Kim describes the issue with a single argument: “If an emergent, ‘M’, emerges from basal condition ‘P’, why can’t ‘P’ displace ‘M’ as a cause of any putative effect of ‘M’? Why can’t ‘P’ do all the work in explaining why any alleged effect of ‘M’ occurred?” That is, if the base level causes what is emergent, then the base level is the cause of its own modification through downward causation. And, in a way, Kim argues that in such a scheme we could ignore emergence entirely. There is a problem, then, in considering the relation between the basal condition and the effect as casual (in the downward causation sense). When the base changes, this change can be entirely explained with no regard to the alleged emergent effect, as analyzed by Kim, in a “simple argument [that] has not so far been overcome by any effective counter-argument.” Kim closes this argumentation with a conclusion that “isn’t encouraging to emergentists: If emergent properties exist, they are causally, and hence explanatorily, inert and therefore largely useless for the purpose of causal/explanatory theories” (Kim, 1999).

Despite the discouraging assessment, however, Kim uses his last (and long) paragraph of his 1999 paper to argue that this doesn't necessarily mean to kill downward causation (and therefore emergence), but "we may try to salvage [it] by giving it a 'conceptual' interpretation." Instead of interpreting the different levels as levels of phenomena in reality, we can interpret them as levels of concepts and descriptions. Then, when a cause is described in terms of concepts at the higher level, in relation to the concepts in which its effect is represented, we can speak of downward causation. This conceptual approach, Kim admits, brings up new questions and it might not save "real downward causation". But still he considers that "it may be a good enough way of saving saving 'downward causal explanation,' and perhaps that is all we need or should care about." Such an approach proposes a change of focus from the ontology to the level of description and, thus, to epistemology. This is reminiscent of Second Order Cybernetics, and Von Foerster and Maturana's question of the observed reviewed in chapter 4. It relates, too, to the views of emergence that give a paramount role to the maker of the description: the observer; in particular, to Peter Cariani's emergence-relative-to-a-model which will be reviewed in Chapter 5.

2.4.3 Final Remarks

Summarized, Kim's approach to emergence can be described as follows:

- Emergent properties are not reducible to their lower-level bases.
- Emergent properties need downward causation in order not to be mere epiphenomena.
- Downward causation combined with upward determination (emergence) is circular and inconsistent if synchronic.
- Downward causation combined with upward determination (emergence), if diachronic, is inert and largely useless, since the 'emergent' part can be skipped in the causal explanation.
- We can salvage (sic) downward causation (and therefore emergence) if we give it a conceptual explanation.

Kim does not elaborate in the idea of the conceptual explanation, and it does not seem to offer a satisfactory way to come to terms with the notion of emergence and emergent phenomena. A less discouraging approach, if one seeks to advocate for emergence, might be to use the concept of channeling that will be presented in section 2.6 as introduced by John Protevi. The use of the notion of channeling as a way to explain how the emergent properties influence the behavior of the elements at the basal level liberates emergence from such a strong commitment to downward causation as Kim demands. It is consistent also with some scientific approaches to emergence such as J. A. Scott Kelso’s, reviewed in the following chapter.

2.5 A Process-Based Model: A Critique on Kim’s Physicalist Metaphysics.

Along with those philosophical discourses that are directly connected to complexity sciences and artificial life, the philosophic discourse on physicalism is also a fecund ground for the discussion on emergence and the possibilities of emergent properties. As said above, one of the most prolific and influential authors is Jaegwon Kim, who, according to Richard J. Campbell and Mark H. Bickhard, writes in the context of physicalism. It is in relation to this context that these two authors elaborate a critique of his work and of physicalism in general. In doing so, they set the ground rules for what they call a process-based metaphysics which would allow for an explanation of emergent phenomena, as opposed to what they consider a false metaphysical framework in which physicalism is based.

2.5.1 Emergence and Physicalism

According to Campbell and Bickhard, the main thesis of physicalism is that everything that is relevant in causal terms happens “at the fundamental level of concrete physical particulars” (Campbell and Bickhard, 2001). That is, all of what happens at superior levels is causally redundant. Events such as those related to human intentionality, here, ought to be reduced to events at the physical level. Within this strong claim, the authors identify two main positions in physicalism; reductive and non-reductive. The reductive physicalism claims that the mental phenomena can actually be reduced to the physical, while the non-reductive denies such possibility. However, in any of both versions, the causality is always and exclusively in the physical level: “even though there are no scientific laws by which mental phenomena could be ‘reduced’ to physical phenomena, the underlying causality of the world remains entirely physical” (Campbell and Bickhard, 2001).

For the non-reductive approach, there’s a difficulty in how to explain that mental properties and events cannot be reduced to physical properties and events, and yet they are dependent on them. It is in this context that the ideas such as supervenience or emergence can be found useful. However, the authors state that this metaphysics underlying physicalism is faulty, and they base their claim on the work of Jaegwon Kim who, according to them, “has identified a ‘reductio ad absurdum’ of physicalist metaphysics in general” (Campbell and Bickhard, 2001). What is interesting here in regards to the study of emergence is that, in doing so, Kim “opens the logical space for a fecund notion of genuine emergence.” According to them, Jaegwon Kim invalidates the claim of non-reductive physicalism with the arguments discussed in the previous section. In short: by pointing out the circularity problem of combining downward causation with upward determination diachronically.

In the case of reductionist physicalism, the problem is much more obvious, since the authors consider that any model that denies any kind of emergence is self-refuting. The arguments based on elementary particles (to which everything could be reduced) are also incoherent with the latest developments in science, which demonstrate that "there are no elementary 'particles', fundamental events, or some such particulars." On the other hand, the acceptance of emergence "in the sense of higher-level, causally efficacious powers that are not explicable in terms of the lower-level powers of physical constituents" is incoherent with the basic claims of reductive physicalism. Therefore, it is an impossible position to sustain. In fact, their claim is even broader at this point: "Any metaphysics that respects contemporary physics has to accommodate genuine emergence, and not treat it as merely 'apparent'" (Campbell and Bickhard, 2001).

According to Bickhard and Campbell, once we accept an ontology based on different levels of reality such as the one underlying the theory of emergence, certain issues must be taken into consideration. The authors note that this scheme assumes a layered picture of the world, a mereological model of nature in which "entities, characterized by their distinctive properties and processes, emerge (in some sense) out of the entities, properties, and processes of the levels below it" (Campbell and Bickhard, 2001). One of the risks of such a scheme, they say, is what the authors call the causal drain.

The causal drain is the discontinuity, in terms of causality, between the (emergent) properties at higher levels and the constituents of lower levels. That is, the causal powers at the lower levels are the basis from which emergent phenomena depend on (supervene from). But these causal powers must be based on another more elementary level, and this goes on until a layer of elementary particles. Otherwise, the causal drain happens. But since, as said above, Campbell and Bickhard consider that elementary particles are incoherent with the current scientific theories, such a scheme is rendered invalid in principle. Regarding all this, according to them, Kim in fact endorses emergence: "[Kim has] reached an important insight in recognizing that 'configurations of constituents are what generate the emergence of higher-level causal powers.'" (Campbell and Bickhard, 2001).

Along these lines, they argue, Kim opens up the possibility for mental properties to be emergent too, "precisely because they 'are' 'configurations' – organizations of neuro-physiological processes. … Mind, or mental properties, 'emerge' in specific 'organizations' of lower level processes" (Campbell and Bickhard, 2001). The authors note that Kim doesn't really elaborate on this claim they consider to be so important. The reason, they say, would be that after all Kim is a physicalist.

2.5.2 A Process Model of Emergence

Finally, Campbell and Bickhard propose a process-based metaphysics in which "the mere possibility of emergence need no longer be regarded as problematic" (Campbell and Bickhard, 2001). Allegedly, their metaphysics accommodate complex systems in both their internal workings and their interactions with the environment. It is a metaphysics that applies to far-from-equilibrium dynamical systems. In such systems, they argue, genuine examples of emergence can be found.

Far-from-equilibrium systems exhibit processes with three main characteristics, which the authors list, and which in turn become basic ground-rules for this kind of emergence. First, they are necessarily open. That is, they interact in meaningful ways with their environment. These interactions are crucial for the processes to persist. A candle flame is a good choice of example here. As long as it has fuel, its interaction with its environment is what sustains it. It is what allows it to manifest persistence and stability. Second, these systems exhibit self-maintenance. So long as some minimum boundary conditions are met, both within the system producing the process and its environment, the process contributes to its own persistence. In the case of the candle flame, the induction of convection currents to pull in oxygen is an example of it. For Campbell and Bickhard, this is an example of an emergent process that is causally relevant: “The ability to be self-maintaining is an ‘emergent causal power’ of the organization of the candle flame; it cannot be explained simply as the physical resultant of the causal properties of its distinct constituents.” (Campbell and Bickhard, 2001). This self-maintenance capability is what in other contexts, mostly regarding living systems, is known as autopoiesis. It is the ability of a living system to maintain its own life, and it is often presented in relation to the idea of emergence. Finally, there’s a third characteristic exhibited by some systems which is adaptation (or recursive self-maintenance). That is, some of these systems “can switch to deploying ‘different’ processes depending on conditions they detect in the environment” (Campbell and Bickhard, 2001).

All this entails an ontology which is contrary to the reductionist model. Biological systems –including humans– are not reducible to entities constituted by cells, which are made of elementary particles. “They are open, organized action systems, in ‘essential’ interactions with their environments, such that we cannot say what they ‘are’ without taking those interactive processes into account.” This ontology questions the idea of the causal base for emergence, if the processes and patterns of interaction among elements are not taken into consideration. John Holland (1998) makes a very similar claim in his account on emergence, in considering the interactions among agents to be a crucial element to understand such phenomena[13]. Therefore, it is in the organization of the constituents –in their patterns of behavior– and not on the constituents alone where the emergent properties are to be found.

In this process-based model of emergence that Campbell and Bickhard present, a property such as the candle flame’s ability to maintain itself is “a genuinely ‘emergent’

---

[13] See chapter 3.

causal power.” Along the same lines, biological (autopoietic) systems are complex organizations of processes which “persist ‘only so long as’ they are able to maintain appropriate interactions with their environments, by which to sustain their existence.” According to the authors, in stable far-from-equilibrium process systems we can find “the most interesting and intriguing of emergent causal powers: life, consciousness, and self-consciousness” (Campbell and Bickhard, 2001).

2.6 John Protevi

John Protevi is a Deleuzian philosopher who has written on emergence from the philosophy of Science point of view. Protevi situates emergence in the philosophical discourse stating that it “plays a crucial role in debates in philosophical reflection on science as a whole (the question of reductionism) as well as in the fields of biology (the status of the organism), social science (the practical subject), and cognitive science (the cognitive subject)” (Protevi, 2006). In discussing downward causation, he introduces the notion of channeling as a means to overcome the difficulty in combining the ideas of emergence and causality.

2.6.1 Complexity Theory and Chaos Theory

Protevi proposes a rather elaborate yet interesting definition of emergence. In its simpler formulation, he defines it as “the (diachronic) construction of functional structures in complex systems that achieve a (synchronic) focus of systematic behavior as they constrain the behavior of individual components” (Protevi, 2006). He summarizes complexity theory in one paragraph in which, without using the word, he defines emergence in terms of dynamical systems and complexity: “complexity theory models material systems using the techniques of nonlinear dynamics, which, by means of showing the topological features of manifolds (the distribution of ‘singularities’) affecting a series of trajectories in a phase space, reveals the patterns (shown by ‘attractors’ in the models), thresholds (‘bifurcators’ in the models), and the necessary intensity of triggers (events that move systems to a threshold activating a pattern) of these systems. By showing the spontaneous appearance of indicators of patterns and thresholds in the models of the behaviour of complex systems, complexity theory enables us to think material systems in terms of their powers of immanent self-organization” (Protevi, 2006).

He distinguishes complexity theory from chaos theory, a distinction which is very relevant regarding the importance and the context in which emergence can be understood: “Complexity theory is not chaos theory. Chaos theory treats the growth of unpredictable behavior from simple rules in deterministic nonlinear dynamical systems, while complexity theory treats the emergence of relatively simple functional structures from complex interchanges of the component parts of a system. Chaos theory moves from simple to complex while complexity theory moves from complex to simple.” Thus, both chaos and complexity theory operate in some respects in similar terms, and they are strongly related. Nonetheless, the distinction is important as it serves to distinguish two phenomena that might otherwise be taken for one another: emergence and what is known as Sensitive Dependence on Initial Conditions.

Protevi states that, just as the two mentioned disciplines must be differentiated, so must these two phenomena. His point is that although emergent phenomena and Sensitive

Dependence on Initial Conditions might seem to be very similar in some respect, there's an element of predictability in the latter that invalidates a strong comparison with emergence: "Although the behavior of chaotic systems is unpredictable in quantitative detail it is sometimes predictable in the long run or 'qualitatively'". And later on he insists: "We have an irreducible element of 'chance' [i.e. a fractal attractor], even though the system is thoroughly deterministic" (Protevi, 2006).

2.6.2 Synchronic Emergence

Protevi identifies, like Kim, a series of key issues around the concept of emergence. Among them, the most relevant is again downward causation, but he takes a radically different approach here than that of Kim, to the extreme of labeling it "the false problem of 'downward causation'" (Protevi, 2006). He proposes to understand downward causation in terms of quasi-causes (a term borrowed from DeLanda). Whilst authors such as Kim understand it as a case of efficient causality, Protevi proposes this idea of a more subtle influence of the emergent phenomena to its causal base.

He bases this argumentation on the idea of channeling, which would be here a sort of 'soft' downward causation: a mechanism that effects the behavior of the element in the lower level by channeling its behavior, rather than causing some change on its behavior in a direct manner. In other words, the emergent phenomena would generate an attractor for the behavior of the agent in the causal base. This notion of channeling is key here. It allows the author to solve the central problem articulated by Kim in his critique of emergentism and the importance of downward causation. If the emergent phenomenon channels the behavior of the base level, the phenomena must be taken into consideration in order to explain the overall behavior of the system. It also connects nicely with authors such as Kelso, who also explains emergence in terms of pattern formation.

This idea of channeling as soft downward causation can also be read in terms of self-organization emergence. If read from the perspective of the boids example (presented in chapter 1), some ideas of downward causation would imply that the rules that govern each individual agent in the simulation change because of the overall pattern of behavior generated, which is not the case. In contrast, channeling would mean that the behavior of the agent is affected by the pattern that is generated. And indeed the agent will end up going towards one direction or another depending on the overall direction that the group takes. In other words, the emergent pattern becomes an attractor to the behavior of the agent.

In any case, Protevi does not insist so much on this point as in its general classification of emergence, which starts with three different notions that will gain some complexity as the discussion advances: synchronic emergence (order), diachronic emergence (novelty) and, closer to his interest in Deleuze and Guattari, a "third form of emergence, transverse emergence in assemblages, [which] is what I call 'political physiology'."

The first case would be the more typical idea of emergence, “the emergence of ‘order out of chaos’.” In short, it is the same idea as Bedau’s nominal emergence (discussed below), or Kelso’s process of pattern formation[14]. Here, the system is kept inside a behavioral pattern: “What keeps a system inside a behavior pattern –represented by the trajectories inhabiting a basin of attraction– is the operation of negative feedback loops that respond to system fluctuations below a certain threshold of recuperation by quickly returning to the system to its pattern. … These quickly recuperating systems are called ‘stable’” (Protevi, 2006). When this stability is broken, systems can either die or adapt (learn) by showing resilience. That is, when the system is pushed beyond its comfort zones (its stability) it will create new attractors representing new behaviors. This development can be seen as evolution[15], as adaptation in cybernetic or enactive terms or, in Protevi’s terms, ‘diachronic emergence’. Protevi defines this type of emergence as a “structure which enables focused systematic behavior through constraining the action of component parts”. Note the idea of constraining, as equivalent to channeling.

As said above, the idea of channeling is very important to solve the problem (false problem for Protevi) of downward causation. He compares this notion to that of the final cause in Aristotle. The emergent phenomena, according to him, channels the behavior of the base level elements in the same way the Aristotelian final cause channels the development of the form: “The final cause or end state channels development; the infant does not intend to grow into an adult. It is this notion of channeling which is the key to understanding systematic constraint and focused behaviour in synchronic emergent functional structures. In other words, synchronic emergence is a misnomer; there is always a coming into being of functional structures which needs to be conceptualized. The final cause or end state channels the development” (Protevi, 2006). It also resonates with Kim’s proposed solution to the circularity argument. The emergent phenomena and its influence on the basal level can’t happen exactly at the same time, so there is a time delay between one and the other, the ‘coming into being’ of Provevi. Thus, this ‘coming into being’ introduces time and process in the conceptualization of emergence. From here, Protevi moves naturally to the next two types of emergence: Diachronic and transverse.

2.6.3 Diachronic and Transverse Emergence

Protevi notes that synchronic emergence uses the notion of reality organized in levels. When talking about diachronic emergence, one must account also for different time-scales among levels. “But the perspective of diachronic emergence shows that the time scales of each level are staggered, so that what appears as a systematic unity on a

---

[14] See chapter 3.

[15] See Chapter 5 on Cariani.

specific level is an event, a process, from the perspective of another level with a longer time scale. We can call this heterochrony: cells come and go but the organ stays (relatively) the same; people die but the social body lives on, and so on" (Protevi, 2006). Finally, rooted in Deleuze and Guattari, he introduces transversal emergence, which is a concept related to the cited authors' assemblages. These are systems within a territorial context operating as base levels to create greater potential for emergent phenomena. The details, background and implications of this characterization are beyond the scope of this study.

With all this, Protevi, apologizing for a "rather barbaric terminology", articulates a classification of the types of the emergence as follows:

- Homeostratic[16] synchronic transversal emergence
    - organic (symbiosis among organisms; ecosystems among groups of organisms)
    - social (institutions forming a larger entity: e.g., colleges forming a university)
    - technical (e.g., computers and routers forming the Internet)
- Homeostratic diachronic transversal emergence
    - organic (symbiogenesis: Margolis's theory of the origin of the eukaryotic cell)
    - social (system change: e.g., change of the university from education of an elite into a centre for mass vocational training/ military-industrial research)
    - technical (system change: e.g., from ARPANET to Internet to world wide web)
- Heterostratic synchronic transversal emergence (a bio-socio-technical assemblage)
- Heterostratic diachronic transversal emergence (mutation and coevolution of such assemblages in 'machinic phyla')

One of the relevant points of this classification is the differentiation between processes that happen among different levels (heterostratic) and those happening in the same

---

[16] The terms homeostratic and heterostratic refer to the layered organization of reality. The first is concerned with phenomena occurring within a single level, whilst the second concerns those in which different levels are involved.

(homeostratic). With this, he connects with the authors that, like Cariani, argue that emergence is not necessarily a matter of processes at one level (basal level) influencing on a superior one.

2.7 Mark A. Bedau

Mark Bedau is, like Kim, one of the contemporary philosophers who has a significant body of work on Emergence. His argumentation, too, is rooted on ontological concerns and on the stratification of reality. The notions of micro and macro levels are paramount in his discourse. One of Bedau's main goals is to make emergence compatible with the scientific discourse. Like Protevi, he articulates his discussion in the context of complexity sciences and, in his case, especially within artificial life.

2.7.1 Weak Emergence

In his work, Bedau distinguishes three kinds of emergence: nominal, weak and strong (Bedau, 1997; 2008). The three notions share a 'core component', which is "the notion of a kind of property that can be possessed by macro objects but cannot be possessed by micro objects" (Bedau, 2008). Nominal emergence is just that: a macro property that cannot be a micro property, but as Bedau says this is too broad a concept, one that applies both to resultant and to emergent properties, in the terms of Lewes' classical distinction. In order to find more restricted notions of emergence one would need to turn to strong and weak emergence.

The distinction between weak and strong emergence is not exclusive from Bedau. Assad and Packard use the same terminology when they define the scale of emergence that goes from non-emergent to strong emergence[17] (Assad and Packard, 2008). The basic idea is that strong emergence would relate to the actual ontologically valid concept of emergence (i.e. the ontological existence of emergent phenomena), whilst weak emergent would be the understanding of emergence as an epistemological construct; as a useful tool to understand certain phenomena, for example though the use of computer simulations.

The distinction finds its roots in John Searle's distinction between strong and weak Artificial Intelligence (Pattee, 1989). According to Howard Pattee Searle distinguishes among a 'simulation school' and a 'realization school' in AI, which would relate to strong and weak AI respectively. Indeed, the difference that Searle established between both AI schools is relevant to the analysis of the different approaches to emergence within Artificial Life. Searle states that, while weak or 'cautious' AI understands that the computer is a very powerful tool in order to understand how the mind works, and thus simulates it with this objective, strong AI understands that these computer programs, if properly structured, create an actual mind: "according to strong AI the computer is not merely a tool in the study of the mind; rather, the appropriately programmed computer really is a mind, in the sense that computers given the right

---

[17] See section 3.3.4.

programs can be literally said to understand and have other cognitive states. In strong AI, because the programmed computer has cognitive states, the programs are not mere tools that enable us to test psychological explanations; rather, the programs are themselves the explanations" (Searle, 1980).

Searle's arguments are classically presented against the imitation game of Alan Turing as a test to determine whether or not a machine can think. The Turing test was presented as the 'imitation game' in (Turing, 1950). It is a thought experiment that was conceived, in Turing's words, to discern whether or not a machine was capable of thinking. According to him –and to latter (strong) AI researchers, says Searle– the answer to this question will be positive if the test is passed by one.

Turing proposed to think of a game where three persons are connected via text-based terminals: a player, a man and a woman. The game is that the player has to guess, from the answers gotten from the other two, who is a man and who is a woman. Either man or woman will always give truthful answers to the player, while the other will try to deceive the player. In the text-based terminal, man and woman are identified as person X and person Y by the player, so he or she knows who is answering what, and they all see the whole conversation. Simplified, the Turing test consists in substituting the deceiver by a computer. Then, if the computer is good enough at tricking the player according to some predefined conditions, we can conclude that the computer can think.

In opposition to this idea, Searle proposed in (1980) the counterargument of The Chinese Room. It is another thought experiment. In it, someone who does not speak any Chinese is given a large set of Chinese character combinations and a set of rules that relate these combinations among them. Then, if someone asks in writing questions in Chinese, the room's inhabitant (who has no contact to the outside world besides this) uses the rules to correlate these questions to answers, unknowingly of any of the meanings of either. If the rules were sophisticated enough, the correlation would be successful, so the person outside the room perceives the answers as correct. If strong AI is right, says Searle, then we would have to agree that the person inside the room actually speaks Chinese, as much as the Turing test computer actually 'thinks'. The problem is, according to Searle, that strong AI gives ontological validity to what is only a correlation between input and output in a computer program: "The Turing test is typical of the tradition in being unashamedly behavioristic and operationalistic, and I believe that if AI workers totally repudiated behaviorism and operationalism much of the confusion between simulation and duplication would be eliminated" (Searle, 1980)[18].

Similarly, the distinction between strong and weak emergence coincides with the actual existence of such properties or the concept that is a useful tool for explaining certain

---

[18] Despite the relevance of the discussion, it is beyond the scope of this thesis to delve into it further.

phenomena. According to Bedau, strong emergence is either inexistent or at least scientifically irrelevant. It would be a case in which the emergent properties are brute forces unexplainable by the aggregation of micro-level causes, with an irreducible causal power onto the micro-level elements (strong downward causation).

For Bedau, this is problematic mainly because it competes with causation (and with causal fundamentalism), but also because of the competition that would occur among micro and macro causal powers in such a case of emergence (the exclusion argument) – a similar problem to that of the circularity identified by Kim. Consequently, he situates strong emergence outside scientific inquiry and, therefore, outside of his interest in reconciling the problematic of emergence with the scientific discourse: “strong emergence starts where scientific explanation ends” (Bedau, 2008).

The intermediate notion between strong and nominal emergence is weak emergence. It is given in a system when its global behavior derives from the processes which operate on the micro-level, but this derivation is not easily explained due to the complexity of the micro-level element’s interactions. Similarly to Holland’s notion of emergence[19], Bedau explains: “Weak emergent phenomena can be derived from full knowledge of the micro facts. Weak emergence attributes the apparent underivability of emergent phenomena to the complex consequences of myriad non-linear and context-dependent micro-level interactions. … Weak emergent phenomena are ontologically dependent on and reducible to micro phenomena” (Bedau, 2008). That is, what results from the multiple interactions is too complex to be determined only through observation of the lower level elements in isolation, but it is in fact, ontologically, determined by it in its entirety.

In the ant colony or the slime mold examples, this means that what happens at the higher level –the complexity of the colony or of the organism– is not explainable if the elements are observed on its own –individual ants or cells–. However, in ontological terms, according to Bedau, the causal link exists and reductionism is still a valid notion in the explanation. Therefore, according to this weak emergence displaces the problem of emergence from ontology to epistemology. In Bedau’s conceptualization of weak emergence the ontological perspective is assumed to be unproblematic. Emergent phenomena are reducible to –and thus explainable in causal terms from– the causal base. It is in explaining this scientifically that the difficulties appear; in elaborating the explanation of the phenomena through observation.

Another important aspect of the discussion is that Bedau argues that this notion of weak emergence allows him to solve the problem of conjunctly validating what he previously introduced as the two hallmarks of emergence: the need for emergent phenomena to be both dependent on and autonomous from underlying processes (Bedau, 1997). Later, he

---

[19] See chapter 3.

notes: “There is nothing metaphysically illegitimate about combining this explanatory autonomy (irreducibility) with ontological and causal dependence (reducibility), so weak emergence dissolves the problem of emergence” (Bedau, 2008).

2.7.2 The Importance of the Simulation and the Game of Life

Once this has been established, Bedau proceeds to give a definition of weak emergence that is very relevant to ALife art systems, since he uses the idea of simulation as a central concept (and uses Game of Life as its main driving example): The difference between a nominal emergent and a weak emergent macro property is that the latter is derivable from the micro facts ‘only by simulation’: “Assume that ‘P’ is a nominally emergent property possessed by some locally reducible system ‘S’. Then ‘P’ is weakly emergent if and only if ‘P’ is derivable from all of ‘S’’s micro facts but only by simulation” (Bedau, 2008).

This is neither about only artificial systems nor about simulations that are possible or have been done, but about plausible simulations both in natural and artificial systems. In fact, in this view natural evolution (or even the course of time) would be nothing more, ontologically speaking and in this context, than a simulation. It is a notion that fits perfectly, as Bedau himself notes, with complexity science, which uses computer simulations in the study of its complex systems. And like Varela, Thompson and Rosch (Varela et al., 1993) or Wolfram (2002)[20] he exemplifies his theory with cellular automata, and especially the Game of Life, which, in this context, serves as a perfect formal system, a very simplified and controlled model of reality that allows researchers to speculate about how patterns form from simple rules and how they relate to the elements at the base.

Bedau uses Game of Life to exemplify weak emergence. He examines some of the configurations of this cellular automaton in order to discern whether or not they exhibit infinite growth (i.e., they generate a pattern that creates new cells ad infinitum). This would be the case of the glider gun, a configuration that generates gliders –small structures that move towards the edge of the simulation space– at regular intervals (Flake, 1998). According to him, with some simple configurations it is easy, once someone knows the rules of the simulation, to anticipate whether or not they will generate infinite growth. A row of three cells will generate a blinker, constantly changing the row from vertical to horizontal, but no growth or anything else. A two by two block of cells will do nothing. But he then proposes to consider the structure know as R pentomino. A very simple structure made out of five cells shown in figure 2.4, which is known to exhibit a “wildly unstable behavior.” But the answer to whether it will exhibit infinite growth is, according to Bedau, only attainable by simulation. That

---

[20] See chapter 3.

is, by letting the Game of Life "'play' itself out with the R pentomino as initial condition. That is, one has no option but to observe the R pentomino's behavior" (Bedau, 2008).

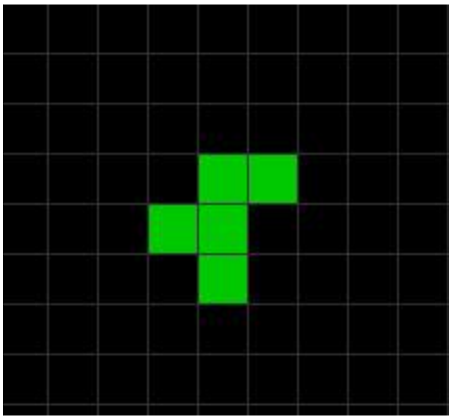


Figure 2.4: The R pentomino.

As it turns out, simulating this will led to a halt, precisely, at step 1103. This is, according to Bedau, a case of weak emergence: "The halt of the growth of the R pentomino is a weakly emergent macro state in the Game of Life." Therefore, according to Bedau's formulation, whether or not a certain configuration's outcome is weakly emergent depends on whether or not one needs to run the simulation in order to anticipate it. The R pentomino, as wild as it may seem to be, always does the exact same series of steps. A screenshot of the simulation at one particular time will always be the same (see figure 2.5). So the question is here, in fact, epistemological.

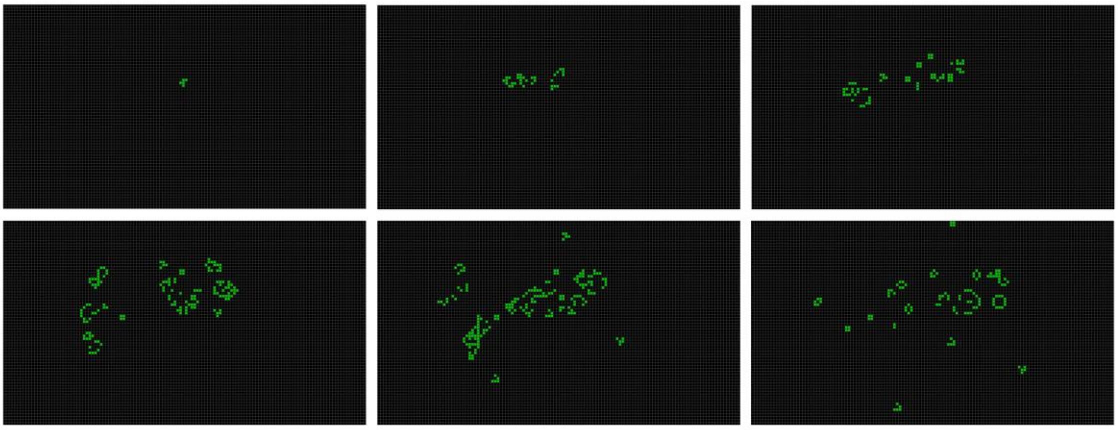


Figure 2.5: R pentomino's simulation. From left to right, up to bottom, steps 0, 40, 80, 120, 160 and 200.

The problem with Bedau's formulation is that it is unclear where the need for the simulation starts and where it ends. With the row of three cells, for instance, it is save to say that any informed observer will arrive at the conclusion that it will result in a loop of horizontal lines of three cells alternating without the need for simulating it –i.e. without iterating through the process, whether mentally, in paper or with a computer–. And it is also coherent to state that most likely no one will arrive to the conclusion that the R pentomino stops at step 1103 without running a simulation. But consider the form of figure 2.6, a modification of the configuration known as the beehive. This configuration of seven cells will exhibit a pattern that goes into a halt in exactly sixteen steps. Some of them are shown in figure 2.7.

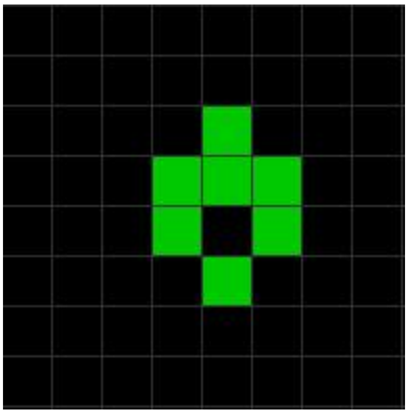


Figure 2.6: Modified beehive

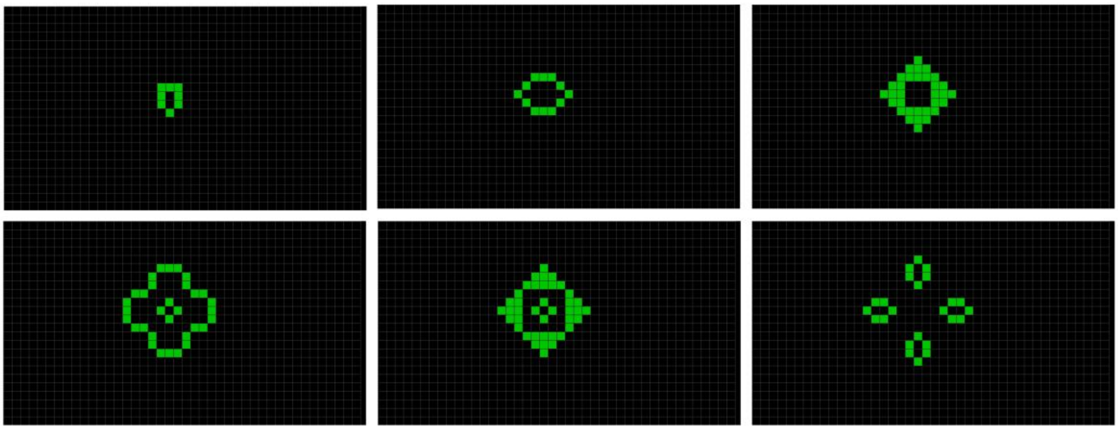


Figure 2.7: Modified beehive pattern. From left to right, up to bottom, steps 1, 4, 7, 10, 13 and 16.

Thus, the question here is: are sixteen steps still too much to deduce without simulating, or could an expert in cellular automata and Game of Life anticipate, without the need of a simulation, that the configuration will indeed come to a halt in this exact number of steps? If so, how exactly is anticipating sixteen steps different from anticipating two, as in the case of the three cell blinkers, or anticipating over one thousand, as with the R pentomino?

Whatever the answer might be for this particular case, it is clear that this notion of weak emergence leaves too much space to intuition. The threshold where the need for the simulation starts and ends is not only unclear but impossible to communicate among researchers. Once again, this illustrates how approaches like Cariani's emergence-relative-to-a-model, which might seem restrictive under certain points of view, are the only way to create a frame of reference which can be explicated and shared among different people in order to discern when phenomena are emergent and when they are not.

2.7.3 Weak Downward Causation and Autonomy

In this categorization of emergence, as Bedau goes on to argue in detail, both the autonomy and downward causation problems are solved. First, with a notion of 'weak downward causation', in which according to him there is no violation of micro-laws and its cause-effect relations: "higher level properties can causally influence the conditions by which they are sustained, but this process unfolds over time" (Bedau, 2008). Secondly, he argues in favor of autonomy. In some cases, autonomy is nothing but epistemological. It is what he calls an explanatory autonomy, which signals nothing in the macro structure in reality. But in other cases he identifies "robust macro regularities that can be described and explained only at the macro level." This is strongly reminiscent of Kelso's patterns reviewed in chapter 3.

For Bedau, this sort of autonomy has ontological validity: "This kind of robust weak emergence reveals something about reality, not just about how we describe or explain it. So the autonomy of this robust weak emergence is ontological, not merely epistemological" (Bedau, 2008). His notion of weak emergence does not intend to explain all the aspects over which emergence is projected, but it does attempt to solve and clearly explain a big part of it.

A very interesting point in his approach is the insistence in the importance of time, which connects him to Bergson and Protevi. His emergence is not about wholes being more than the sum of its parts, but about wholes (or macro) arising over time from the interactions of the parts (the micro): "The primary focus of weak emergence is diachronic. It concerns how the macro arises over time from the micro, i.e., the causal process (derivation) by which the micro constructs the macro. This is a bottom-up generative process, rooted in context-sensitive micro-level causal interactions" (Bedau, 2008).

This relates to the importance of the notion of simulation, discussed above. An importance that implies that the notion of weak emergence, in combination with this crucial role of simulation, and although Bedau doesn’t explicitly acknowledge this, represents a displacement of the problem of emergence from ontology to epistemology, and could be even linked to Cariani’s claim that emergence must always be framed within an observation of a system in action (through the looking glass of a clearly predefined model). However, Cariani would exclude emergence from a case like the Game of Life; once a pattern has been observed, it is incorporated into the model and ceases to be emergence. Bedau, in contrast, elaborates his theory precisely with Game of Life as a key example. But then in Bedau, like some other authors, the issue of what exactly is the complexity that makes a phenomenon emergent, as opposed to a mere resultant, to use the original term, remains unclear.

2.8 Concluding Remarks: The Philosophical Emergence

Emergence is still a strongly debated concept in Philosophy. Whilst in disciplines such as complexity sciences or artificial life emergence is used sometimes without much discussion on what it exactly means and implies, philosophers discuss its definitions and implications in detail. This is done mostly within the discourses of philosophy of science, with some ramifications to physicalism, which studies how mental properties and their physical base relate in terms of causality.

Emergence can be either an epistemological or an ontological construction to explain some aspects of reality that escape the reductionist accounts based on causality. The epistemological approach is, in general, less controversial. Often the debate focuses on whether emergence is simply a consequence of lack of knowledge or a legitimate source of knowledge.

The argument that emergence is just a temporary black box for something that might eventually be explained was one of the main arguments against the British Emergentists (Nagel, 1961; Whitelaw, 2006). And indeed, as explained above, the wrong choice of examples by this first group of emergentists, leading to the explanation in scientific terms of what they had labeled as emergent as twentieth century science advanced, did seem to validate this claim.

As a legitimate source of knowledge, epistemological emergence has been elaborated by Peter Cariani and his emergence-relative-to-a-model, discussed in chapter 5. As an ontological phenomenon, which is how many philosophers read it, emergent phenomena need a much more elaborated philosophical discourse either to be defended or argued against (Ali, Zimmer and Elstob, 1998; Bedau and Humphreys, 2008).

As said above, supervenience and downward causaition are among the key concepts in the discourses on emergence. As noted by Kim, it is especially problematic to combine the two directions in which causality needs to go for emergence to make sense and be relevant. With supervenience (or upward determination) causality goes from the basal level to the emergent phenomena. But it is an irreducible causality, and therefore unexplainable in classical terms.

That is, for some, no causality at all. But still when this is accepted, there is the problem of explaining downward causation. That is, how whatever appeared influences back to the elements in the basal level. Authors such as Bedau or Protevi do consider this to be in the ontological domain, but their argumentations need to be elaborated and refined to counteract Kim's arguments on the circularity of combining both supervenience and downward causation meaningfully. Also, when their explanations are driven towards epistemology, such as in Bedau's weak emergence, then the argumentations of what exactly is this emergence need to be very clear and explicit and, as demonstrated with the Game of Life example, this is not always the case.

This thesis does not attempt to solve all this discussions or to propose an alternative philosophical view. The aim here is to build a frame of reference in order to discern whether or not emergence occurs in interactive systems, either as self-organization or in the creation of systems that will be able to generate new behaviors. It is in pursuing this objective that an epistemological approach, mainly based on Peter Cariani's emergence-relative-to-a-model, will be proposed in the final chapters.

3. EMERGENCE IN SCIENCE

In 1972, Phil Anderson, who would receive the Physic's Nobel Prize five years later, published an important paper in Nature, entitled *More is Different* (Anderson, 1972), in which he questioned reductionism by introducing the idea that it is unable to completely explain complexity at different levels of reality, as some laws are not valid in all of them. The title of the article refers to the classic idea of emergence, which states that there are cases in which a whole is not explainable as a mere aggregation of its parts. The whole is more or, at least, different than the sum of the parts. But even written by a soon-to-be Nobel Laureate, emergence, had not been, was not, and would not be considered an important scientific topic for some time.

As previously discussed, emergence was not identified as a phenomenon in academic discourse until the mid–ninetieth century, when John Stuart Mill used the concept to distinguish different types of causation, but it still remained a marginal concept for a long period. In the Newtonian Science paradigm, emergence was unknown and unknowable, since reductionism was an indisputable method. But even when the once revolutionary 'modern' science was reshaped by the Twentieth Century revolutions of Relativity and Quantum Mechanics, emergence remained an outsider to scientific discourse. It was not until the second half of the twentieth century that the work of some rather unorthodox scientists prepared the context for emergence to appear in its contemporary form. By the end of the century, it was a central concern in the Complexity Sciences (ALife, dynamical systems theory, neural networks, etc.).

The first thinkers and experimenters to reintroduce, if not emergence per se, a context for it to be relevant, were the 1950s British Cyberneticians (Pickering, 2007; 2010). Some of these artists and experimenters were the predecessors of Artificial Life, a discipline in which emergent properties acquired a fundamental role. A little more than ten years after writing the mentioned paper, Anderson would be one of the founders of the New Mexico's Santa Fe Institute, an innovative interdisciplinary center for what they decided to call Complexity Sciences, in which disciplines such as Artificial Life would appear (Waldrop, 1992). It was within Chaos Theory and the Complexity Sciences of the 1980s that emergence and emergent properties gained a central status as a scientific matter.

This chapter is structured in three parts. First, there is a very brief account of emergence in cognitive science which uses the concept of emergence to explain the irreducibility of some of the processes of cognition. The second section of the chapter covers a series of scientific approaches to emergence in disciplines outside the scope of Artificial Life and Cybernetics. The reviewed accounts are relevant to this study, as they help to contextualize the discussions in respect to the artistic practices developed around the idea of emergence. Finally, the third and last part is dedicated to Artificial Life as a scientific discipline and the centrality of emergence in it. Artificial Life Art will be reviewed in chapters 6, 7 and 8.

3.1 A note on Cognitive Science

The mind, viewed as an emergent phenomenon in respect to the physical neuronal interconnections of the brain, is a classic example of emergence. In this way emergence has a relevant role in some approaches to Cognitive Science. Due to its general relevance, this approach is noted here. However, delving into it is beyond the scope of this thesis. The following is an overview of how the idea of emergence is presented in one of the foundational works of the field.

3.1.1 Enaction, Emergence and Cognitivism

In their highly influential *The Embodied Mind*, Francisco Varela, Evan Thompson and Eleanor Rosh present an overview of cognitive science which includes emergence as central concept. Although their goal is to present the enactive paradigm as a more elaborate proposal than that of emergence, their characterization of emergent phenomena is worthy of mention here.

They characterize cognitive science as consisting of three successive historical stages: cognitivism, emergence and enaction (Varela et al., 1993: 6). The first is basically related to the mental representations, and works with the metaphor of the digital computer as a manipulator of symbols. Here the idea is that reality exists for an observer to represent it through symbols: “The mind is thought to operate by manipulating symbols that represent features of the world or represent the world as being in a certain way” (Varela et al., 1993: 8). This well defined domain, however, finds two main alternative views: emergence and enaction. The first is, according to the authors, synonymous to connectionism, a name “derived from the idea that many cognitive tasks (such as vision and memory) seem to be handled best by systems made up of many simple components, which, when connected by the appropriate rules, give rise to global behavior corresponding to the desired task.” In contrast to symbolic processing, which is localized, the connectionist models trade this localization for “distributed operations (ones that extend over an entire network of components) and so result in the emergence of global properties resilient to local malfunction.” Representation, here, is not a function of particular symbols in isolation. Instead, it is found in the correspondence between the properties of the world and such an emergent global state.

Although they do dedicate a fair amount of space in their book to cognitivism and emergence, the goal of the authors is to defend the enactive paradigm as the most sophisticated one. A view in which “cognition is not the representation of a pregiven world by a pregiven mind but is rather the enactment of a world and a mind on the basis of a history of the variety of actions that a being in the world performs” (Varela et al., 1993: 9).

3.1.2 Emergence in Cognitivism

Varela and colleagues identify two main reasons why a look into self-organization is worthwhile in terms of overcoming the limitations of cognitivism. The first is to find a paradigm for understanding the process of cognition which can suit parallel processing. Symbolic operations are based on the application of sequential rules, which are applied one at a time. This creates a "dramatic limitation when the task at hand requires large numbers of sequential operations (such as image analysis or weather forecasting)" (Varela et al., 1993: 86). In these cases the symbolic operation processing, according to the authors, is inadequate. A second limitation is localization. As mentioned above, symbolic processes are localized. This means that "the loss or malfunction of any part of the symbols or rules of the system results in a serious malfunction." A distributed approach, in contrast, offers a more resilient solution to the localized problems, since it does not depend as much on particular parts of the system.

The idea is, then, to "build a cognitive system not by starting with symbols and rules but by starting with simple components that would dynamically connect to each other in dense ways." (Varela et al., 1993: 8). Each component, here, operates exclusively at the local level, but there is a global cooperation that emerges spontaneously due to the system's networked configuration. This is a typical characterization of emergent properties which, as the authors note, has received different names in different disciplines, from self-organization in cybernetics to synergetics: "This passage from local rules to global coherence is the heart of what used to be called self-organization during the cybernetic years. Today people prefer to speak of emergent or global properties, network dynamics, nonlinear networks, complex systems, or even synergetics" (Varela et al., 1993: 88).

From this point on, they move to the analysis of Wolfram's work on cellular automata[21] to demonstrate how even very simple systems can have rich self-organizing capacities. The idea is to point out that "the emergence of global patterns or configurations in systems of interacting elements is neither an oddity of isolated cases nor unique to neural systems. In fact, it seems difficult for any densely connected aggregate to escape emergent properties" (Varela et al., 1993: 90). And if this is valid for such simple systems, it will also be for complex ones such as the brain, for which operation, according to the authors, emergent properties are fundamental.

From here on, they move to an analysis of how this view copes with the process of cognition and the concept of self, to later argue for their enactive paradigm as a better approach to cognition. For the purposes of this study, it is sufficient to note how emergence (understood basically as self-organization) is portrayed as an intrinsic part of

---

[21] See section 3.3.3.

the cognition process. As in other scientific cases (e.g. Wolfram, 2002; Solé and Goldwin, 2001; Langton, 1986, 1988) this kind of emergence, and parallel concepts like downward causation, supervenience, etc. are not questioned or put under strict scrutiny as it often happens in philosophy (e.g. Kim 1999, 2006; Bedau 1997, 2008; Campbell and Bickhard, 2001). The identification of emergence with self-organization, whether explicit or implicit, allows the authors to assume that it is obvious enough a phenomenon not to polemize about its existence.

3.2 Order out of Chaos

The order-out-of-chaos formulation is another concept commonly associated with emergence. Indeed, it easily relates with self-organization emergence and spontaneous formation of processes of organization. This is one of the departure points of Complexity Sciences and a research space where Artificial Life and non Artificial Life related disciplines converge. For a clarity purpose, I will explain here the approaches that are not strictly developed within ALife, to move to these latter ones in the next section.

3.2.1 Prigogine: Time and Irreversibility

Ilya Prigogine was one of the first scientists to situate emergent properties, or self-organization, in the center of his research on the formation of order in Nature. He wrote about the importance of the 'nonclassical' sciences of complexity, which started, according to him, in the early nineteenth century with thermodynamics (Prigogine and Stengers, 1984). For Prigogine, it is in 'far from equilibrium thermodynamics' where we can find structures that originate spontaneously, dissipative structures in which the elements forming it (the particles) self-organize. In short, dissipative structures here are to be understood as systems which perform some sort of exchange (of energy or information) with their environment. That is, they are necessarily open systems. These are also structures which, because they are far from equilibrium, need in general to be highly energized.

A classical example of it would be chemical clocks: "Oversimplifying somewhat, we can say that in a chemical clock all molecules change their chemical identity 'simultaneously', at regular time intervals" (Prigogine and Stengers, 1984:13). A chemical clock, then, is a process which defies the second law of thermodynamics since it keeps creating states in a stable cyclic process of creation and destruction of order. One of the most well-known examples of a chemical clock is the Belousov-Zhabotinsky reaction. It is a complex mix of chemicals which produces a continuous oscillation of states. In such a process, there is a sort of communication among molecules in chemical clocks. That is, the parts of the system interact, and this creates an overall behavior of the system, which is not explainable solely by the (reductionist) analysis of the particles in isolation: "To change color all at once, molecules must have a way to 'communicate'. The system has to act as a whole" (Prigogine and Stengers, 1984:148).

According to Prigogine, thermodynamics is the first non-classic science in the sense that it deals for the first time with irreversible processes. I.e. processes in which the direction of time is not irrelevant: "It is based on the distinction of two types of processes: reversible processes, which are independent of the direction of time, and irreversible processes, which depend on the direction of time. … It was in order to distinguish the two types of processes that the concept of entropy was introduced, since entropy increases only because of irreversible processes" (Prigogine and Stengers, 1984:12). Within thermodynamics, these irreversible processes were at first considered to be of minor scientific interest. But as Prigogine notes, towards the last decades of the twentieth century the importance of these far from equilibrium processes arose as a context in which new types of structures can originate spontaneously.

Here is where emergent properties fit in Prigogine's context: "The famous law of increase of entropy describes the world as evolving from order to disorder; still, biological or social evolution shows us the complex emerging from the simple" (Prigogine and Stengers, 1984: xxix). This is a common approach in Complexity Sciences: emergent processes originate in the boundary conditions of far from equilibrium systems, where a large number of elements interact with each other. At a certain point, these elements self-organize around some sort of macrostructure that is generated from within the system: an 'attractor' that channels the organization of the group of individual elements.

In these processes, time plays a crucial role, as Prigogine notes while referring to Bergson's work, which he considers to have dealt with not less than "the central problem of Western ontology: the relation between Being and Becoming" (Prigogine and Stengers, 1984:310). The importance of Bergson is to advocate for duration in the scientific discourse. As explained above[22], the French Philosopher defended the importance of time against the 'cinematograph method' of a reductionist science which splits time in order to analyze the elements in isolation: "For Bergson all the limitations of scientific rationality can be reduced to a single and decisive one: it is incapable of understanding duration since it reduces time to a sequence of instantaneous states linked by deterministic law" (Prigogine and Stengers, 1984:92).

Finally, Prigogine objects to the idea of a pure objective knowledge and places the observer inside the system of knowledge, thus presenting "[a] conception of knowledge as both objective and participatory" (Prigogine and Stengers, 1984:299). This situates him, at least in terms of the study of emergence, closer to the epistemological views rather than the ontological. The observer is always within a system in which observation is performed. Simply put: "Nature cannot be described 'from the outside', as if by a spectator. Description is dialogue, communication, and this communication is subject to constraints that demonstrate that we are macroscopic beings embedded in the physical world" (Prigogine and Stengers, 1984:300). This notion of participatory knowledge resonates, too, with second order cybernetics, and the conceptions of knowledge of authors like Maturana and Von Forester. It is what Andrew Pickering, in describing the work of the British Cyberneticians, has called the performative ontology[23]. It also

---

[22] See section 2.3.

[23] See chapter 4.

relates to Peter Cariani's emergence-relative-to-a-model which will be discussed in chapter 6.

3.2.2 Dynamical Systems and Chaos

The study of Dynamical Systems is based on two main concepts. The first is the notion of state (i.e. the information that tells us about how the system is composed at a certain moment). The second is a dynamic, a rule that describes how the system's state evolves over time (Crutchfield et al., 1986). In other words, we need to know what it is about the system that changes over time and which are the rules that determine how this change happens (Flake, 1998). The changes in these systems are explicated through a series of states in which the system can be at a particular moment or the fluctuations in which it is oscillating.

The combination of forces and environment characteristics that push a system to one of these states are called attractors. Dynamical systems can be very complex, but also as simple as a pendulum or a ball falling down a slope. In these examples, the attractor would be the position where the pendulum is at rest or the lowest point in the slope respectively. In relation to it, the basin of attraction is a concept that determines the area of influence of such attractor. If the ball is thrown out of the slope, the lower point doesn't act as an attractor to it anymore[24].

The pendulum or the ball falling down the slope are examples of dynamic systems exhibiting a fixed point motion –which includes the event of the system staying still. These fixed points can be stable, like in the two examples described, or unstable, like in the case of the ball being at the top point of the slope, ready to fall to another place as soon as something moves it. Another type of motion, with its correspondent attractors, is a limit cycle, a motion in which the system constantly oscillates from one state to another and back to the first. Chemical clocks fit this category. These motions can also be either stable or unstable, but in both cases the repetition of cycles is what defines the state changes of the system. One more degree of complexity is found in quasiperiodic systems, which perform cycles of states that never repeat themselves exactly: "The moon orbits the Earth, which orbits the sun, which, in turn, orbits the galactic center, and son on" (Flake, 1998).

All of these systems, no matter how complex, are deterministic and fully predictable if enough knowledge is obtained of them. They are ordered structures that can be fully accounted for. All dynamic systems were, up to the second part of the twentieth century, though to be of these kinds. Any unpredictability was attributed to lack of knowledge,

---

[24] See the glossary on Appendix A for a more detailed description of these concepts.

as Laplace's famous claim illustrated[25]. But in the 1960s, starting with Lorenz's discoveries in studying weather systems, Chaos Theory was developed, which explained yet another kind of motion: an infinite-period limit cycle which, despite being deterministic –not stochastic– and predictable for short periods of time, was unable to perform long term predictions. Thus, while classical science was built upon the believe that any perceived indeterminacy in nature was due to a lack of knowledge, chaos theory came along with a "striking discovery: Simple deterministic systems with only a few elements can generate random [unpredictable] behavior. The randomness is fundamental; gathering more information does not make it go away. Randomness generated in this way has come to be called chaos" (Crutchfield et al., 1986).

A classical example of a chaotic system is the weather system. It was indeed studying the complexities of weather behavior when Edward N. Lorenz, in 1962, discovered the first example of what came to be called a strange attractor. Among other things, this implied the existence of what is popularly known as the butterfly effect, or Sensitive Dependence on Initial Conditions. The idea here is that there are certain systems that have an intrinsic degree of indeterminacy which makes long term predictions impossible. Added to this, any error or lack of accuracy in the initial measurements or calculations can add up to a very distorted prediction. Hence the 'sensitiveness' of the dependence on these initial conditions in chaotic systems: "Any effect, no matter how small, quickly reaches macroscopic proportions. That is one of the basic properties of chaos" (Crutchfield et al., 1986).

This intrinsic indeterminacy is of major importance in terms of how science is understood. It is not anymore that the unknown is so because it is not yet known. This is an unknown that is not fully knowable: "The existence of chaos affects the scientific method itself. The classic approach to verifying a theory is to make predictions and test them against experimental data. If the phenomena are chaotic, however, long-term predictions are intrinsically impossible" (Crutchfield et al., 1986). As stated in chapter 1, most of the scientific efforts up until the second part of the twentieth century deploy reductionism. Complexity Sciences, Chaos Theory and Emergence in one degree or another, contest the validity of reductionism.

The reductionist approach is to break down any system into parts in order to study them in isolation. Once the parts are understood, the system can be understood as a sum of the parts. Complexity Science does not deny the value of this approach, but it does show that it is inadequate in certain circumstances. When the scientist is faced with complex dynamic systems there is an important degree of indeterminacy that renders the reductionist approach invalid. This indeterminacy can come in two forms: Sensitive

---

[25] Laplace claimed that "given for one instant an intelligence which could comprehend all the forces by which nature is animated and the respective positions of the beings which compose it … nothing would be uncertain, and the future as the past would be present to its eyes" (Laplace, cited in: Flake, 1998).

Dependence on Initial Conditions and emergent phenomena (Solé and Goodwin, 2001). Therefore, Sensitive Dependence on Initial Conditions and emergence are different kinds of indeterminacy. The first is characteristic of strange attractors and, as discussed in chapter 1, it is epiphenomenal. As the following section explains, emergence in the context of dynamic systems is, mostly, synonymous with self-organization. One of the main issues here is "whether one can predict the behavior of a system of many interacting components when the properties of the individual components themselves are understood" (Solé and Goodwin, 2001: 13). The response of Complexity Sciences and the advocates of emergence to this question would be negative.

3.2.3 Emergence as Self-Organization in Dynamic Systems

Dynamic Systems literature is one of the cases were emergence is most clearly identified with self-organization, and so it is adequate to label it as self-organization emergence in the sense that the term is proposed here. It is a context in which emergence is not necessarily related to novelty. In fact, this self-organized order is an emergent property that can actually be expected in such systems when certain conditions are given (Kauffman, 1993). The idea is that in complex dynamic systems, where there exist a large number of coupled elements, self-organization emergence represents the spontaneous appearance of "enormously ordered dynamic behavior" (Kauffman, 1984). The elements, through their interactions, form patterns of behavior without any agent controlling the process on its own: "The system organizes itself, but there is no 'self', no agent inside [or outside] the system doing the organizing" (Kelso, 1995).

J.A. Scott Kelso develops the idea of emergence in his own domain, which is the study of the brain and of human behavior, theorizing on self-organizing processes and examining not how molecules or particles are in isolation but in how they interact with each other. "When we talk about patterns, we step back from things themselves and we concentrate on the relations among things" (Kelso, 1995: 3). In this context, emergence is a central concern in his work, as he considers it to be an important feature of complex systems in general: "Emergent properties are a significant feature of all complex systems in nature" (Kelso, 1995: 2).

In Kelso, these processes of self-organization are understood through the notion of 'control parameters' or collective variables (which are basically equivalent to the attractors mentioned above). Such parameters can have different values at different iterations of a process, and they are uncontrollable in that sense. But at the same time they are what gives order to an otherwise disordered process. One simple example of a control parameter would be the direction in which a group of particles in a heated liquid roll in boiling. If the liquid were to be sliced at one particular point, one could observe how this direction can either be clockwise or counter-clockwise. There is no way to know in advance in which direction a group of particles will roll, but once the motion starts the particles organize around it –they are ‘attracted’– and the order is formed. And the overall process is indeed predictable.

With this example we can see how simple self-organization can deal with emergent properties unproblematically. Here the process can be surprising to a new observer. But once the process is understood, there is no more mystery to it. Rather, it is an inevitable degree of indeterminacy which is perfectly consistent and coherent with the non-reductionism approach. Still, the overall behavior is not explained entirely looking at the parts in isolation. The system must be taken into account as a whole in order for the process to be described: “We make it intelligible by recognizing how it is consistent with lower-level properties and by finding appropriate mathematical descriptors. But in doing this we don’t ‘reduce’ a whole to the properties of its parts and their interactions” (Solé and Goldwin, 2001: 18).

As noted by Jaegwon Kim, this is a mereological conception of reality. There is a separation of levels which is fundamental in the understanding of the emergent processes, since it is thought these differences between levels in which they are described. Again, we find the key philosophical concepts of emergence, although in the context of dynamic systems some of them are more easily (and happily) explained. This is the case of ‘downward causation’[26]. Within the theory of Dynamic Systems, downward causation is a consequence of the order parameters or attractors. These are created by the coordination between the parts and made apparent at the macro levels. And in turn, they influence the behavior of the parts. There is here a circular causality which is not problematic. The order parameters are simply de relevant degrees of freedom of a behavioral pattern (Kelso, 1995). In this context, these collective variables are just a part (important, nonetheless) of the understanding of the levels of organization, which start “with the knowledge of basically three things: the parameters acting on the system (which are sometimes equated with the term ‘boundary conditions’), the interacting elements themselves (a set of primitives[27]), ant he emerging patterns or modes (cooperativities) to which they give rise” (Kelso, 1995: 18).

3.2.4 Crutchfield’s Intrinsic Emergence

In (Crutchfield, 2008), James P. Crutchfield introduced the concept of intrinsic emergence to indicate processes in which whatever emerges from the interactions of the system’s constituting elements (the parts) becomes important for the overall behavior of these elements as a group. It is essentially the same idea as downward causation: the attractors influencing the behavior of the system’s elements that generated it. It was

---

[26] See chapter 1 and chapter 2 for detailed descriptions of the term.

[27] The notion of primitives will acquire a central role in Cariani’s work. See chapter 5.

introduced in an attempt to go beyond the differentiation between emergence and Sensitive Dependence on Initial Conditions. Crutchfield emphasizes the role of the observer for a generic understanding of emergence, but not to find what he identifies as intrinsic emergence. Pattern formation alone is not enough to explain this kind of emergence, as these patterns can be important only externally, to the observer (or to whom would interact with the system as a whole): "In the emergence of coordinated behavior, though, there is a closure in which the patterns that emerge are important within the system. That is, those patterns take on their 'newness' with respect to other structures in the underlying system. Since there is no external referent for novelty or pattern, we can refer to this process as 'intrinsic' emergence."

He distinguishes three notions, or levels, in which intrinsic emergence would be the most elaborate:

1. "The intuitive definition of emergence: 'something new appears';
2. Pattern formation: an observer identifies 'organization' in a dynamical system; and
3. Intrinsic emergence: the system itself capitalizes on patterns that appear."

There must be a process in order to arrive at intrinsic emergence. Crutchfield points at this as the main difference between emergence and discovery. The latter, a fundamental concept of modern science, is essentially atemporal. In contrast, on accounting for emergence, the process, and therefore time, plays a central role: "Emergence is meaningless unless it is defined within the context of the processes themselves; the only well-defined notion of emergence would seem to be intrinsic emergence" (Crutchfield, 2008). This relates to second order cybernetics and also with Bergson and Prigogine. Discovery belongs to what Bergson labeled as the 'cinematograph method', although not necessarily in relation to the idea of the division of reality into frames. It is, in any case, a non durational account of reality.

Emergence, instead, is intrinsically related to duration, and therefore to Bergson's becoming or Prigogine's irreversible processes. Another connection that is relevant here is to Mark Bedau's importance on the concept of simulation in order to discern whether or not a particular phenomenon is emergent. For Bedau, the only way to discern this is through observation of the system in motion (either in reality or in a simulation). Thus, duration is for him of paramount importance, too. As discussed in the previous chapter, Bedau's formulation is problematic due to the subjective implications of the necessity of duration[28]. In any case, it is sufficient here to state that he considers that duration is a crucial element in the understanding of emergence.

---

[28] See section 2.7.2.

Finally, since all the argumentation is within the concept of systems and processes within systems, the link to information and computation is, for Crutchfield, also clear: “A process undergoes emergence if at some time the architecture of information processing has changed in such a way that a distinct and more powerful level of intrinsic computation has appeared that was not present in earlier conditions” (Crutchfield, 2008).

Once the basic notion has been established, Crutchfield distinguishes among four levels of mechanics, each of which would build on the previous one, to explain how intrinsic emergence would be found on the last one. These levels are: deterministic, statistical, computational and evolutionary mechanics. Deterministic mechanics is linked to dynamical systems theory, while statistical mechanics is related to probability theory. It is in this second level where the emergence of chaos theory (self-organization) is found: “Statistical mechanics is engendered by deterministic mechanics largely due to the emergence of irreducible uncertainty.” Computational mechanics, linked to the theory of structure for statistical mechanics would allow for evolutionary mechanics, which he links to what he calls “dynamical theory of innovation”, in respect to the appearance of novelty. Here is where his most powerful version of emergence, ‘intrinsic emergence’, is found: “Evolutionary mechanics concerns how genuine novelty occurs. This is the first level at which emergence takes on its intrinsic aspect. Building on the previous levels, the goal is to delineate the constraints guiding and the forces driving the emergence of complexity” (Crutchfield, 2008).

Crutchfield is here interested in the role of emergence both as a self-organizing process and as a generator of novelty. Very much along the lines of Peter Cariani’s emergence-relative-to-a-model, and of second order cybernetics, too, he insists in placing the observer inside the group of elements to be taken into consideration. If this is done, and a detailed account of what happens in the system can be given, then the answer to his paper’s title *Is Anything Ever New?* is positive: “I would answer ‘most definitely.’ With careful attention to the location of the observer and the system-under-study, with detailed accounting of intrinsic computation, with quantitative measures of complexity, we can analyze the patterns, structures, and novel information processing architectures that emerge in nonlinear processes. In this way, we demonstrate that something new has appeared” (Crutchfield, 2008).

3.3 Emergence and Artificial Life

Simon Penny (2010) summarizes the approach and lines of work of Artificial Life, a discipline which has produced evolved artwork, chatbots, robotic work, or crowd behavior (such as the flocking behavior) among other things: “The term ‘Artificial Life’ arose as a descriptor of a range of (mostly) computer based research practices which sought, among other things, alternatives to conventional Artificial Intelligence methods as a source of (quasi-) intelligent behavior in technological systems and artifacts[29]. These practices included reactive and bottom-up robotics, computational systems which simulated evolutionary and genetic processes, and are range of other activities informed by biology and complexity theory. A general desire was to capture, harness or simulate the generative and ‘emergent’ qualities of ‘nature’ - of evolution, co-evolution and adaptation. ‘Emergence’ was a keyword in the discourse” (Penny, 2010).

As will be discussed in the following chapter, Artificial Life has a clear antecedent in the work of some of the cyberneticians that, decades before the descriptor was used, anticipated many of the concerns and experimental frameworks of the Artificial Life field such as autonomous agents or adaptation. As a field, Artificial Life started in 1987, during a 5 day workshop in the Santa Fe Institute, New Mexico, a center dedicated to the study of Complexity Sciences (Waldrop, 1992). The workshop was organized by Christopher Langton, an early spokesman who several years earlier had been fascinated with computer simulations such as the Game of Life, and who found in the Santa Fe Institute a suitable environment in which to present his ideas and invite fellow researchers to build up a scientific community. The group was highly interdisciplinary, and mostly centered in computer simulations that modeled natural environments and behaviors. It was a field which “premised on a perpetual reevaluation of its founding category –‘life’– and of the experimental and theoretical implications of its methodology –computer simulation” (Helmreich, 2001) [30]. The simulations with which they modeled life were constructs that, as they expected, would help understand and make explicit the existing cultural constructs around the notion of ‘natural’ life (Helmreich, 2001).

3.3.1 Christopher Langton

There are few disciplines in which a single leading figure is so clearly distinguishable as Chris Langton in Artificial Life. He invented the name and even the discipline itself

---

[29] See section 6.6.

[30] Stefan Helmreich conducted during the 1990s an anthropology field study of Santa Fe’s Institute Artificial Life scientists. He noted how, very often, they described themselves as fathers of their digital creatures or as microcosmic gods, and how they even projected sometimes colonial imagery in their accounts of creating and ordering digital life in the virtual space (Helmreich, 2000, 2001).

when he went to Santa Fe and organized the first workshop, hoping for a good turnout, which was what indeed happened (Waldrop, 1992). Langton understood Artificial Life as an interdisciplinary field which “as a whole represents an attempt to increase vastly the role of synthesis in the study of biological phenomena” (Langton, 1995). He held a strong conviction that the synthetic modeling of life would become a very powerful tool to understand life itself. The notions of artificiality, life and emergence are deeply intertwined in his work, which was developed around the study of cellular automata (Langton, 1984; 1986; 1990).

Langton worked extensively in cellular automata studying and simplifying Von Newnan’s self-reproducing models (Langton 1986) and the possibilities of such structures to transmit, store and modify information (Langton, 1990). One of the main things he hoped to achieve with such work was to understand the very origins of life as an emergent phenomenon: “Biochemistry studies the way in which life emerges from the interaction of inanimate molecules. … We have looked into the possibility that life could emerge from the interaction of inanimate ‘artificial’ molecules” (Langton, 1984). Here life is that something that sustains itself as a process that exists on top of these inanimate molecules (natural or artificial) in an autopoietic process: “The collections of inanimate molecules … constitute living organisms [that] interact with each other to maintain and perpetuate the living state” (Lehninger, cited in: Langton, 1984). Therefore, life is an emergent autopoietic property of such natural and artificial systems.

Cellular automata are dynamic systems which, for Langton, can be described as logical universes with their own local physics. Such universes are populated by objects controlled by some function that serves as a rule, and an initial state. Once this has been prepared, the simulation is run in order for the researcher to observe the resulting global behavior in it (Langton, 1984). These global behaviors, according to Langton, are synonymous of emergence: “Any behavior that appears on scales larger than that of a single cell will be ‘emergent’ behavior”. Therefore, he doesn’t necessarily link emergence with novelty. Rather, his idea of emergence is the same as was described in the first chapter regarding the example of Conway’s Game of Life. It is a simplified model of self-organization. The cells are given a set of rules, and from these rules global patterns emerge, as if the cells organize themselves in order to create the patterns. Like in the ant colony, the global behavior is not specified by anyone, but rather appears as the individual agents interact with each other and with the environment.

This illustrates the essence of the bottom-up approach of Artificial Life, which aims at generating complex behavior not directly, but through the multiple (and simple) interactions among a series of agents, through the process of self-organization. In such an approach, a stratified model of reality is inherent. Agents are considered to be at an inferior level than the global behavior of the group, and this in turn allows for terms such as supervenience and downward causation to enter the discourse.

This approach is so deeply assumed by Langton that he actually proposes to interpret the term ‘artificiality’ in accordance with it. Whilst the literal meaning of artificial is ‘made by man, rather than by nature’ (Langton, 1988, 1995), he offers a more complex definition, as formulated by Herbert Simon: “Artificiality connotes perceptual similarity but essential difference, resemblance from without rather than within. The artificial object imitates the real by turning the same face to the outer system… imitation is possible because distinct physical systems can be organized to exhibit nearly identical behavior… Resemblance in behavior of systems without identity of the inner systems is particularly feasible if the aspects in which we are interested arise out of the ‘organization’ of the parts, independently of all but a few properties of the individual components” (Simon, cited in: Langton, 1988).

Therefore, the artificiality of Artificial Life lies in the modeling, not of the system’s components in particular, but of the global behavior. That is, cellular automata are, for Langton, not only mathematical games but legitimate abstractions of the emergent processes in nature (in the sense of self-organization emergence): “Artificial Life studies natural life by attempting to capture the behavioral essence of the constituent components of a living system, and endowing a collection of artificial components with similar behavioral repertoires. If organized correctly, the aggregate of artificial parts should exhibit the same dynamic behavior as the natural system” (Langton, 1988). This is, according to Langton, the main concern of Artificial Life: to emulate the behaviors observed in natural life through a bottom-up approach, constructing systems of artificial agents that will be able to exhibit “’essentially’ the same as some behavior exhibited by a natural system.”

3.3.2 Wolfram’s Cellular Automata

Langton’s work was being developed in parallel to that of Stephen Wolfram, who was also studying cellular automata, and whose writings would become of paramount importance in the field. Wolfram, too, understood the importance of cellular automata as models to investigate self-organization and, thus, emergence understood as such. Like Langton, Wolfram saw them as models for dynamic systems –and as dynamical systems themselves–, although he wasn’t particularly interested in their applications in Artificial Life or in respect to biology, but as mathematical objects of interest in themselves: “Cellular automata are mathematical idealizations of physical systems in which space and time are discrete, and physical quantities take on a finite set of discrete values” (Wolfram, 1983).

One of his major achievements was to classify the types of one dimensional cellular automata into four categories (Wolfram, 1984; 2002; Flake, 1998). In one dimensional cellular automata cells are arranged in the following manner:

- Cells can only have one of a series of infinite states –often only two for the sake of simplicity.
- The cells are arranged in a one dimensional grid, so they can have only either one neighbor –those at the edge– or two.
- At each time step, each cell will compute its new state depending on its local neighbors, according to a predefined rule.

The first and third conditions are in fact common to all cellular automata. It is the second condition what defines it as one dimensional. Game of Life is an example of a two dimensional cellular automata. This one dimensional characteristic makes it very easy to visualize the evolution of the system. It can be done by drawing vertically the progression of the cells over time. Typically, the top row will be the original state and each row below will represent the evolution of the cells' states at the subsequent steps.

What Wolfram discovered was that these automata exhibited, in all cases, one of four behavior, which he dubbed Class 1 to 4 automata, depending on what kind of attractor conditioned its evolution in time. Class 1, 2 and 3 automata contained attractors that were "roughly analogous respectively to the limit points, limit cycles and chaotic ("strange" attractors) found in continuous dynamical systems" (Wolfram, 1984).

Cellular automata of class 1 are those which, starting from a random distribution of black and white cells, will quickly reach a full stop in the evolution of the system, by each cell going permanently either to black or to white. This final state is their fixed point attractor. Almost all initial conditions, in such systems, lead to the same final state.

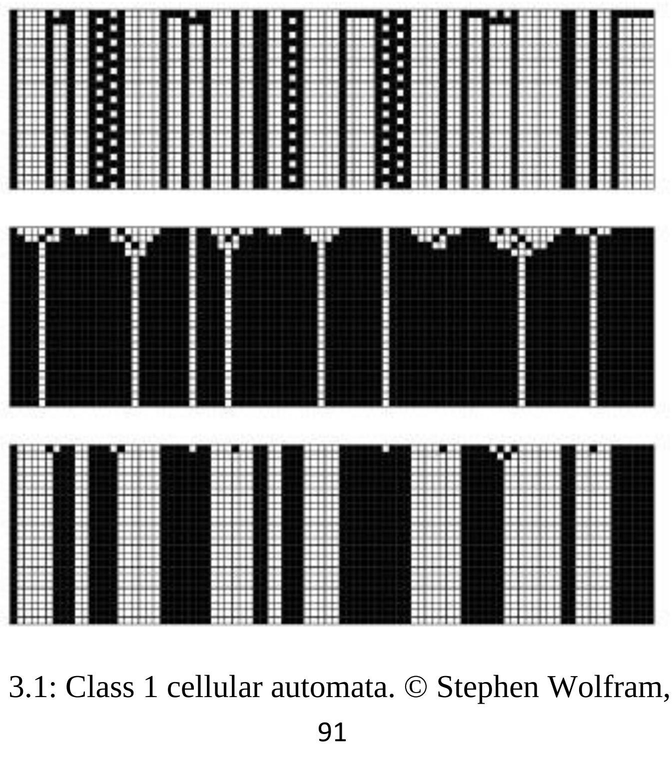


Figure 3.1: Class 1 cellular automata. © Stephen Wolfram, LLC.

Class 2 automata will, instead, find that some of the cells get locked into a periodic cycle of change which will run ad infinitum. Thus, after a few steps from the random start, the system will exhibit what graphically appears as a continuing vertical pattern towards the bottom of the image.

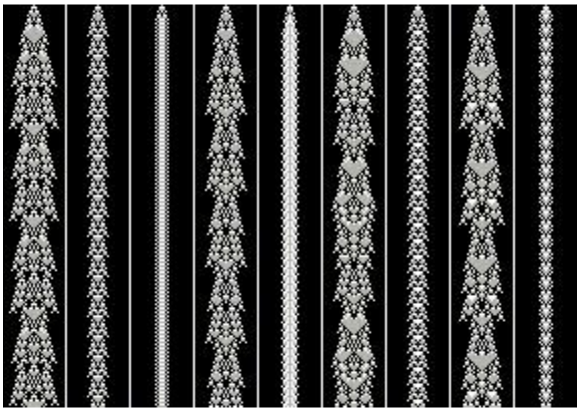


Figure 3.2: Class 2 cellular automata. © Stephen Wolfram, LLC.

The third class is governed by chaotic attractors. Therefore, they are systems which never settle into a stable state. Cells keep changing quasi-randomly, with triangles and small structures eventually appearing. However, what is observed is, in fact, chaotic behavior. These systems exhibit Sensitive Dependence on Initial Conditions (Wolfram, 2002), but here the initial conditions are very easy to repeat, so it is very easy to verify how the system is, in fact, deterministic.

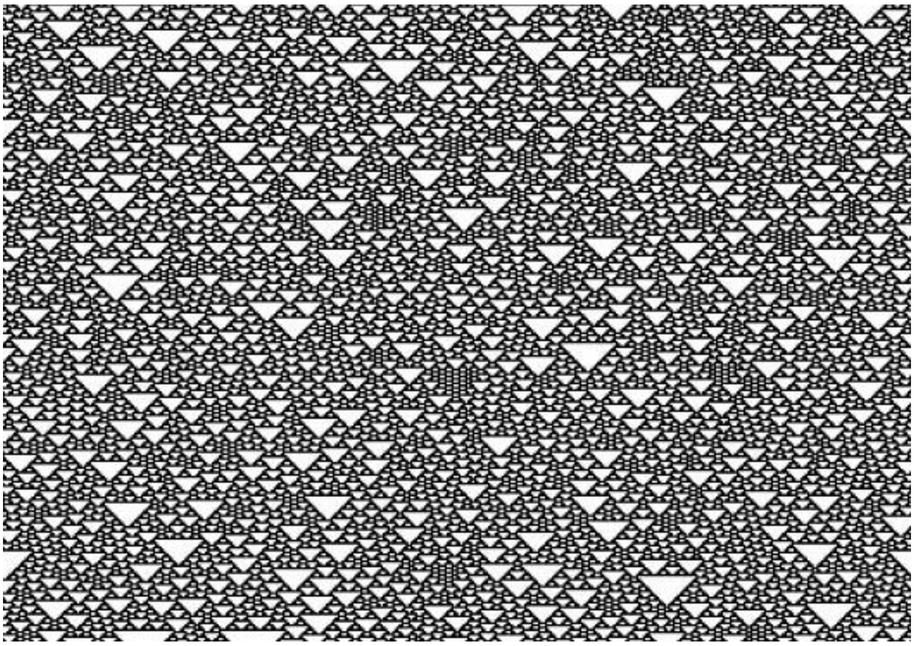


Figure 3.3: Class 3 cellular automata. © Stephen Wolfram, LLC.

Finally, class 4, which “involves a mixture of order and randomness: localized structures are produced which on their own are fairly simple, but the structures move around and interact with each other in very complicated ways” (Wolfram, 2002: 235).

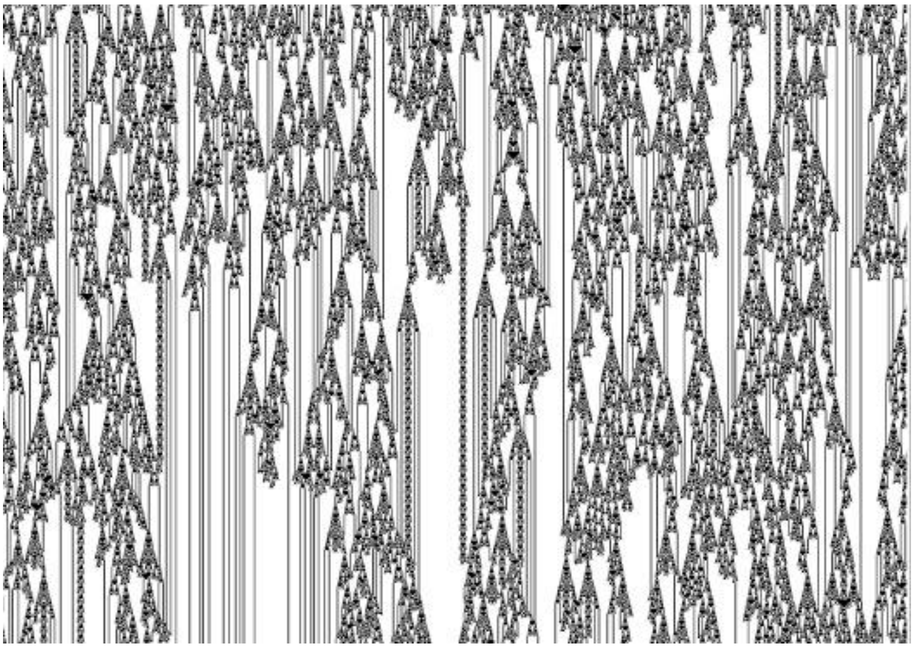


Figure 3.4. Class 4 cellular automata. © Stephen Wolfram, LLC.

Wolfram argues that class 4 cellular automata can be “conjectured to be capable of universal computation, so that their evolution may implement any finite algorithm” (Wolfram, 1984). The idea is that these systems can be viewed as Turing Machines, i.e. as theoretical computers. Their initial configurations represent data, and this data is processed by time evolution. Wolfram notes that some cellular automata, like Game of Life, have been shown to have the property of computational universality, which is in fact one of the main goals under which Von Neumann started the field in the 1960s[31].

Wolfram defines the behavior of these systems in terms that resonate with some of the approaches on emergence reviewed above. Wolfram notes that, while classes 1, 2 and 3 exhibit behaviors that can be algorithmically predicted given the initial circumstances, this is not the case in class 4. In these last cellular automata the value of a particular cell depends on many other variables, and “may apparently be determined only by an algorithm equivalent in complexity to explicit simulation of the cellular automaton evolution. For these cellular automata, no effective prediction is possible; their behavior may be determined only by explicit simulation” (Wolfram, 1984). It is not surprising that this same idea is in Mark Bedau’s account of weak emergence, since actually

---

[31] Further discussion of the issue of the computability in cellular automata is beyond the scope of this study.

Bedau’s work is strongly based on the study of cellular automata, including Wolfram’s work.

3.3.3 Holland: Emergence, Self-organization and Perpetual Novelty.

When Artificial Life started to form as a discipline, John Holland quickly became a central figure. In the 1970s he had invented genetic algorithms (Holland, 1992) which became not only central in (scientific) Artificial Life but also in Artificial Life Art. Genetic algorithms are a computational modeling technique inspired by the theory of evolution. They were developed as a computing technique to find optimized solutions to complex algorithmic problems. As a technique, it emulates a simplified description of genetic reproduction.

Briefly described, genetic algorithms usually start with a population of elements (agents, programming subroutines, etc.) created randomly according to certain predefined parameters (the genotype). Each one of these elements (phenotypes) is then evaluated according to some predefined criteria (the fitness function). Then, the elements which are the most successful are selected to create a new generation of population, which will in turn be evaluated. The process is iterated for a certain number of generations, or until the fitness criteria are met. In the process of creating new individuals, the characteristics (essentially a series of programming variables) of each of the successful phenotypes are recombined. In this manner, the newly created individuals, although different from their progenitors, will inherit the characteristics that made them successful. Each generation, therefore, should be closer to the optimal solution than the previous one. In addition to all these combinations, there is also the possibility of mutation: a random change in some of the variables that define each of the individuals. Mutation is a very important aspect of genetic algorithms, because it allows for the introduction into the system of new possibilities that were unanticipated by the system's designer. If these new randomly introduced possibilities that enter the genetic pool through mutation turn out to be successful (if they help create a phenotype closest to the fitness), then they will enter the evolutionary process.

Genetic algorithms are crucial in understanding Artificial Life and Artificial Life Art. They are the basis of evolutionary programming, which is one of the main ways in which these disciplines have pursued the creation of novelty. Indeed, novelty is one of the key issues that Holland identifies with emergence. His account of the concept, summarized in an excellent monograph (Holland, 1998), looks at the two aspects of emergence to which the present work is dedicated: self-organization and generation of novelty. In this work, Holland gives an account of self-organization emergence. The first definition we find in the book is close to that which is most generally accepted, which states that the whole is more than the mere sum of its parts: emergence is “much coming from little” (Holland, 1998). A definition that works as the opposite of the idea of reductionism, which states that the whole can be strictly and entirely understood through the analysis of its parts.

Holland proposes a framework for understanding emergent phenomena that is based on the creation of multi-agent systems, and the study of the interactions among them, the state of the model in different times (and the transitions between states), and the rules followed by the agents. In this context, he argues for a different kind of reductionism, one which is not based on the analysis of the isolated agents but on the interactions among them and with the environment: "Emergence is above all a product of coupled, context-dependent interactions. … Under these conditions the whole is indeed more than the sum of its parts. However, we can reduce the behavior of the whole to the lawful behavior of its parts, 'if' we take the nonlinear interactions into account." (Holland, 1998: 122). Here, nonlinearity refers to system states that exhibit changes that are not simply incremental. More input does not necessarily represent a proportional increment in the output. The phase transitions of water becoming gas, or the slime mold going from individual cells to slugs are examples of nonlinear behavior.

Holland considers that truly emergent phenomena exist, and that they can be understood according to the parameters just described, but he also admits the difficulty in determining exactly which phenomena are emergent. His idea of emergence is similar to that of Bedau. For Holland, emergence does have an ontological validity, but it is equivalent to Bedau's weak emergence in that it is explainable solely through simulation of the nonlinear interactions of the elements in the micro level. Thus, the nonlinear interaction of the elements must be simulated in order to be properly taken into account, and this implies the necessity of taking duration –time– into account. In this respect, it resonates too with Crutchfield's emphasis on duration (along with Bedau's on simulation), and also with Prigogine and Bergson.

But besides the importance of the self-organization emergence, Holland also emphasizes the importance of novelty in emergence. Indeed, he identifies some emergent processes as generators of perpetual novelty. To illustrate this point, he reviews the work of Arthur Samuel who, during the 1950s, designed and implemented an artificial checkers player with a very Artificial Life-like approach (i.e. bottom up): the player didn't have but very basic pre-specified behavior, and it was through playing that it learned which were the best moves to perform at every given situation (Holland, 1998). Samuel's player was, therefore, not designed to follow a particular set of strategies, but to evaluate each game situation and its own actions in respect to how they changed these situations and, eventually, to whether or not they lead to winning or loosing the game. Through this approach, the checkers player would be constantly changing its model of the game of checkers and of its opponent. With this, according to Holland, it was able to cope with the perpetual novelty of the game: the enormous space of possibilities that exists in it, which otherwise would have been impossible to anticipate.

In his own work, Holland incorporated some of these ideas with genetic algorithms, which are entities that, as a group, are also designed to constantly adapt to new situations and find optimal solutions. Holland has been influential, not only in Artificial Life but also in many other disciplines. As it will be reviewed in the second part of this thesis, his genetic algorithm technique has been extensively used by many of the most prolific artists in Artificial Life Art.

3.3.4 Computational Emergence

The term computational emergence refers to the study of emergence within the context of computer simulations, which as stated above is the playing ground of Artificial Life. One of the most complete attempts to define it was articulated by Assad and Packard. In the same text they mentioned the conceptual link of emergence to the presocratics (Assad and Packard, 2008), they also mentioned the importance of Darwin who implicitly included the idea of emergence in his theory of evolution. This inclusion would be made explicit, they say, by Bergson.

They start with the critique of a distinction made by the father of the Artificial Life field, Christopher Langton, between emergent and pre-specified behavior, which would be related, according to Langton, to Artificial Life and Artificial Intelligence respectively. This distinction, according to the authors, is not so clear “upon further inspection” (Assad and Packard, 2008). Once contextualized, they introduce computational emergence, which, they argue, is too vaguely defined: “Quite generally, emergence usually refers to two or more levels of description. (...) We will call the level on which the model is defined in terms of interactions between components the ‘microscopic level’. The level on which phenomena emerge, as a global property of the collection of components, is the ‘macroscopic level’.”

Within this structure, they argue, in Artificial Life emergence is said to occur whenever an “unexpected macroscopic behavior that is not immediately predictable upon inspection of the specification of the system” appears. According to them, this is “certainly not a rigorous definition”, but it is pervasive in the ALife literature. It is at best an epistemological definition, but the lack of concretion makes it problematic, since what is emergent from one perspective might not be from a different one. After acknowledging this problematic, they obviate giving a new definition, and “in this spirit of relativity” they give instead a scale to determine the emergence of a behavior in a computational system, while recognizing that this does not circumvent the lack of precision problem just mentioned. At the root of the theory is the distinction between weak and strong emergence.

In their scaling system, first of all they clarify what non-emergent behaviors would be. If a behavior can be deduced immediately by inspecting the rules that generate it, then it’s not emergent. Regarding emergent behavior, they distinguish three kinds. Weakly emergent behavior would be that which is "deducible in hindsight from the specification after observing the behavior," whilst more strongly emergent would be that which is "deducible in theory, but its elucidation is prohibitively difficult." They present weak and strong emergence as endpoints of a scale. Therefore, the two categories are presented not so much as a binary distinction but as two stages in a continuum that starts with the non-emergent, passes through weakly and strongly emergent and ends with the maximally emergent behavior. This behavior would be "impossible to deduce from the specification."

The question of what this capability of specification means is unresolved. It is not clear whether or not they are taking an epistemological or an ontological approach to emergence. A similar question to that which was presented to Bedau's weak emergence could be presented here: what is the specification exactly and what are the conditions in which we can say deductions are made?[32]

Be that as it may, once the scale has been defined, the question is: what does actually emerge in these systems were emergent behavior is manifested? To explain this they propose a scheme in which there are three levels, in which each level implies the previous one (but does not imply the successor). That is, each level is a necessary but not sufficient condition of the next one.

- Structure: "Emergence of a patterned structure, either in space time configurations or in more abstract symbolic spaces." E.g. the flocking behavior of the boids or the gliders in Game of Life.
- Computation: "Emergence of computational processing, over and above the computation automatically implemented in the formation of a structure." E.g. Holland's genetic algorithms or Tom Ray's Tierra.
- Functionality: "Emergence of functionality, where functionality is defined in terms of actions that are functional, or beneficial to the microscopic components."[33]

This description of emergence, like Holland's work, covers the two main aspects of emergence. On the one hand, the emergence of structure and functionality is the formation of patterns and the effects these patterns have on the agents at the local level: self-organization emergence and downward causation. On the other, the emergence of computation can be read as combinatoric novelty, a concept which will be discussed in chapter 5.

---

[32] The account of Peter Cariani to emergence is, in part, a response to these kinds of questions and ambiguities.

[33] It is, again, the idea of downward causation, or Crutchfield's Intrinsic Emergence.

3.4 Concluding Remarks

Outside the domain of artificial life, emergence is, within the scientific discourse, mostly understood as a synonymous with self-organization. As stated above, this identification, along with the fact that self-organization is not in dispute as an existing phenomenon, drive the concept of emergence away from the polemics that it encounters, for instance, in philosophy, as seen in chapter 2. Most researchers on complexity, chaos theory or dynamical systems assume that emergent phenomena exist, and they are concerned with how this affects the dynamics of the systems. The scientific interest lies in observation, so it sits in the realm of epistemology, thus taking the edge out of the ontological discussions of the philosophers.

Within Artificial Life, which finds emergence to be such a central concept some clearer efforts to define the term have arisen (e.g., Langton, 1984; 1988; Assad and Packard, 2008; Ronald et al, 1999; Holland, 1995; 1998). For some authors, such as Langton or Wolfram, emergence is again basically self-organization, while others like Holland point out too that this concept can be, and is very often, linked to generation of novelty. His position can be read as having a good balance of both approaches. It was within this context that Peter Cariani, whose work will be reviewed in chapter 5, developed his discourse, in part as a critique of some of these formulations, which were often rather ambiguous. His emergence-relative-to-a-model abandons the self-organization schemes in favor emergence as a generator of novelty, and offers a framework for scientific discussion on emergence that does aim at circumventing the lack of concretion that authors such as Assad and Packard had identified.

4. THE PIONEERING FIELD OF CYBERNETICS

As a discipline, cybernetics covered a vast range of topics, setting the ground for many of later computer-related disciplines that acquired an autonomous status in the second part of the twentieth century such as Human-Computer Interaction or Artificial Intelligence. As it will be explained, the connection between Cybernetics and concepts such as emergence and complexity is clear. One of its main concerns was the study of self-organizing systems, which was also a central interest among the aggregate of disciplines that comprise Complexity Sciences. Self-organization emergence is a moving force of both movements, and one of the main concepts that connects them.

The goal of this chapter is not to cover the history of the cybernetics but to account for its importance as an experimentation and theoretical ground for many of the ideas and practices related to Artificial Life and Artificial Life Art. It starts with a brief overview of cybernetics and of its most relevant interpretations in this respect, then moves to the analysis of the cybernetic devices: experimental prototypes and artworks that some cyberneticians built in order to put in practice their ideas on homeostasis and adaptive devices. The final part of the chapter will review some of the examples of cybernetic art and its role as a context for early computer art.

4.1 An Overview of Cybernetics

Cybernetics was one of the most important and influential scientific discourses during the decades that followed the end of World War II. One of the causes of its early success (and probably of its later decay) was that it covered a much larger ground than most of the scientific disciplines. It was, in fact, conceived as an ur-discipline, and its areas of influence ranged from engineering to social sciences, psychology, or even politics, and as such it was a precursor for more recent interdisciplinary thinking. It appeared in parallel with Information Theory and became a highly influential discipline both for scientific and humanistic disciplines (Gere, 2008). It was a central theory in the technologically-related discourses during most of the 1950s and 1960s. It found a fertile ground in the General System's Theory, which was presented by Ludwig Bertalanffy in 1937 and was at its theoretical basis during the founding years of the discipline (Penny, 2009). It was also the ground for the theorization of control theory, neural networks and for precursor research in digital computing and Artificial Intelligence.

The term 'cybernetics' was coined by Norman Wiener in 1947, in the context of a series of scientific encounters known as the Macy Conferences (Pickering, 2010). Originally, the word was taken from the Greek equivalent of steersman (kybernetes). It was, therefore, etymologically linked to the idea of control, and this is indeed a concept that has commonly been associated with it. In general terms, 'cybernetic systems' refer to (biological and) technological systems which self-regulate. That is, they interact with their environment in terms of feedback loops in order to maintain their activity in the desired terms: "A mechanical system equipped with feedback techniques to control its operations would be cybernetic, that is, self-regulating, without the need of human monitoring" (Murphie and Potts, 2003: 109). This self regulation was called homeostasis, and was regulated by feedback mechanisms. The concepts of feedback and homeostasis were central to Cybernetics.

Control here was understood as directing (steering) an action towards a goal. And this control was exerted "by minimizing error –for any course of action is a series of approximations" (Trask, 1971). These corrections were made using feedback: a back and forth information loop between the system and its environment, in which the first would acquire knowledge on the results of its actions. This information would allow the system to readjust its future actions in order to keep steering towards its goal. This process of control understood as continual adjustment is at the heart of Cybernetics. Cybernetics is indeed the quintessential closed loop, dynamical and processual; systems act on their environment and in turn their behavior is affected by it.

A thermostat connected to a heating system is a classical example of a simple cybernetic system. If the temperature goes below a desired threshold, the thermostat turns on the heater. If it reaches the desired temperature, it turns it off. Any technological system that has some sort of built in auto-regulation can be framed within Cybernetics. Another classic example of such mechanisms is James Watt's governor, a centrifugal valve inside Watt's implementation of the steam engine that opened and closed in order to maintain the a certain speed in the rotary motion of the engine. The governor is generally regarded as the first deliberately contrived feedback device (McNeil, 1990). It represented a significant advancement in technology, as it provided control over the consumption of energy. Through its feedback loop, which allowed it to control how much steam should pass through the valve in order to maintain the desired speed, the Watt's governor allowed the steam engine to become self-regulating (Bernstein, 2002). Because of the indisputable importance of the Watt's engine in the industrial revolution, the fact that it was the first implementation of a feedback system, and its name, the Watt's governor became a perfect example for Cybernetics.

More modern technological developments have also served as recurrent examples. It is the case of the gyro stabilization for aviation and navigation, or the World War II's anti-aircraft gun devised by Wiener and Bigelow, which aided the human gunner with a radar and a statistical model that was able to predict an airplane's most likely position twenty seconds after a ten second tracking concluded (Jerison and Stroock, 1995). It has been noted that such exemplary devices, along with the ideas of control device, and feedback loop, predate Cybernetics as they have a history that in most cases starts at least to the first part of the twentieth century (Mindell, 2002: 8-11).

Starting from the central idea of the feedback loop, Cybernetics soon became a very broad and complex discipline when Wiener, in the 1948's book *Cybernetics; or, Control and Communication in the Animal and the Machine*, "tried to tie together all sorts of more or less independent lines of scientific development: digital electronic computing (then still novel), information theory, early work on neural networks, the theory of servomechanisms and feedback systems" (Pickering, 2010: 3). In his early cybernetic works, Wiener "formulated the idea of information and feedback as the bases of a paradigm for understanding biological, machinic and social processes" (Gere, 2008). This opened the scope of the discipline to virtually any scientific and humanistic ambit.

But while it had an indisputable importance during a significant part of the twentieth century, Cybernetics did not earn a strong position of continuity as a distinct scientific discipline, despite the fact that it continued to strongly influence many aspects of scientific and technological practices. Since the general notions of system, self-organization and feedback are key ideas on technical language, and to the extent that control theory was a product of cybernetics, it did remain a mainstay on engineering. In ecology and sociology, too, mainly through Luhmann in the latter, it remained a strong force. Part of the reason why Cybernetics faded from the first line was that it lost a battle of rhetoric and fundraising with emerging discourses such as Artificial Intelligence and cognitivism, more clearly aligned with the rise of digital computing, although actually very close to the cybernetic discourse. This association of Cybernetics with analog technologies and Artificial Intelligence with digital, among other reasons and issues which are not within the scope of this study, was part of the reason for this decay in importance[34].

Another reason lies in the same motive that made it such a rich movement: the lack of unity in the academic arena, which deprived it to find a solid space in the discipline-oriented universities, and the lack of agreement on fundamentals, causing that every subfield within Cybernetics developed its own theories, methods and languages (Umpleby, 2005). As said above, Cybernetics was an ur-discipline, a theoretical construct which had a broad influence. But also this broad scope was one of the reasons why it was unable to persist as a distinct field of study, and many of its aspects were dissolved into other disciplines, assimilated or temporarily forgotten, only to be rediscovered a few decades later[35]. As Andrew Pickering put it, “the field is not much discussed these days, and the temptation is to assume that it died of some fatal law. In fact, it is alive and well and living under a lot of other names” (Pickering, 2010: 15).

4.1.1 Second Order Cybernetics

Second Order Cybernetics was a label adopted by philosophically inclined cyberneticists during the early 70s, as they felt the need to distinguish themselves from the mechanistic approaches of computer sciences and engineering, which by then were starting to become fully independent from Cybernetics (Heylighen and Joslyn, 2001). Second Order Cybernetics focused, in contrast, on the importance of autonomy, self-organization and cognition, and a very important emphasis on the role of the observer in modeling a system. Indeed, the difference between first and second order has been also phrased as the difference between the Cybernetics of the observed systems and that of the observing systems (Pickering, 2010).

Modeling was an important issue here. It was the recognition that we deal with systems by simplifying them[36], and moreover, that the system itself is an agent on its own right, which interacts with the observer, which is in turn another interacting agent. There is no full separation, therefore, between the observer and what is observed. Most importantly, this undermines conventional notions of objectivity central to modern normal science, which is one of the reasons second order cybernetics was often perceived as eccentric. Here the observer is acknowledged as a part of the system to be studied, and this implies that his or her point of view is also taken into account, as it is recognized that it cannot

---

[34] This entire paragraph is strongly based in comments by thesis director Simon Penny.

[35] According to Umpleby (2005), this is the case of the Santa Fe Institute’s work on complex adaptive systems, whose writers “rarely refer to the early work in cybernetics and systems theory.”

[36] The necessary simplification and abstraction of modeling is what the phrase “relatively abstracted from its surroundings for study” refers to in (Rosenblueth et al., 1943) cited below.

be a purely neutral one. As Von Foerster, who produced the shift towards second-order Cybernetics, put it, objectivity is “the illusion that reality could be observed without an observer” (Von Foerster, cited in: Krippendor, 1995). Similarly, Humberto Maturana asserted that “everything is said by an observer” (Maturana, 1987).

It has also been argued that, in fact, the distinction between first and second order Cybernetics is a rather misleading one. More than a breaking point with a previous approach, Second Order Cybernetics is the continuation of a previously started body of work, often by the same authors. More than a breach, then, there is “a continuous development towards a stronger focus on autonomy and the role of the observer” (Heylighen and Joslyn, 2001). Except the cases of the earliest authors in Cybernetics, such as Norbert Wiener or Ross Ashby, the detailed distinction among first and second-order Cybernetics authors varies amongst interpretations and categorizations and many authors would actually belong to both groups. This discussion is beyond the scope of the work presented here.

4.1.2 Cybernetics and the Black Box

The black box is an important concept in early Cybernetics. It is typically understood as any device, system or part of them that remains opaque to whoever tries to understand how it works. It is “a system whose inputs and outputs are known but with an unknown internal mechanism” (Reichardt, 1968: 14). In sympathy with Behaviorism, the importance of the black box is that you ‘do not need to know’ what is inside the system, just its inputs and outputs. One can and need only to know what comes in and what comes out of it, but not what happens on the inside. The idea was first developed in engineering, but it later on was generalized, mainly under the influence of systems theory discourse in early Cybernetics. Along with this generalization, the black box became a metaphor of any part of a system which remains invisible to whomever is analyzing or interacting with it. Since its early adoption by Cybernetics, it has remained an important idea and metaphor to theorize about interfaces and interactive systems, and notions as important as input or output.

The concept of the black box and that of interface mutually define each other. Black boxes must have interfaces. These are the parts with which the user or experimenter interact with the box. Buttons, dials or sliders as inputs or lights and sounds as outputs are the interface with which a black box is to be analyzed. On the other hand, an interface is what allows one system to connect to another. A computer screen and a keyboard are interfaces, but it also is a USB port or a software library. When one defines an interface, the systems which it connects are, precisely, black boxes. Any zooming in into the analysis of the systems connected is to step away from the interface. The black box, then, is essential in the definition of the interface, and vice-versa. This distinction between black box and interface is equivalent to that of component and interface found in John Haugeland's *Mind Embodied and Embeded* (1998). In this text, Haugeland elucidates this distinction when commenting on Herbert Simon's account on how to approach the issue of the decompositions of systems into subsystems in (Simon, 1996). Simon distinguishes between interactions among systems and interactions within systems, and Haugeland uses this to define three key terms which, he argues, should be understood in relation to each other: component, interface and system. A component, when considered within a system, is the equivalent to the black box. It is a self-contained element which interacts with other components in the system through interfaces. A component is, in principle, replaceable by other components with equivalent functionality. An interface is "a point of interactive 'contact' between components such that the relevant interactions are well-defined, reliable, and relatively simple" (Haugeland, 1998: 213). It is the point where information passes from one component to another. Components and interfaces form systems. A definition of these three concepts, according to Haugeland, allows us to define what level of analysis to apply to a certain aggregate in order to consider it a system, and to determine what function each of these parts does. The fact that the three concepts involve each other, then, does not involve any circularity, since "they collectively involve the further notions of relative independence, simplicity, relevance, and interaction" (Haugeland, 1998: 213).

Arguably, the connection between the idea of the black box and Cybernetics is as old as the discipline itself. Despite the later appearance of the term, the Cybernetics' discourse dates from at least 1943. This is the publication date of *Behavior, Purpose and Teleology* by Arturo Rosenblueth, Norbert Wiener and Julian Bigelow, a paper which is usually considered to be the founding text of the discipline[37]. In the first paragraphs of this text, the authors defined their object of study, a behavioral approach to knowledge, in the terms of a black box-like entity: "Given any object, relatively abstracted from its surroundings for study, the behavioristic approach consists in the examination of the output of the object and of the relations of this output to the input … omitting the specific structure and the instrinsic organization of the object" (Rosenblueth et al., 1943).

In his *Introduction to Cybernetics*, W. Ross Ashby dedicates an entire chapter to the idea of the black box, starting with a reference to its origins in electrical engineering: "The engineer is given a sealed box that has terminals for input, to which he may bring any voltages, shocks, or other disturbances he pleases, and terminals for output, from which he may observe what he can. He is to deduce what he can of its contents" (Ashby, 1957: 86). Black boxes are predictable, and this predictability implies that the relationship between inputs and outputs has to remain the same over time. The experimentation with them is the essence of reverse engineering: the experimenter maps

---

[37] Andrew Pickering argues that this honor should be granted to *Adaptiveness and Equilibrium*, a previous essay published by William Ross Ashby in 1940 (Pickering, 2010).

inputs to outputs in order to understand what happens inside the box. This will allow him or her to recreate (or emulate) it if necessary. By doing so, the engineer creates what Ashby calls a protocol, which contains all the knowledge that can be obtained from this back and forth of the black box's inputs and outputs. According to Ashby, the black box theory is, despite its engineering origin, applicable to a far wider range of activities, related to professional studies but also to far more common examples, such that of a child trying to open a door. The child has a handle as an input device and, without inspecting the inner workings of the system (the door opening mechanism), he or she has to learn how to produce the desired effect (the correct movement of the latch). The infant learns to do so by means of trial and error, just like the electrical engineer does when faced to a black box.

Ashby's point is that, in fact, we are interacting with black boxes all the time: "In our daily lives we are confronted at every turn with systems whose internal mechanisms are not fully open to inspection, and which must be treated by the methods appropriate to the Black Box" (Ashby, 1957: 86). These methods consist, in a nutshell, first in the definition of both the inputs and outputs of the system under examination, and then in the experimentation with those in order to establish the relations among them: the protocol, in Ashby's terms. The goal is to find the regularities and repetitiveness in behavior that will inform the experimenter of the inner workings of the box: "When a generous length of record has been obtained, the experimenter will look for regularities, for repetitiveness in the behaviour" (Ashby, 1957: 89). From these processes, we make inferences and generalizations on how certain interfaces work and what is to be expected from then.

4.1.3 Unpredictable Black Boxes

As said above, in general black boxes need to be predictable and the relationship between inputs and outputs has to remain the same over time. Otherwise the task of the experimenter testing it in order to understand how it works would be an impossible one. In the context of creativity, which is the subject matter of this investigation, however, the black box can be a very useful concept if we allow an element of uncertainty into it. The black box can be here also, as in the Rosenblueth et al. paper, a metaphor for any space or aspect which remains unknown. This metaphor is useful in order to understand creativity and emergence in the context of interactive art. By removing the constraining rule, indispensable in the engineering context, that the relationship among inputs and outputs must remain always the same, the door for unpredictability is opened, as systems can be created so that these mappings change over time.

These spaces for uncertainty could be labeled as unpredictable black boxes (Soler-Adillon, 2011). That is, systems which the user interacts with through inputs and outputs that are not fully predictable, and the inside of which is not only unknown but unknowable. The inner workings of such boxes are not reducible to the analysis of its parts and connections upon an eventual opening of the box (or zoom in into the system), as it is the case of the general idea of a black box. This doesn't mean that these systems are totally unpredictable and stochastic in their behavior. The relations between inputs and outputs are neither always unpredictable nor necessarily the processes in its behavior unrepeatable. One can make informed guesses on future behaviors and responses, and test sequences of inputs and outputs. The difference is that the mappings between them would not always give back the same results, as they can change over time.

These unpredictable black boxes are often found within the context of Cybernetics, in the form of adaptive devices. Most certainly not all black boxes in cybernetic discourse are of this kind. In the above cited example, the child doesn't know the insides of the door but these are knowable once the door's handle-latch system is open for inspection. The black boxes that Ashby found everywhere can be of both types, and are actually usually of the first. But what is important here is that they can be viewed as a central idea on the understanding of the world[38] upon which Cybernetics is built.

Another characteristic of the cybernetic approach to the unpredictable black box is that it is not only about dealing with these systems, but also about creating them. Indeed, this is something that some cyberneticians did, as will be shown in the following section on the cybernetic devices. The idea is to create something that will appear as a black box to its own creator. That is, a device that will surprise its designer, in terms of its behavior and of the relationships between the inputs it receives and the outputs it produces. In it, even though the designer has thought through, build and programmed the device or system, the relationship between what it perceives (the inputs) and how it responds to it (the outputs) become unexpected at a certain point.

Here the ideas of the unpredictable black box and emergence can be linked. If the result of the unpredictable black box is not mere randomness, in some cases in which the system creates something that was not explicitly built on it by its designer, emergent behavior can be identified. This is not something that a designer of a conventional computational system would desire, but it can be the case in digital art practices. In generative art and especially in Artificial Live Art, it is often sought by the artist to create systems or processes which exceed her expectations. The idea is to do so not through some blind trial and error, but through the creation of the conditions of emergent phenomena to occur: "The basic principle of emergence is that organization (behavior/order/meaning) can arise from the agglomeration of small component units which do not individually exhibit those characteristics" (Penny, 1996). The examples that will be reviewed in the following section, as well as the Artificial Life Art pieces analyzed in chapter 6, can be understood with the idea of the unpredictable black box.

---

[38] The ontology in Andrew Pickering's terms, as will be reviewed in section 4.1.4.

4.1.4 Cybernetic Ontology and Emergence

Andrew Pickering has vindicated the legacy of Cybernetics. He considers that this discipline has been marginalized in the academia, and stigmatized to being too strongly and exclusively related to military and industrial applications. *The Cybernetic Brain* is a monograph published in 2010 which aims to compensate for this. In his previous work (Pickering, 1995; 2002) Pickering had elaborated a theory of how the scientific practice is determined by its own processes and even measuring devices. In it, he argues how science is not so much about representing and mapping the world and articulating knowledge about it, as it is an interplay of human and material agency. A performative idiom, as he labels it, rather than a representational one. He presents his idea in contrast to the traditional scientific paradigm, built on the basis of the reductionism and linearity of the Newtonian paradigm, which is a representational approach, or "a venture into 'epistemology'" (Pickering, 2002).

His alternative is, instead, based on a performative ontology: "a decentered perspective that is concerned with agency –doing things in the world– and with the emergent interplay of human and material agency" (Pickering, 2002). His use of emergence here can be loosely understood as a sort of self-organizing result of the decentered activities that constitute what Pickering identifies as 'the mangle'. The individual and local actions add up to a result that is not specifically intended by any of the agents performing these actions. He uses the term 'temporal emergence' to stress how plans and goals in scientific practice appear over a period of time, as scientific agency advances in its particular actions (Pickering, 1995).

In any case, the issue here is not Pickering's idea of emergence but the differentiation between the representational approach of Newtonian science and the performative ontology on which Cybernetics is based, according to him. Within this context, Cybernetics appears as a discipline paradigmatic of the latter idiom. The discipline as a whole, but especially the English cyberneticians' practice, "is all about this shift from epistemology to ontology, from representation to performativity, agency and emergence, not in the analysis of science but within the body of science itself" (Pickering, 1995).

The performative idiom, according to Pickering, is an ontology in its own right, and a nonmodern one in the sense that it differentiates itself with the reductionist approach of the scientific method (the modern ontology). In contrast to Bruno Latour's characterization of modernity as being determined by the dualism of people and things, the cybernetic approach (Latour, 1993), in which this frontier would be blurred, "stages for us a nonmodern ontology, in which people and things are not so different at all" (Pickering 2010: 18). As he has explained (Pickering, 2008; 2010), this ontology allows Cybernetics to propose an image of the world that is performative rather than representational. He calls it the 'ontological theater', a nonmodern understanding of the world and our relationship with it that implies "a vision of knowledge as 'part of' performance rather than as an external controller of it" (Pickering 2010: 25).

And here, the idea of the black box is useful to understand how this is brought into practice. According to Pickering, Cybernetics proposes a theory of knowledge that is largely built up through a performative relationship with what we can understand as black boxes, and many of them are of the unpredictable kind discussed above. Rather than about control in a classical sense, “the entire task of Cybernetics was to figure out how to get along in a world that was not enframable, that could not be subjugated to human designs – how to build machines and construct systems that could adapt performatively to whatever happened to come their way” (Pickering 2010: 30-31). This is how he understands some of the devices that will be subsequently discussed: Walter’s tortoises, Ashby’s Homeostat or Pask’s Musicolor, electrochemical devices and Colloquy of Mobiles.

In this context, he argues how Cybernetics assumes in fact an ontology that, contrary to the reductionist approach, involves a certain degree of unknowability, as it “tries to address the problematic of getting along performatively with systems that can always surprise us” (Pickering 2010: 23). These are what Stafford Beer labeled exceedingly complex systems. That is, systems which are, unlike simple and merely complex systems, neither predictable nor susceptible to be treated by the methods of modern science and engineering. Exceedingly complex systems, like the interior of the unpredictable black boxes, are unknowable. They are too complex to be grasped representationally, and they change over time, so that future behavior cannot be anticipated trough current knowledge.

This element of surprise, of impossibility of a full anticipation, clearly resonates with emergence in any of its formulations. So if Cybernetics was, as Pickering argues, about that, it was not so much about control, as some literature has portrayed it, as it was about the study of the conditions under which the interaction among different parts of technology could result in these complex, unanticipated patterns of behavior. Some of the devices built by the cybernetic practitioners were clearly built in order to experiment in this direction, and were in this respect anticipative of the Artificial Life Art movement that would build around concepts such as emergence in the 90s, as it will be shown in the rest of the chapter.

4.2 Cybernetic Devices

This section is dedicated to the exploration of a series of experimental devices build by three of the Cybernetics' authors that were active both in their writings and in building of devices to test and put in practice their ideas. The selection is not intended to be exhaustive, but representative of devices that illustrate Pickering's performative ontology of Cybernetics, and which are in turn clear predecessors of the Artificial Life and Artificial Life Art devices.

4.2.1 Ashby's Device: The Homeostat

William Ross Ashby was a medical doctor, psychologist and psychiatrist who developed most of his career within these disciplines, studying and working with the pathologies of the brain until 1961, when he moved to the United States after accepting a position as a professor in the Department of Electrical Engineering at the University of Illinois, were Von Foerster was developing his research (Pickering, 2010). His work on Cybernetics was very extensive and influential. And like other cybernetic authors, his approach to it was conditioned by his own field and interests. His device, the homeostat, was conceived as a model that should help understand some of the inner workings of the brain (Pickering, 2010). Indeed, he extensively described it in a book entitled *Design for a Brain*, were he stated that the properties it exhibited were "of special interest to the physiologist and the psychologist" (Ashby, 1976: 100). Other accounts give the homeostat a wider scope of ambition. According to Ashby's colleague Grey Walter, "it was created to study the mechanism whereby an animal adapts its total system to preserve its internal stability in spite of violent external changes" (Walter, 1950). In any case, it was conceived by Ross Ashby in order to experiment with systems that would find ways to stabilize themselves. He started working on the first experiments towards it in 1946 and it would become the center-piece of his cybernetic work for the following years (Pickering, 2010).

The homeostat aimed to model the behavior of the brain as it adapted to the changes of the environment. But when it was presented at the Macy conferences there were discussions around what it did exactly model, if anything. For most participants in the conferences, the value of the machine resided precisely in its capacity to model something, whilst Asbhy was more interested in the homeostat per se and its stabilization mechanisms. According to John Johnston, these differences in criteria might be explained with Pickering's notions of the representational idiom of science, which would guide the approach of the majority of participants, and the performative idiom that these early cyberneticians were inaugurating (Johnston, 2008). The father of Cybernetics, Norbert Wiener, described the homeostat as the "brilliant idea of the unpurposeful random mechanism which seeks for its own purpose through a process of learning" and qualified it as "one of the great philosophical contributions of the present day" (Wiener, 1989: 38). The homesostat was, indeed, an unpurposeful machine in regards to the classical sense in which machines have purposes. It was built in order to interact with other homeostats, and to experiment with homeostasis by doing so. Ashby's goal was to build a self-regulating system, and he was successful in doing so.

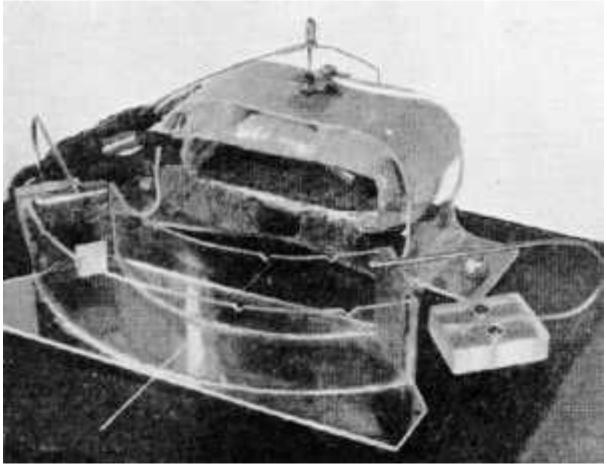


Figure 4.1: The top part of a homeostat unit (Ashby, 1976).

The homeostat was an electromechanical device which converted electrical inputs to electrical outputs. It consisted of four identical units (Ashby, 1976). In each unit, the input passed through a coil that was inside the machine, which generated a magnetic field that exerted a torque on a needle located at the top of the machine, which rotated to one direction or another. The position of the needle was to be interpreted as the main variable indicating the state of the homeostat (Pickering, 2002; 2010). The unit could either be in a stable position, with the needle resting in a central position, or in an unstable one, with it going out of range (a deviation of above 45 degrees from the resting position), which provoked the circuitry to randomly switch to another state (another set of values in the circuitry) and seek for stability in it. If it didn't find it, is switched again, and so on[39].

Isolated, a homeostat unit did close to nothing. It was when connected with others that a system of dynamic feedback loops started to build up, thus forcing each unit to look for its own stability within a changing (dynamic) system, were each unit was affecting all others. As a group of several units, it was a way to prove what Ashby labeled 'ultrastability': a consequence of the process of adaptation of each unit to the system (Ashby, 1976). When the four units came to a stable state simultaneously, ultrastability had been reached.

---

[39] For a detailed account of the homeostat workings and circuitry, see Ashby (1960) and Pickering (2010).

Figure 4.2: Four units forming the homeostat (Ashby, 1976).

The homeostat as a system, as Wiener pointed out, lacked a purpose beyond its own search for stabilization. What was, then, the motive in its building? As Pickering points out, the homeostat was a proto-brain designed to model adaptive behavior, “a model of the brain as an adaptive controller of behavior” (Pickering, 2002). The search of each unit in the homeostat for a stable setting within the system is a process of adaptation. Each unit struggles, though its open ended search of possibilities, to find a sate which will be stable as it interacts with the other units, whilst the other units are performing the exact same search. Ultrastability, the moment in which all of the units come to a stable state (i.e., when the system attains homeostasis), can be understood here as the result of a self-organizing process and, therefore, as an emergent property.

Despite the limited number of units in Ashby’s description, we can imagine that the homeostat could theoretically grow its unit number and the system would behave in the same manner. Like in the case of the Craig Reynold’s Boids, presented in chapter 1 as a case study, each unit interacts solely with its neighboring units. It is the sum of all the interactions what causes ultrastabilty to appear, like the flocking behavior appears in the simulated boids. No central command is directing the operations and no individual unit has any advantage or stronger influence on the others. The process of ultrastabilization as a whole is a result of self-organization.

In this respect, Ashby succeeded in creating a machine that could go beyond a fixed repertoire of stimulus-response reactions (Johnston, 2008), thus creating a system that was open-ended enough to constantly change responses. Despite the fact that the homeostat was itself not interactive respect an external user, it can be argued that in terms of internal interactivity (i.e. regarding one unit in respect of all the others) it went beyond the database paradigm, and moved towards a more complex approach of creation of non-predefined behaviors.

4.2.2 Grey Walter’s Tortoises

William Grey Walter was, like Ross Asbhy, a British Neuroscientist and a cybernetitian whose interests and career orbited around the study of the human brain. In the 1930s he became a specialist in the then emerging field of electroencephalography, and this remained a central activity for him for the rest of his career (Pickering, 2010). In studying the functioning of the brain, Walter soon became aware that it posed a problem of complexity. The brain was a system which escaped the traditional (reductionist) approach in the sense that, in it, the atomic understanding did not translate, by accumulation, into a global one. It was what another of his colleagues, Stafford Beer, would take for a typical case of an ‘exceedingly complex system’ (Pickering, 2010). As he described it: “one took an anatomical glance at the brain, and turned away in despair” (Walter, cited in: Pickering, 2010: 41). Building a model of such a complex system, with ten thousand million units to connect, was not a realistic prospect that Walter contemplated (Walter, 1950; Holland, 2003). Instead, he decided on an opposite approach: a device in which “the number of components … was deliberately restricted to two in order to discover what degree of complexity of behavior and independence could be achieved with the smallest number of elements connected in a system, providing the greatest number of possible interconnections” (Walter, 1950).

For Walter, the building of the model was an essential part of the cybernetic approach. According to him, propositions such as ‘reflexive behavior gives an impression of purposefulness,’ or ‘stability can be achieved by negative feedback’ could easily be misinterpreted[40]. Modeling was, in this context, a practice that was aimed at avoiding the possible ambiguities that could lead to such misinterpretations. Modeling consisted in embodying these abstract or theoretical propositions in ‘hardware’ so as to remove the ambiguities of language: “When these propositions are embodied in working models their content is unequivocal and their implications are open to test and verification” (Walter, 1963).

The cybernetic models that Walter defended had three main characteristics. First, they where as simple as possible. ‘Pure,’ in Walter’s terms, in the sense that they applied the principle of parsimony (Ockham’s razor). Secondly, and derived from this simplicity, they were ‘transparent’. The models were build with visible components, each of which had a clearly defined function. The aim of the models was that the scientist could use them to look at a problem, so any embellishment that could get in the way of this objective was discouraged. Thirdly, the models were ‘semantically brittle,’ in the sense that, because of its simplicity and unambiguous design, they did not hide their flaws. As

---

[40] Walter adds: “particularly when translated into a foreign language” (Walter, 1963), in a gesture that is reminiscent of a time when English was not yet considered to be the only valid scientific language, and so English writers were more concerned with translations than they are nowadays.

Walter put it, “when [the model] fails it breaks neatly and does not bend like words do” (Walter, 1963).

At the end of the 1940s Walter built two robotic devices, known as the tortoises, that have to be understood as an example of such modeling practice: a pragmatic material experiment that would serve as the basis for and orderly and practical classification of complex phenomena (Walter, 1963). The devices that Grey Walter built were two electromechanical automata that, by their shape, resembled tortoises. He dubbed them Elmer (for Electro MEchanical Robot) and Elsie (Electro mechanical robot, Light-Sensitive with Internal and External stability).

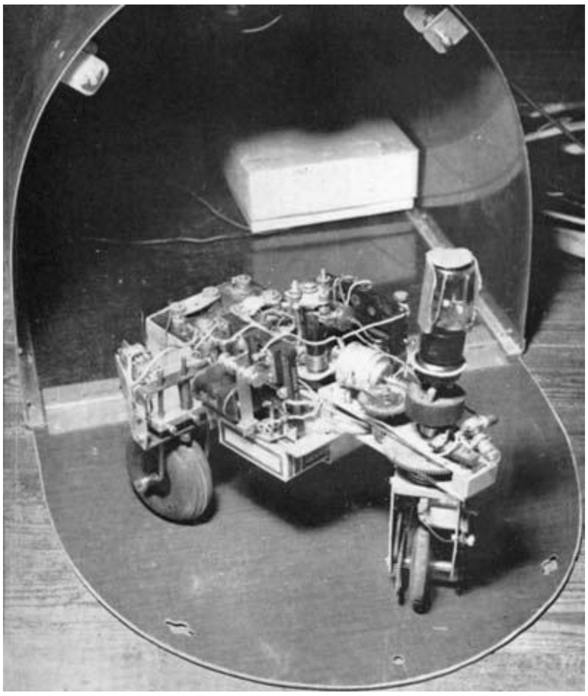


Figure 4.3: A tortoise without the shell (Pickering, 2010).

The tortoises had two sensors, for light and touch, and two effectors, for crawling and for steering. They were powered by batteries that they carried with them, and had a hutch were these could be recharged. The technical details of these electromechanical tortoises were described in detail by Walter himself (1950; 1951; 1953) and by Owen Holland (1997; 2003), who reconstructed one of the tortoises and has defended importance of Walter’s work as a pioneer (an antecessor, in fact) of Artificial Life.

In terms of behavior, Elmer and Elsie’s main goal was to look for a light source. If it was not found, they would wander around the space while a touch sensor allowed them to avoid obstacles trough a very simple procedure: if they couldn’t move forward and detected that they were touching an obstacle, they would perform a series of little bumping movements until the pressure was released, therefore continuing their wandering movement. Along with these movements, a light sensor would be spinning around in cycles, looking for a light source. Once it was found, the whole device would steer and move towards the light, amplifying this action linearly with the intensity of the light source, until a certain threshold was reached. When this happened, the tortoise would saturate and turn away from the light and start over. When their batteries were low, however, the changes in the circuitry would drive them towards the brightest light possible. As it turns out, the hutch that Walter built for them to recharge their batteries was strongly illuminated. Thus, this behavior allowed them to go back 'home' to recharge, and retain full autonomy[41].

From this simple scheme, as Walter himself soon noted, "the behavior of Elmer and Elsie [was] in fact remarkably unpredictable" (Walter, 1950). This unpredictability was intentional in the design. When Walter described his devices, he pointed out that one of his main objectives was to experiment with one of the aspects of animal behavior: "the uncertainty, randomness, free will or independence so strikingly absent in most well-designed machines." The tortoises would exhibit different behaviors as the environment changed and they adapted to it. For instance, when they carried a light source with them, they would interact with each other or with their image on a mirror in manners that resembled the behavior of an animal when it recognizes one of its own kind, as they cyclically approached and avoided each other (or its own image on the mirror).

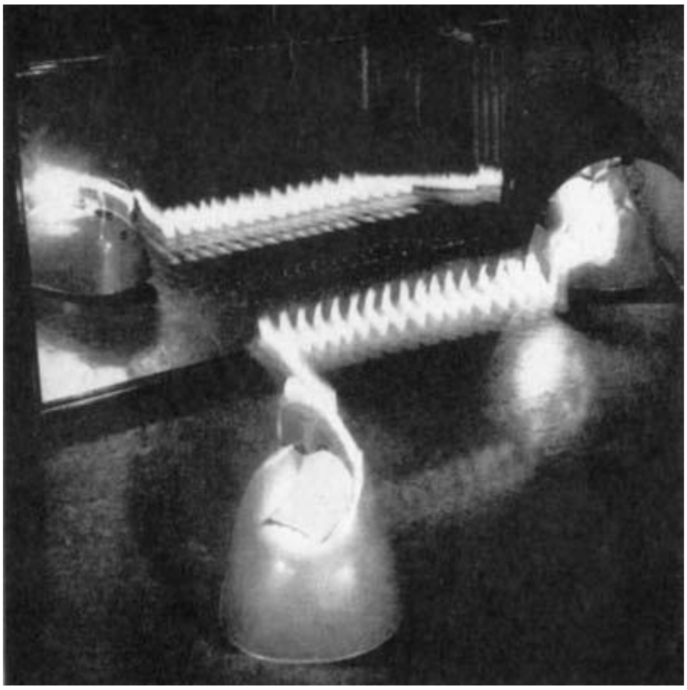


Figure 4.4: A tortoise's path in front of a mirror (Pickering, 2010).

---

[41] It is quite remarkable how this idea took sixty years to become a commodity in the home vacuum cleaning robots popularized by iRobot, the Roomba series, which also proceeds by wandering through space while avoiding obstacles, and returns to its base when it needs recharging. The company iRobot was co-founded by roboticist Rodney Brooks, a major figure in robotics research, who has recognized the influence of Walter's *Design for a Brain*, which he read as a child (LeBouthillier, 1999).

As Walter admitted, his robots were meant to only approximate animal behavior (Burnham, 1968), but nonetheless he was convinced that this resemblance was perfectly valid to regard them as a model of behavior, of the adaptiveness of the brain to the environment or even, as he titled his 1950 Scientific American article: *An Imitation of Life*. In any case, they are a very clear predecessor of Artificial Life and also a landmark in robotics, and the complexity of their behavior can be regarded as emergent (in the sense of self-organization emergence) in respect to the simplicity of the parts involved in generating them. Along these lines, Andrew Pickering describes them as having "emergent properties relative to what Walter had designed into them" (Pickering, 2010: 50). The local interactions of the sensors with the circuitry, as the interactions of the boids or the ants' colony, generate the patterns of behavior that are not explicitly designed, but appear only as the system is set in motion and the experimenter can observe them.

The 'remarkable unpredictability' of Elmer and Elsie's behavior can be read as the appearance of these patterns of behavior that Walter observed and described in detail (Holland, 1997). One didn't know what Elsie or Elmer would exactly do, but could only expect them to behave along the lines of one of the previously observed patterns. According to this, in the cases were a new pattern of behavior is discovered, the idea of emergence as generation of novelty could also be applied. Walter noted, for instance, that he had been taken by surprise by the tortoises' behavior in front of the mirror and with the other tortoises (Walter, 1953). These behaviors would, therefore, be emergent relative to the model of expected behavior that Walter had elaborated with his first observations.

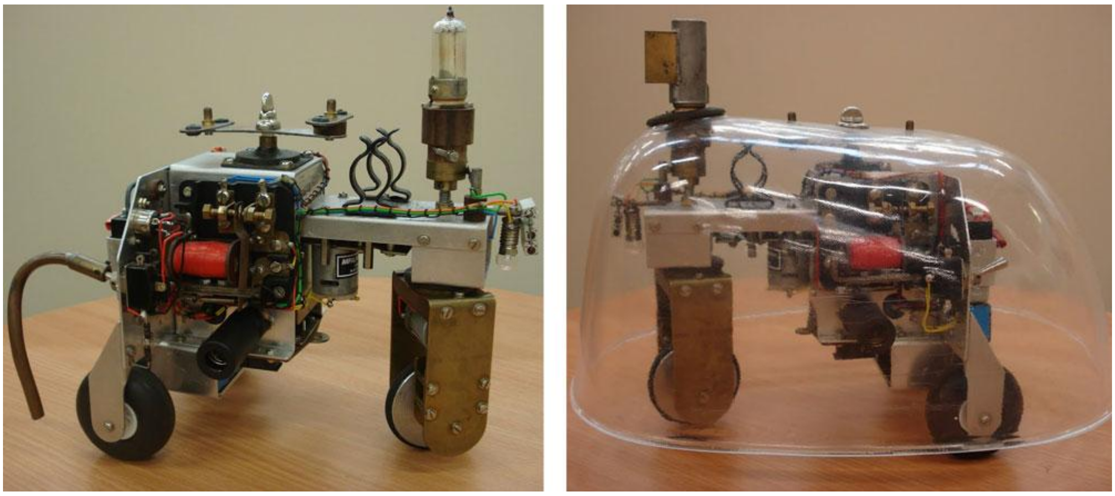


Figure 4.5: In 2008-2009, the Bristol Robotics Lab created Amy and Ninja, two replicas of Walter's tortoises. © Bristol Robotics Laboratory. Used with permission.

4.2.3 Gordon Pask's Devices

The third and last author to be reviewed in this section is Gordon Pask. Born in Derby on 1928, during his career he worked in several universities both in Europe and the United States, but he developed most of his projects in a research organization of his creation called System Research (Pickering, 2010). His interests in cybernetics ran in parallel with an avid interest in theater and with art in general, and this is reflected in his numerous experimental devices, only three of which will be reviewed here. As a whole, all of his devices are a reflection of his particular way of working but also of Cybernetics as a discipline. Pask built his devices much in the spirit with which Walter described the idea of a cybernetic model. They were vehicles not only to test and put in practice his ideas but to develop them. For these cyberneticians there was no separation between theory and practice. This is what Pickering calls the performative ontology (Pickering, 2008; 2010), which drives all of the creations revised here and many others.

4.2.3.1 The Musicolour

Pask's first device, and one with capital importance on his path into cybernetics, was the Musicolour. Home-built at the beginning of the fifties, it anticipated the idea of computer visuals accompanying musical performances. I was inspired "by the concept of synaesthesia and the general proposition that the aesthetic value of a work can be enhanced if the work is simultaneously presented in more than one sensory modality" (Pask, 1971). Neither of the ideas was new, but as Pask notes, in the early 1950s the idea of augmenting sound by light was not yet as overused as it would be in the psychedelic years.

The Musicolour was first demonstrated in 1953. The machine, constructed largely out of war surplus analog electromechanical equipment, was essentially a transducer which, through the input of a microphone, generated a series of visuals from a "predetermined vocabulary of visual symbols; coloured forms which were projected on to a large screen in front of the performer and an audience." As the experiments with it advanced, Pask lost interest in the idea of synaesthesia and the focus shifted towards the learning capabilities of the device (Pask, 1971). The Musicolour became a machine that engaged very strongly with the performer, rather than amused the audience, as musician and machine learned from each other as the performances advanced. This feedback loop "had an almost hypnotic effect upon the performer," whilst, in contrast, it was sometimes disappointing to the audience.

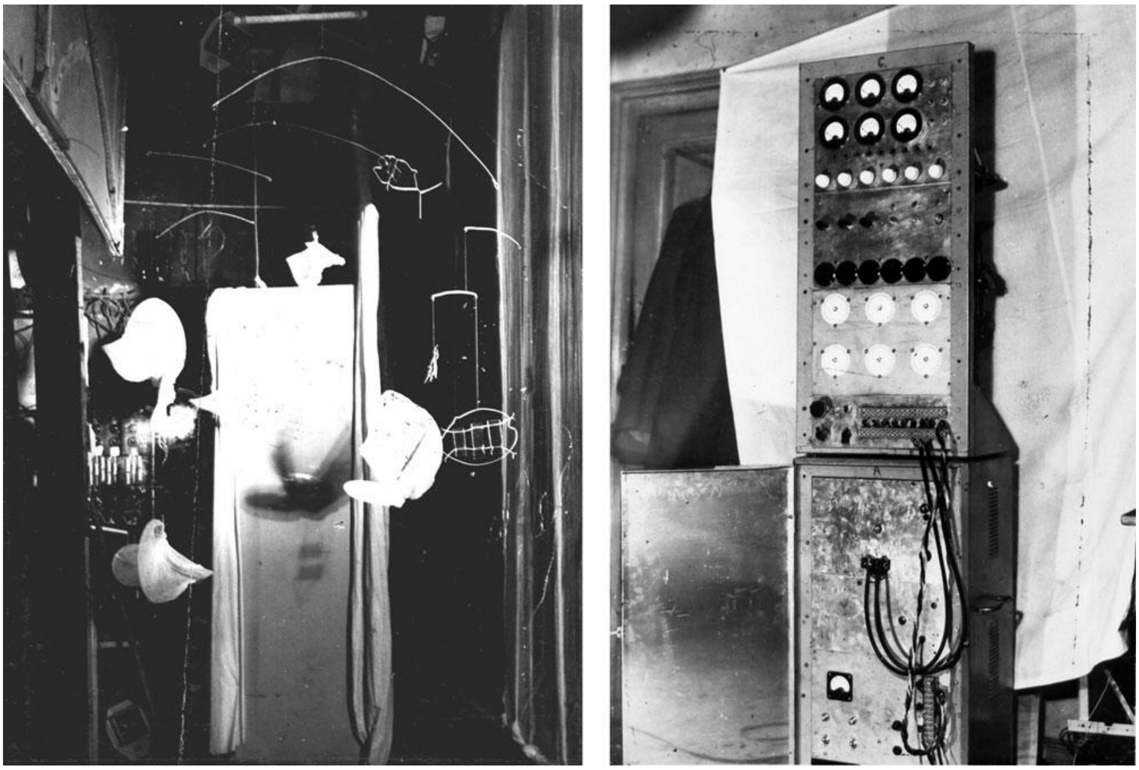


Figure 4.6: The Musicolour's scenography (left) and machine (right). © Gordon Pask Archive at the Dept of Contemporary History, University of Vienna, Austria. Used with permission.

It is easy to understand how these two effects were related. The Musicolour, in its last versions, moved away from the linear relations between sounds and colors. Instead, it exhibited more complex behaviors. Most remarkably, if the performer became too repetitive in trying to provoke a particular response by the system, the Musicolour would cease to respond, as if it got bored with the reiteration (Pickering, 2002). With this, it would encourage the performer to try something new, and so to never repeat his or herself too much. As Pask noted, the effect of this complexity of behavior on the audience was rather negative. Whilst the performer could be aware of what was going on, the audience would perceive too much a degree of randomness in the relations between sounds and lights. The perceived randomness would cause them to lose interest in the device as a real-time generator of visuals interpreted as readings to the music. This fail to interest the public can also be read as a symptom of a pre-Cage audience (Soler-Adillon and Penny, 2014).

But as said above the interest had shifted to the relationship between device and performer. The latter engaged in an experience that Pask describes as a game in which the machine became one with the performer: "He trained the machine and it played a game with him. In this sense, the system acted as an extension for the performer with which he could co-operate to achieve effects that he could not achieve on his own" (Pask, 1971). Thus, the coupling of the performer's actions and the machine's responses created a stability that was unattainable if one of the two parts was missing. In this respect, these stability, which in this case is the performance itself, would be emergent in the same way that the stability in Ashby's homeostat is. Local interactions create a result that is not intentionally mediated by any of the parts in particular. This interpretation would be analogous to that proposed in section 4.1.4: a series of individual localized actions create a result that is not intended specifically by any of the agents involved.

Thus, Musicolour is a case of self-organizing emergence, in which each individual action adds up to a complex pattern of behavior, but it is also very close to the generation of novelty, at least theoretically. If the Musicolour was open-ended enough, the new actions of the performer, could result in non-previously observed responses by the system. In fact, Pask's interest in the Musicolour was as much as on it being a performance object as on it being a model for learning. The devices that followed demonstrate both these two interests. The device's last appearance was in 1957. From them on, Pask concentrated in other projects and never revisited the experiment (Pask, 1971).

4.2.3.2 Learning: The Electrochemical Ear

After the Musicolour was shelved, Gordon Pask started to work on a series of electrochemical computing devices or, as he liked to refer to them focusing on their properties rather than on the materials they were built of: organic computers (Pickering, 2010). Pask's electrochemical computing devices were extremely open-ended systems with which he aimed to describe the process by which machines think. A thinking that had little to do with what Alan Turing had proposed at the beginning of the decade (Turing, 1950), but rather with the formation of conceptual categories that allow the thinking entity to separate different objects into them (Pask, 1958). For Turing, thinking was roughly the equivalent of reasoning in terms of responses given to certain inputs (i.e. functioning as a black box). The machine which eventually passed the Turing test would do so because it would respond as if it processed equivalently to the process of a human intelligence, up to the point where it would mistakenly be taken by human. The enormous complexity of such device had to be implemented in its design from the beginning. Pask's approach was radically different. The thinking of his device had very little to do with emulating human thinking or behavior. Like Walter's tortoises or Ashby's Homeostat, the electrochemical ear was designed with very simple sensing and acting capabilities, and the complexity of behavior was expected to appear emergently through the combination and iteration of these simple assemblages.

The electrochemical devices were intended to create an analog control system. They consisted of electrochemical assemblages of various aqueous solutions of metallic salts, in which he introduced electrodes to apply current to them. As the current passed, a series of filaments of iron grew from the tip of the electrodes into the liquid. Pask called these groups of filaments 'threads'. As they grew and encountered others, the threads ended up creating connections among the electrodes. These connections would account for the assemblage's computation capabilities (Cariani, 1993; Pickering, 2010).

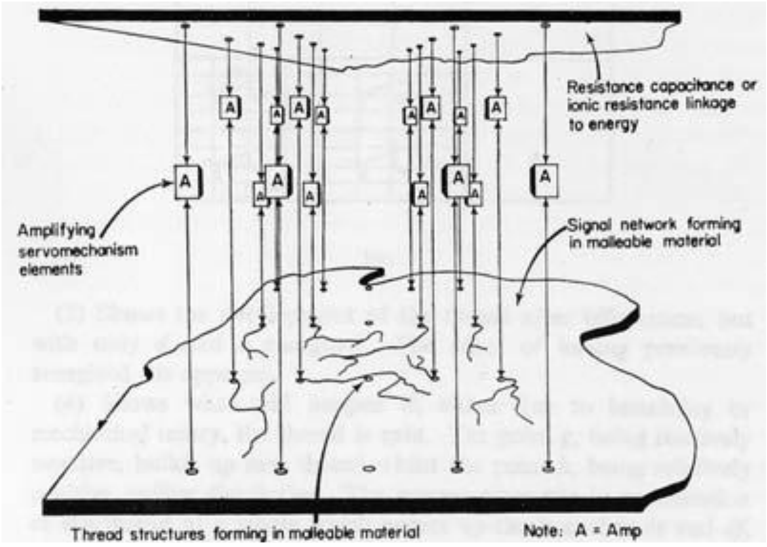


Figure 4.7: The formation of threads (Pickering, 2010).

Gordon Pask worked extensively on these assemblages, although in general the details of the exact compositions with which he worked remain obscure. By 1958, he had a working demonstrating device (Cariani, 1993). However, beyond the technical details, there is a very interesting aspect here in the sought open-endedness of it. The approach was to create a machine, a control system, which was capable of creating its own relevance criteria. That is, that the relations among inputs received and system's reactions would not be completely well-defined, but instead the device itself would be capable to choose to what it would react to and how. It would evolve its own sensors "to choose, independent of the designer, those aspects of its external environment to which it would react" (Cariani, 1993). And indeed the device was capable of doing so. As he presented it in 1958, it could either be trained to recognize magnetic fields or sound. In about half a day, it was capable of adaptively grow its own connections in order to do so. In the case of sound, once this was done it could also rapidly gain the ability to distinguish between two different frequencies (Pask, 1960). Hence the reference of Cariani and Pask himself to it as an 'ear'.

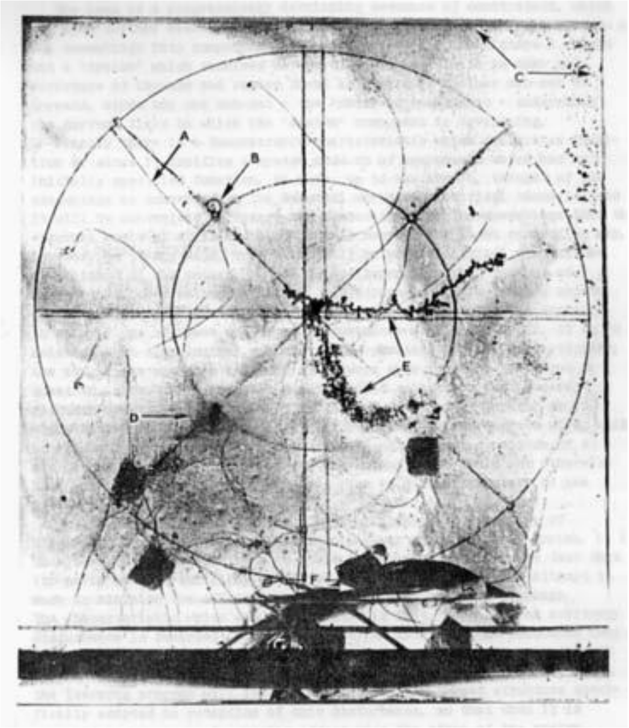


Figure 4.8: Threads growing on the electrochemical computer (Pickering, 2010).

The electrochemical ear is a very important example in Cariani's theory of emergence-relative-to-a-model that will be extensively reviewed in the following chapter. It is a proof that artificial systems which evolve their own sensors are a real possibility. The structural autonomy of Pask's device allows it to be informationally open, with a degree of epistemic autonomy that allows it to choose its own relevance criteria. With this, in creating its own sensors the device is able to incorporate new observables into the system of computation, and thus, it is an example of what Cariani calls creative emergence. It is one of the two ways in which emergence as generation of novelty can be achieved in artificial systems (the other one is Cariani's combinatoric emergence). It is certainly very difficult to attain, and it is precisely this difficulty what makes Pask's device such an important example.

4.2.3.3 The Colloquy of Mobiles

Finally, a last cybernetic device will be examined here. It is one that very nicely connects to the next section on Cybernetics and art. Indeed, it could have been included in that section as well, as could have been the Musicolour, which was in fact the first self-proclaimed cybernetic artwork and, indeed, the first interactive artwork (Penny, 2010). As a matter of fact, the distinction between artistic and non artistic is, like many other distinctions, blurred in Cybernetics. Many of the cybernetitians' projects, like those mentioned above were at least in the experimental verge between science and art. In fact, cybernetics was since its origins a highly (and truly) interdisciplinary field, which created "machines and theories [that] were neither art nor science; they utilized and exceeded both" (Fernandez, 2008). If Walter's tortoises were a predecessor of Artificial Life, Gordon Pask's Colloquy of Mobiles is undoubtedly a case of interactive Artificial Life Art avant-la-lettre. Not only because the theme of the piece, but also of its approach and realization.

The Colloquy was presented in London at the historically unique exhibition Cybernetic Serendipity in 1968, curated by Jasia Reichardt. As Pask described it, it was an 'aesthetically potent' environment. An environment that seeks the artistic enjoyment of the viewer or hearer, and that should be able to interest him or her into exploring it. The piece was aimed towards this goal. Among such environments Pask differentiated between those which are passive and those which are reactive, and proposed his installation as an attempt to go one step further in that direction (i.e. as an interactive piece)[42] (Pask in: Reichardt, 1968: 34).

The installation was made out of five pieces hanging from a platform that was suspended from the ceiling. Two of the mobiles had the ability to emit light beams, whilst the other three didn't. Instead, they had mirrors to reflect them. The first kind of mobiles were "as a whimsy", as Pask claimed, labeled as males and the second as females.

---

[42] With this, Pask is proposing his Colloquy to go further in that direction as an interactive piece, and thus, situates interactivity as the third kind of environment. This distinction among active, reactive and interactive is also at the heart of the concept of interactivity that will be elaborated in chapter 7 and which has been previously presented in (Soler-Adillon, 2012).

Figure 4.9: The Colloquy of Mobiles. © Gordon Pask Archive at the Dept of Contemporary History, University of Vienna, Austria. Used with permission.

Each of the mobiles was wired up to have a main goal, and would have to learn how to achieve it, either through competition of through collaboration with the other mobiles. Very much along the lines of Artificial Life Art, Pask built his system around the metaphor of sexual reproduction, or, rather, of mating rituals. Each mobile had two urges or drives that it would try to satisfy, labeled O and P, for the colors they represented them; orange and puce[43] (Fernandez, 2008). To reduce either of the drives, the male was required to project the corresponding light beam and to have it projected back to a specific part of its body. This was something for which it needed collaboration from a female, which was equipped with a mirror. The females, according to their

---

[43] As Fernandez notes in (2008), these two letters also "suggest forms of sexual stimulation."

current state, would also be required to fulfill either one or the other drive. So if there was a coincidence in the goals of male and female, the latter would offer collaboration. During all this, the males would compete with each other, blocking the competitors signal in order to gain the attention of the females for themselves. Sound and different intervals of light beam would send different messages in quite a complex sequence of events, in which each mobile learned how to best satisfy its particular urges.

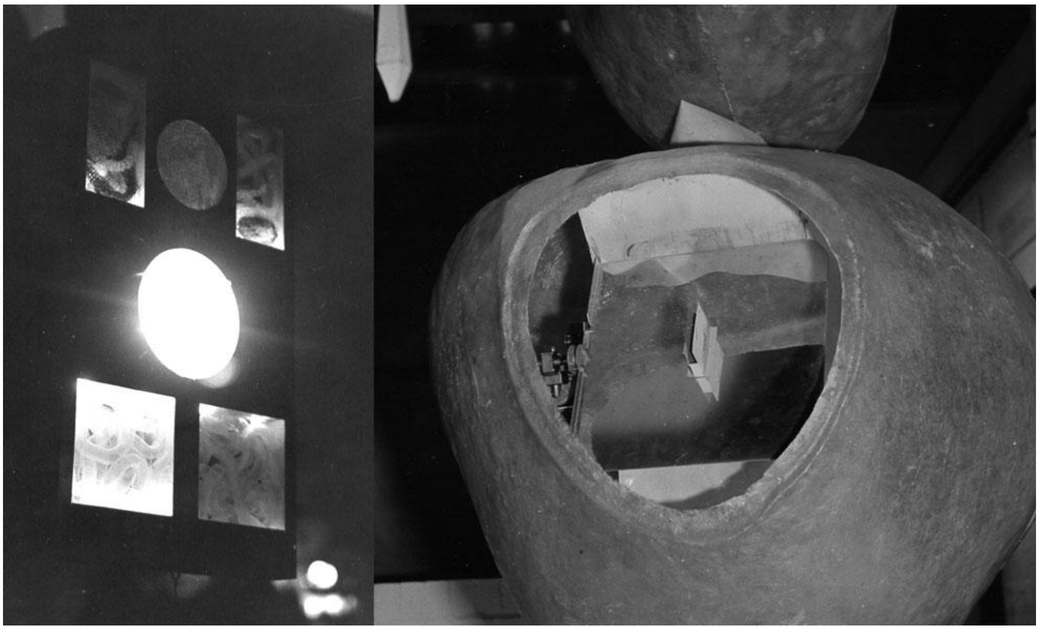


Figure 4.10: Detail of a male (left) and female (right) in the Colloquy of Mobiles. © Gordon Pask Archive at the Dept of Contemporary History, University of Vienna, Austria. Used with permission.

The Colloquy as a whole was a self-organizing system, as local decisions and interactions created the overall behavior, which could be labeled as emergent in this sense, again much like the case of the ants or the boids. And, on top of all that, there was the possibility for the visitor to the installation to intervene. Stepping into the installation space users could block light signals or redirect them with their own mirrors. As noted in (Pickering, 2008), the women’s make-up mirrors were a good resource for interacting with the piece, and some visitors did spend significant amounts of time on it.

Thus the piece had a double level of activity, which resonates with other interactive A-Life inspired pieces that will be reviewed in chapter 6 and 7, like Simon Penny’s Sympathetic Sentience or my own Digital Babylon. First, there is the level of the piece running on its own, creating its own equilibrium of self-organized behaviors. And second, there is the possibility for interaction, for the public to enter the space and affect the piece, creating with disturbances that enter the activity loop that forms the overall behavior of the system.

In conclusion, the Colloquy is remarkable for anticipating some of the latter developments in computer art, Artificial Life Art, as mentioned above, or even the notion of real-time interactive installations (Myron Krueger’s responsive environments) as noted in (Fernandez, 2008), as Pask integrated in it the active participation of the audience in his aesthetically potent environment.

4.3 Cybernetics and Art

Cybernetic art is both the predecessor and the first stage of computer art. It is a predecessor because it anticipates it thematically and in the way in which technology becomes central in the process of creation. And it is its first stage because many of the first projects that used computers with an artistic fashion were framed within, or at least related to, the cybernetic movement. The relationship between art and technology is as old as art itself, but technology as a central theme in artistic disciplines dates from the avant-gardes of the beginning of the twentieth century, mostly in the futurist movement.

The decade of the 1960s saw the appearance of Cybernetic Art during best years of Cybernetics in terms of popularity (if such a word can be used in academy), but also the beginning of its dissolution into a myriad of other disciplines. As the discipline reformulated itself as second-order cybernetics, cybernetic art remained marginal in respect to the mainstream art scene during the 1970s and 1980s. According to Gere (2008), during that time among the art practices with a strong technological component only video art was able to sustain interest in the art world. However, authors such as Dan Sandin or Woody Vasulka were directly influenced by cybernetics, and the system's aesthetics of Burnham remained influential for a long time. Also, other authors who would not identify themselves as cybernetitians were starting to become interested in systems and feedback. It is a prominent example the extensive work of Myron Krueger on responsive environments (Krueger, 1991). The rest of this section offers a partial overview of some of the most important practices of cybernetic art, and has no intention of being exhaustive.

4.3.1 Cybernetic Serendipity

Cybernetic Serendipity was the name of a now legendary exhibition which took place in 1968 at London's Institute of Contemporary Art. Jasia Reichardt worked on the preparation of the exhibition during the preceding three years. Although it is has a reputation of being the first computer art exhibition, there were several exhibitions that preceded it. According to (Klütsch, 2005), in February 1965, the Studiengalerie der Technischen Hohchschule Stuttgart held the first computer art exhibition, and others followed again in Stuttgart, Frankfurt and Darmstadt (Germany), New York and in the then Czeckhoslovakian cities of Brno, Jilhava and Cottwaldov.

It has also been noted that it was not so much an exhibition of computer art but on cybernetic art –it was indeed strictly about Cybernetics. Part of the focus was on how computers could be used rather than on how they were actually being used. The purpose of the exhibition was not to show what cybernetic systems –not exclusively computers– were capable of doing. It dealt "with possibilities, rather than achievements" (Reichardt, 1968: 5). In fact, according to MacGregor (2002), "there were only two digital machines in the exhibition and much of the work was produced using analogue technology."

Despite all of this, however, the importance of the event is undeniable. Among the participants were Nam June Paik, John Cage, Nicholas Negroponte or John Whitney. The event reached such a status that Umberto Eco traveled from Italy to visit it (MacGregor, 2002).

Cybernetic Serendipity was quite an eclectic assemblage of works, ranging from technical demonstrations of computing to analog robotic pieces like Gordon Pask's Colloquy of Mobiles to the Bell Labs still images created by computers. And it also had a strong theoretical component. When she planned the exhibition, Jasia Reichardt decided not to separate the artists and the engineers, so the visitors would not know, when looking at a piece, if it were made by a mathematician, an architect, an engineer or an artist. She didn't want this to influence how the visitors perceived the pieces (Reichardt, 1968; 1971). The works were organized in three categories, which are quite illustrative of the variety of the pieces shown:

1. Computer-generated graphics, computer-animated films, computer-composed and played music, and computer poems and texts.
2. Cybernetic devices as works of art, cybernetic environments, remote-control robots and painting machines.
3. Machines demonstrating the uses of computers and an environment dealing with the history of cybernetics.

Within the first category there were many of the milestone images of early computer graphics: John Whitney's early animations, Charles Csuri's experiments, which would derivate into the first real-time animation system, Knowlton's legendary computer generated nude image –a predecessor of ASCII Art– or Noll's computerized Mondrian-inspired paintings among them. Arguably, the only absent big name of computer generated animation from the exhibition was Ed Catmull, author of the first 3D imagery to appear in a movie and the brain behind most of the 3D graphics early development. Most likely his technical profile, apparently not interested by that time in artistic experimentation, made him less attractive to such an event.

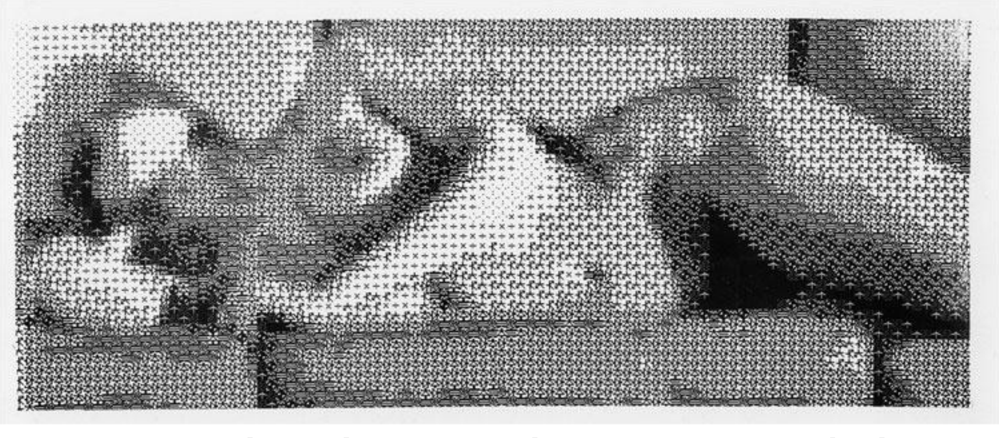


Figure 4.11: Knowlton and Harmon: Studies in Perception (Reichardt, 1968).

The second and third categories congregated a very big variety of works. Perhaps one of pieces that is most representative of what cybernetic art was is Nicolas Schöffer's CYSP 1 (for CYbernetic SPatiodynamics). A big sculptural tower which was "the first 'spatiodynamic sculpture' having a total autonomy of movement (travel in all directions at two speeds) as well as axial and eccentric rotation (setting in motion of its 16 pivoting polychromed plates)" (Reichardt, 1968). Made out of steel and duraluminum, it incorporated an 'electronic brain' designed by the electronics company Phillips, which funded the piece. It had motors to produce output and photo-electric cells and a microphone to read the variation of color and light and sound intensity. The sculpture reacted to its environment with either stillness or movement of its parts and around the space. Some colors would make it stay still, others move in certain ways or spin its blades. In general, darkness and silence would make it react, while brightness and noise would make it still. "Ambiguous stimuli, as in the case of Grey Walter's tortoises, produce the unpredictability of an organism" (Burnham, 1968: 341).

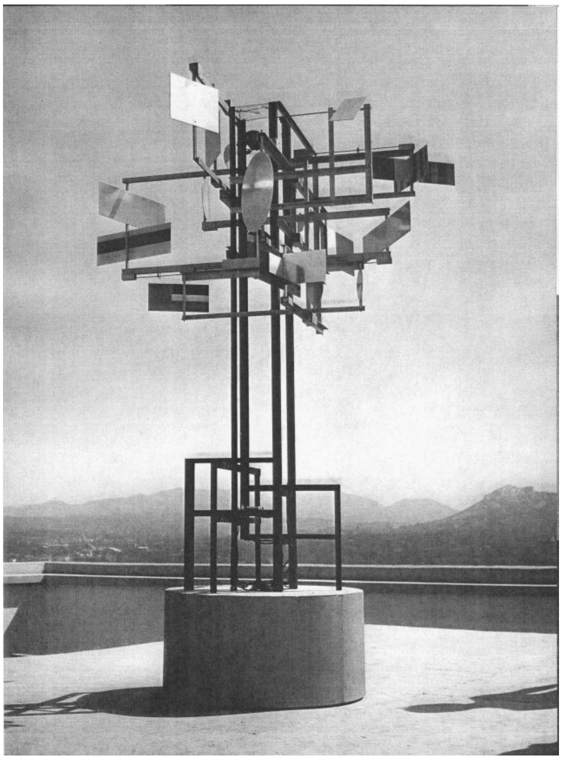


Figure 4.12: The CYSP 1 (Burnham, 1968).

With this piece, introduced to the public for the first time in 1956, Schöffer became one of the first artists to implement the idea of control devices in art (Burnham, 1968). However, this sculpture was not only about a control device, but about a complete cybernetic system. Indeed, much like the Musicolour, the Colloquy or the tortoises, CYSP 1 proposed a conversation with its environment, whether or not human activity intervened in it or not. It was an adaptive system very much in the way that Norbert Wiener had described Ashby's Homeostat: unpurposeful, as art always is in a certain sense.

4.3.2 Cybernetics and the Art of the 1960s

In the 1960s decade, the high point of Cybernetics coincided with a momentum in the art practices which facilitated its adoption as a vehicle in some artistic practices: "It might be argued that in the absence of a complementary aesthetic context, there would have been no common ground for the accommodation of cybernetics to artistic concerns" (Shanken, 2002). The Cybernetics zeitgeist, with its central notions of system, feedback and machine and machine-like control, were influencing all aspects of culture in this decade, and this coincided with the appearance of the art of happenings and performances of Allan Kaprow and John Cage, who were experimenting with the blurring of the distinction between artist (or artwork) and audience. In some of these art practices, participation was a key concept, and this fitted nicely with the cybernetic introduction of computerized systems which were interactive and, thus, participative. Also, disciplines such as kinetic art had opened the doors to the concepts of time, duration and process into the art object (sculpture in particular). In this respect, introducing the idea of behavior was almost a necessary step once authors like Schöffer experimented with the technology that cybernetics was fostering.

All of this prepared a context in which it was natural that many art practitioners gained interest with the ideas that Cybernetics was proposing. An important case of this interest is, according to Shanken (2002), the work of Roy Ascott and especially his writings during the first years after he discovered the work or Wiener and his colleagues. Ascott was, in the sixties, an enthusiast of what he called the Cybernetic Vision in respect to the art practice. For him, cybernetics is used both to explain the creation of interactive and participatory art, but also as a way to explain the art practice (and, for that matter, any cultural practice) as an activity framed within a bigger system. Cybernetics, as an integrative science, was for him a point of collision of many different knowledge, as much as art is a point of collision of experiences (Ascott, 2003). It allowed him to find a reference frame to describe art, the artistic artifact and even the teaching of art. The central idea of feedback, as a means of communication among agents in a system, was used by him as a pivoting point with which he linked all sorts of knowledge and practices. This framework also permitted him to describe the switch of focus that was happening in contemporary art in the sixties as a move from something that could be described as visual (like modern art had been, in the classic disciplines of painting and sculpture) to behavioral (Ascott, 2001). Along the lines of the general art movement, as mentioned above, he understood participation as both a sort of proto-interactivity and as participation of audience as it was being experimented with in the sixties in general[44].

Through this contextualization of the art practice, Ascott describes the artistic artifact as possibly self-organizing, in the sense that the interrelations among artist's creative activity and environment can influence each other through a series of feedback loops, if the activity is not framed in the traditional parameters but more open, in vogue with the sixties practices: "The participational, inclusive form of art has as its basic principle 'feedback,' and it is this loop which makes of the triad artist/artwork/observer an integral whole" (Ascott, 2001). Clearly, this idea of an integral whole as a result of self-organization resonates with the idea of emergence. However, analyzing what this means exactly in Ascott's work falls beyond the scope of this study.

4.3.3 Cybernetics as an Art Form and Cybernetic Art as Early Computer Art

In his seminal book *Beyond Modern Sculpture* (1968) Jack Burnham analyzes Walter's tortoises as an example of the scientific imitation of life, in consonance with the title that Walter gave to the Scientific American article with which he introduced the turtles to the general public. As would happen with Artificial Life, the scientific imitation of life that Burnham describes is not concerned with external appearances but with behavior and performance. Burnham considers that the cybernetic organism is a life form on its own. Walter's tortoises or Schöffer's CYSP 1 are post-kinetic sculptures, in the sense that they go beyond the mechanical repetitive movements of the Swiss clockwork –or of the installations of Swiss artist Jean Tinguely– and incorporate control systems based on feedback (that is, they are cybernetic).

Cybernetic art, therefore, consisted either (or both) in early robotics and Kinetic sculptures that incorporated control systems to react to their environment. But in the context of the sixties, also any artistic practice that utilized any sort of computerized systems would fall under the label of cybernetics. Hence the inclusion of so many computer drawings and animations in Cybernetic Serendipity, although as said above in this exhibition the computer did not have so much a central a role as one would have though at the light of the discourses around it. The opposite case was the exhibition *Software, Information Technology: Its New Meaning for Art*, an exhibition that took place in 1970 at the Jewish Museum of New York, curated by Jack Burnham.

---

[44] According to Simon Penny (personal communication, 2013-12-02), Burnham was in that time significantly more strongly influential than Ascott, who's influence has been claimed by Shanken and himself after the facts, three decades later. Burnham's show Software, the very influential books *Beyond Modern Sculpture* and *The Structure of Art*, and many artforum essays account for this argument.

As the name of the exhibition suggests, in this case the computer was the very center of the exhibition. In fact, it was an exhibition about interactivity. The visitors were invited to use computers, not only to see the results of their work. Nicholas Negroponte and MIT's Architecture Machine Group, Robert Barry, Vito Acconci and Allan Kaprow were among the participants in the event and Ted Nelson created *Labyrinth,* an interactive catalog to which he referred as the first publicly-accessible hypertext (Wardrip-Fruin and Montfort, 2003). With it, users could navigate an interactive database of information of interlinked texts, and then they would receive a print-out of their particular chosen path (Shanken, 1998).

As it turns out, however, the exhibition was shadowed with technical failure. The computer running most of the pieces and the interactive catalog, a DEC PDP-8 Timeshare Computer, collapsed the day before the opening and was inoperative for the first month of the show. This and other problems in the exhibition and with the Museum's board lead to the cancellation of a second venue in which the exhibition had to travel, and precipitated the dismissal of the museum's director Karl Katz. Nonetheless, Software remains, like Cybernetic Serendipity, an important milestone in computer art.

In the years that followed, Cybernetics lost its momentum as it partially dissolved into other disciplines, and some aspects of it were practically abandoned in research programs, only to be rediscovered over twenty years later. In this context, Cybernetic art per se became marginal, and computer art in general suffered from this collapse. As pioneering computer artist Frieder Nake declared, ideas were exhausted and artistic computer graphics were repeating themselves too much. He titled his article *There Should be No Computer Art* (Nake, 1971).

4.4 Conclusion: Cybernetic Art as a Predecessor of Artificial Life Art

When Cybernetics started to dissolve into other disciplines, some aspects found immediate continuity, whilst others were forgotten, only to be picked up years later. The ongoing legacy of Cybernetics, usually covert, is reflected in many aspects of computer science and human-computer interaction related discourses and also in the development of systems ecology, systems sociology (Luhmann), systems thinking and even in meteorology (Penny, 2009; 2010). Among the rediscovered aspects of Cybernetics, there are many of the concepts that, towards the end of the 1980s, would form the theoretical basis of Artificial Life: “Many of the ideas central to cybernetics reappear under slightly different terminology in artificial life discourse. Central to cybernetic thinking were questions of self organization and purposive behavior, the relationship of an entity to its (changing) environment, its real time response and adaptability – interactions characterized as ‘feedback’. In artificial life, these ideas are clad in terms of autonomous agents, reactive insect like robots, simulated evolution in fitness landscapes, emergence and self-organizing criticality” (Penny, 2009).

But beyond this, many of the practices described above anticipate not only Artificial Life but Artificial Life Art. First, it is clear that some of the works of the cybernetitians anticipate it thematically, Walter’s tortoises being the most evident example. All the vocabulary and metaphors that Walter utilizes to describe the tortoises and its behaviors resonates with the descriptions that the ALife Art practitioners would adopt. Evidently enough, the devices themselves are called tortoises and even have names, and their parts are referred to as eyes or shell. Accordingly, their behaviors too are described as the tortoises feeling an obstacle, being independent or spontaneous, or learning (Walter, 1950; 1951). Gordon Pask’s Colloquy is also a clear example of how both the devices and their behaviors are described in Artificial Life Art terms: the males and females and their actions performed in order to fulfill their drives, seen as a clear metaphor as a mating dance between the two genders of a species. As it has been noted in (Fernandez, 2008), this anticipates too an elaboration of narratives to explain the behavior of the agents that were strongly gendered.

Another aspect in which Ashby, Walter and Pask anticipate Artificial Life is in the bottom up approach in which they base their experiments. In all these cases, and more clearly than any in the tortoises, the idea is to build simple systems capable of exhibiting complex behaviors. Walter was quite explicit in this objective in (1950; 1951). The tortoises were built to prove how only a very limited set of connections could result in something that resembled something as complex as animal behavior.

But besides all of this, all of the devices described in the preceding section share with the practices of the artists of Artificial Life Art a an approach to modeling that is often quite free and experimental, in comparison to common scientific modeling practices.

For the scientific cybernetitians, this was always done in order to try to empirically demonstrate a certain aspect of their theories: how very simple aggregates could lead to complex behavior, or how systems sought stability in environments with conditions that were ever-changing, etc. For the Artificial Life artists (as well as for the cybernetic artists such as Schöffer) the goal is more loosely linked to a previous theoretical frame. Artificial Life Art artists borrow techniques and concepts from scientific Artificial Life and play with them. More often than not, however, they are trying to achieve the same results as the early cybernetitians were: to implement systems in which emergent phenomena occur, either as self-organization emergence or in the search for the creation of novelty. The goal for these artists is to create unpredictable black boxes in the sense described above: devices and systems that will hopefully surprise the designer, exceeding his or her expectations in terms of behavior, as they attain a degree of complexity that was not pre-specified in the design.

5. EMERGENCE-RELATIVE-TO-A-MODEL

Peter Cariani is a researcher at Harvard Medical School and Boston University. His research interests include theoretical biology, cybernetics, Artificial Life, auditory neuroscience and the psychology of music. He has elaborated a theory of emergence known as emergence-relative-to-a-model (ERTM).

ERTM is an approach to the problem of emergence that is epistemological and pragmatic. It is epistemological in the sense that Cariani is interested in how emergence can be perceived by an observer, accounted for, and scientifically communicated. The ontological considerations of emergence are left out of the debate in ERTM. And it is pragmatic because the goal is to build devices which can exhibit emergent behavior, understanding emergence as generation of novelty: The problem of emergence is, according to Cariani, the problem of specification vs. creativity (Cariani, 1992b); the process by which new structures and functions come into being (Cariani, 1997). It is about the design of open-ended systems and devices, which is an important issue in order to create systems “that can autonomously find solution to combinatorically-complex and ill-defined problems” (Cariani, 2008).

Cariani is concerned with how new functions can appear in systems or devices that perceive and act on their environment. The basic idea is that this newness can only be accounted for scientifically if, first, the observer of the system defines the states and state-transitions of the system under observation; and second, if these observations are used to make predictions on the future states of the system. Once this has been done, emergence occurs whenever unanticipated behaviors, states or functions appear: “emergence is the appearance of novel entities that in one sense or another could not have been predicted from what came before” (Cariani, 2009a).

This approach to emergence differs from the idea of emergence as self-organization illustrated by the boids or the ants’ colony. The boids are a paradigmatic example of self-organization emergence, but for emergence as generation of novelty, they are not interesting except for the first amusing effect that they produce to a novel observer. Once the phenomenon is seen and understood, there’s nothing new in repeating the simulation. ERTM does connect only to those approaches interested in the potential for emergent behavior to create new structures and behaviors, such as the case of Holland’s genetic algorithms, but in general it is critical to the computational emergence that was at the heart of the early Artificial Life movement. Cariani was in fact involved in the first stages of the movement, and participated the first few workshops organized by the Santa Fe Institute. In these meetings, he raised a critical voice regarding many of the assumptions underlying the idea of emergence in vogue within the group, addressing issues that he considered had not been receiving enough attention (Cariani, 1992b).

Cariani is interested in how his idea of emergence can be applied to robotic devices or, more generally, 'cybernetic percept-action systems' (Cariani, 2011). In order to articulate his discourse, he elaborates a taxonomy of devices, which range from those less inclined to generate novelty to those were novelty finds its way easily. A brief account of the taxonomy follows these lines. For details, see Cariani (1989a; 1991).

The first distinction in the taxonomy is between formal-computational devices and formal-robotic devices. The first are devices that are completely disconnected from their environment in terms of the algorithms they execute. They don't have sensors to read data from their environment, nor have actuators to act on it. The devices of the second type are connected to the environment in terms of sensors ('measurement') and effectors ('control', 'action'), and thus their calculations are affected by these connections. I.e. they have sensors and effectors, and their computations are used to decide how to act on the world according to the sensor's readings of the environment.

Among the last type, there are devices that are non-adaptive, adaptive and evolutionary. The first is a device that cannot modify either its internal structure or computations, or its means of perceiving and acting on the environment. The second, adaptive devices, are systems that can change the relations between what they perceive and how they act, in terms of previous experiences. These can be learning devices. The last is a type of device that can alter its own sensors and actuators, in addition to its internal computations: these are the evolutionary devices in Cariani's taxonomy. These are the most open-ended and prone to emergence of the systems he analyzes. The rest of the chapter elaborates on how these types of devices relate to the idea of emergence.

As said above, Cariani's goal is grounded on pragmatism: his focus is "primarily on developing heuristics for generating useful novelty than in engaging in philosophical debates over the status of emergent novelty vis-à-vis various postulated ontological frameworks" (Cariani, 2012). The goal, then, is not to refute or reinforce these ontological debates. Rather, he focuses on the epistemological aspect of emergence; how it can be observed, and how such observations can be communicated scientifically. "Ultimately we want a theory which will guide us in our construction of fundamentally new ways of observing and acting on the world" (Cariani, 1989a).

5.1 Theories of Emergence

In his survey of different conceptions of emergence (Cariani, 1989) Cariani defines the problem of emergence in relation to the problem of the appearance of fundamental novelty in the world, the origins of order, and the increase in hierarchical complexity. It has been classically involved, he affirms, with the origins of qualitatively new structures and functions not reducible to those already in existence (Cariani, 1989a; 1992b). Along with this, he defines three separate discourses on the problem of emergence, "each with its own interpretation of the problem and distinct view of the world. These three major conceptions of emergence might be called 'computational emergence, thermodynamic emergence, and emergence relative to a model'" (Cariani, 1992b). These three conceptions respectively involve mathematical, physical, and functional accounts of emergence.

5.1.1 Computational and Thermodynamic Emergence

The basic idea of computational emergence is that order arises at one level as a consequence of the ordered activity at an inferior one. It is the classic idea of the whole being more than the sum of its parts. It corresponds with Langton's or Holland's account of emergent phenomena as the result of a large number of interactions in a lower level. Along with the Artificial Life community, according to Cariani, this view is embraced by many connectionists. Arguably, while computational emergence is not well-defined in the literature, it is intuitively related computational complexity and surprise (Peter Cariani, personal communication, 2014-06-23).

Computational emergence is built upon a formal, mathematically based conception of the problem of emergence. For Cariani, this view is reductionist and deterministic. It assumes that the behaviors at higher levels are reducible to a series of rule-like (deterministic) micro behaviors: "Because its ontology only admits of micro-deterministic computational interactions (of which chaotic processes can generate apparently nondeterministic macro-interactions), there is a strong platonic component to the world view: the ideal forms of computational behaviors can be abstracted completely from their material substrates, and the material world can be left completely for a virtual one. In assuming rule-governed, bottom-up organization rather than semi-autonomous levels of organization, computational emergence tacitly incorporates the older reductionist assumption that micro-orders determine macro-orders but not vice-versa" (Cariani, 1992b).

The Platonism that Cariani denunciates can be also applied to Langton's claims that Artificial Life is a legitimate way to study life itself (carbon-based life, as he referred to it), and an extension of theoretical biology: "The dynamic processes that constitute life –in whatever material bases they might occur– must share certain universal features – features that will allow us to recognize life by its dynamic 'form' alone, without reference to its ‘matter’” (Langton, 1988). This Platonism has also been pointed to by Howard Pattee. The idea that a deterministic machine, designed to respond to arbitrary rules, could realize life or evolution is, according to Pattee, sustained in the platonic believe that form is more fundamental than substance (Pattee, 1989).

Within this view, Cariani argues that the computationalists’ strategy is often to design simulations with enough elements and interactions concerned such that it is not easy for an external observer to foresee its outcome. Once the simulation is run, the observers notices high level patterns, which are “literally said to have emerged over the course of the simulation, and ‘the emergent behavior is consequently attributed to the device itself’” (Cariani, 1992b). Therefore, the quality of emergence would here be linked to non anticipation. And as Cariani points out, this dependence on the simulation would be one of the main problems of computational emergence. If this approach is valid, he says, then virtually any computation is emergent, and this would reduce emergence to a synonym of change. This is essentially the same argument that was made in chapter 3 about Mark Bedau’s argument about the importance of the simulation in order to identify emergent phenomena. Indeed, Bedau is thinking in the terms that Cariani frames as computational emergence. Just like there is no clear means of objectively discerning at which point the outcome of a Game of Life simulation ceases to be foreseeable[45], the same happens with mathematical computations.

Cariani is interested in emergence as newness, and in his terms this means essentially the introduction of new primary building blocks in a system’s computation, which can be the source of new behaviors. But in computational systems, by (his) definition, new primary structures cannot appear. Computers are deterministic and need complete specification. Computer simulations “will not create properties which were not encoded in the simulation from the start” (Cariani, 1992b).

The issue of how novelty is generated is at the heart of Cariani’s critique to computational emergence. His understanding of emergence is that it is a relevant concept if it can account for novelty, and his approach implies the necessity of clearly defining what is understood as emergence, so the appearance of emergent phenomena can be objectively observed and communicated. In contrast, he found that the Artificial Life community was often rather ambiguous in defining what they exactly meant by emergence: they simply asserted that their simulations generated emergent phenomena, without specifying a clear frame of reference to account for it (Cariani, 1992b; personal communication, 2013-07-03).

There is a conceptual difference in the emergence of the computationalists and that of Cariani, as Cariani himself pointed out in stipulating three separate approaches to emergence (Cariani 1989, 1992b). While Cariani’s theory is the paradigm of emergence

---

[45] See section 3.7.2

as generation of novelty, Computational emergence is essentially linked to self-organization emergence. And as argued earlier, self-organization emergence is not necessarily related to novelty. It is precisely about local interactions generating patterns of behavior that are observed at a higher level, as happens with the ant colony, the boids or cellular automata. Therefore, in relation to self-organization emergence, computational emergence is no longer problematic, as this identification already implies that novelty is not a necessary condition for it.

Thermodynamic emergence is, for Cariani, less controversial. It is mainly concerned with the generation of emergent structures, while he is interested in the generation of emergent functions. In general terms, this relates too to the difference between self-organization emergence and emergence as generation of novelty. Thermodynamics, as it was discussed in the chapter 4 with the example of Prigogine, understands emergence as the process by which a series of agents (molecules, particles, etc.) generate spontaneous ordered structures that can be identified by an observer, such as the case of chemical clocks: “The primary phenomena to be explained by thermodynamic emergence is pre-biotic evolution, where complex molecular reaction cycles and networks arise out of the chemical flux of form complex, self-propagating structures. Out of the analysis of chemical systems come concepts which are used to explain biological and social structures” (Cariani, 1989).

In conclusion, both thermodynamic and computational emergence can, in general terms, be understood in relation to self-organization emergence. The phenomena they describe are not necessarily related to novelty. The necessary condition for both is an ensemble of interacting agents generating observable patterns at a superior level. Whether or not these patterns are new and in what sense they are is not indispensable to label such phenomena as emergent.

5.1.2 Emergence Relative to a Model

The theory of emergence deployed by Peter Cariani is known as emergence-relative-to-a-model. Cariani presents it in (Cariani, 1989; 1992b) as a third alternative to the analysis of emergent phenomena, with a basis in systems theory (Ashby, 1957; 1981; Klir, 1968), with parallels in the writings of Robert Rosen. However, it is in Cariani’s own work where the most elaborated effort to describe such an approach is found. This section is dedicated to analyze Cariani’s two main influences. The rest of the chapter is dedicated to his contribution.

The theoretical foundation of Cariani’s emergence-relative-to-a-model is Robert Rosen’s idea that emergence is a deviation of a system’s behavior in regards to the behavior predicted by the model of that system. Rosen was a theoretical biologist who made significant advances in relational biology, (he was the last student of the mathematical and theoretical biologist and neuroscientist Nicolas Rashevsky). He developed what he called a relational approach to biology. An approach to biology which is to be read as largely antithetical to the more common analytic experimental approach –the reductionist approach– of biochemistry and molecular biology. Instead of focusing in the physical and chemical constituents of biological systems, Rosen's approach focused on its functional and behavioral aspects: "The relational approach … treats as primary that which is discarded first by physico-chemical analysis; i.e. the organization and function of the original system. In relational biology, it is the structural, physical detail of specific systems which is discarded, to be recaptured later in terms of realizations of the relational properties held in common by large classes of organisms, if not universally throughout the biosphere" (Rosen 1985: 5).

Once the focus was moved from the constituent parts to the functions and structures, Rosen –who was developing his research in a very interdisciplinary context (Rosen, 2012)–, proposed to investigate how these very same method could be applied to social structures and societies. The idea was not to extrapolate biological principles into the human realm, but to see if common modes of organization could be demonstrated between social and biological structures.

To elaborate on this idea, Rosen lays out the basic principles of self-organization. He acknowledges how, in biology, the investigator is almost always constrained to the position of the external observer, "attempting to characterize the infinitely rich properties of life entirely from watching their effects without any direct perception of underlying causal structures. For instance, we may watch a cell in a developing organism differentiate, migrate, and ultimately die … [but we] cannot understand the mechanisms by which the individual behaviors of billions of such cells are integrated into the coherent, adaptive behavior of the single organisms these cells comprise" (Rosen, 1985: 5). On the other hand, as member of social structures in organizations, we are direct participants and, thus, not external observers, of such structures. In this case, "as participants, we know the forces responsible for such improbable aggregation as football games, parades on the Fourth of July, and rush hours in large cities. But how would an external observer account for them?" The relational approach aims to understand both points of view (that of the observer of the whole and that of the parts) with one common framework. By establishing the parallelisms between different realms, we can move from one to another as we account for the similarities in their functions and behaviors.

The main conceptual tool in Rosen's approach is the theory of modeling, i.e., of the relation between a system and the model of that system. A model is a necessarily abstracted and simplified version of the system which is being modeled; a logical and mathematical construction which is created in order to account for its functions and behaviors. As such, it has fewer degrees of freedom than the original system and it is de-contextualized at least to some extent. While systems are open to a large number of interactions with their environment, models are necessarily closed to some of these interactions (Rosen, 1985: 17). These characteristics are what allow the model to be created and to compare models of different systems. But it has too some epistemological implications, one of which being a certain degree of impredictability (Rosen, 2012: xxvi). That is, the unavoidable abstraction of the model from the system being modeled necessarily allows for the possibility of some parameter or degree of freedom, which is not accounted for in the model, to eventually generate a behavior or function that had not been predicted.

The discrepancy between the behavior actually exhibited by a system and the corresponding behavior predicted on the basis of the system's model "subsumes such concepts as 'error', 'system failure' and 'emergence'. In all of these we find an unpredictable and apparently causeless departure from expected behavior, arising precisely because the interactions responsible for such behavior in the actual system have been abstracted out of our model, and hence are invisible in it" (Rosen, 2000).

Therefore, for Rosen emergence is literally a failure of the model to predict the behavior of the modeled system. It is in this sense synonymous to error. However, this error does not necessarily mean that the model was not properly constructed. The possibility of emergence, understood as such, is inevitable with this modeling methodology, because the model is not a complete description of the material system. John Von Neumann's famous phrase comes to mind here: "with four parameters I can fit an elephant, and with five I can make him wiggle his trunk" (Von Neumann, cited in: Mayer et al., 2010). Error, or emergence, is the difference between what the system does and what it would do if it were not open (Rosen, 1985: 311).

According to Rosen, the modeling relation between a natural system and a formal one is generated though an encoding of observables (system state, behaviors, functions, etc.). And this, which is done through observation and evaluation, is necessarily an abstraction. Once the model has been created, the behavior of the modeled system can be predicted. But as time advances and the system interacts both with what was encoded and with what was not, eventually new states of the system or behaviors may appear, making it necessary to incorporate new observables into the model. The introduction of such new observables "is a manifestation of the phenomenon of 'emergence', or 'emergent novelty'" (Rosen, 1985: 305).

Cariani strongly relates to such methodology, and also to its underlying epistemological approach. Rosen is not claiming here to account for intrinsic properties of the system, like the advocates of computational emergence do, according to Cariani. His emergence is purely epistemological, as it is linked to the model of the observer: "What is 'emerging' here is not an intrinsic property of [the system being modeled], but rather the failure of an encoding to model [the sytem] over time" (Rosen, 1985: 305).

Rosen summarizes his conclusions on the theory of error –of emergence– in the following points (Rosen, 1985: 311):

1) The fact that any description of a system is an abstraction is the basis for error. The observer has a limited set of observing instruments or meters, with which he or she elaborates the equivalence classes, and this necessarily leads to the classification of some dissimilar states as identical.

2) This becomes an error once the system's states interact and split the observer's equivalence classes.

3) This appears to the observer as a miss-classification, and thus is interpreted as an error in the system under observation.

4) Error in the system is a discrepancy between observed and expected behavior, each determined on the basis of the descriptions generated by the observing instruments employed.

5) These errors are not a consequence of an incorrect modeling of the system, as these bifurcations "can be regarded as a logical independence of the descriptions which are being compared." It is because of the very nature of modeling that these discrepancies are unavoidable in some cases: "the nature and character of the errors detected by the observer are unpredictable in principle from his descriptions."

6) To deal with errors, the observer must supplement his description in order to take into account the new observables. One way to do this, says Rosen, is to treat it as a system's intrinsic stochastic process, which according to the last conclusion is not entirely true.

7) To fully account for the interactions that produced the error, both the observed system and that with which it interacted should be described. Error is not intrinsic to the system in isolation, but only in relation to its environment. Thus, error "is better regarded as a meter for environmental observables."

The discrepancy between the abstracted system and the system itself, which is open to environmental interactions, will necessarily appear. Furthermore, Rosen notes, in a dynamical context such discrepancy will grow over time. But this discrepancy by itself is not error. Error arises when the underlying variability is sufficient to cause a bifurcation within an equivalence class. That is, two states which were classified as equivalent have enough environmental interaction as to appear as different at some point in time. "Hence, in a sense, error and bifurcation are equivalent terms; they both indicate the inadequacy of a particular mode of system description, and imply the existence of others. When such a bifurcation occurs, the original modeling relation which existed between the system and its description now ceases to hold. As we have seen, such a failure of a modeling relation can also be interpreted in terms of 'emergence;' here too, a given description or model of a system must be replaced by another" (Rosen, 1985: 319).

The relation that Rosen establishes between error and emergence is a departure from the most common conceptualizations of the concept, as seen in the preceding chapters, and can also mislead into a consideration that downplays its importance as an idea. But as will be discussed in the following sections, Peter Cariani elaborated a framework that allowed for this emergence to maintain its most vindicated quality, newness, while moving it away from the ambiguities and polemics that are often found in the discourses around it. Cariani's approach is, like Rosen, epistemological and pragmatic. It wants to identify the appearance of the new for the informed observer, not identify properties ontologically attributable to the systems under analysis.

The 'relative to a model' of Cariani's emergence-relative-to-a-model is to be read as an elaboration of Rosen's idea of error. For Cariani, too, emergence is the discrepancy between actual and expected behavior. Before assessing that something is emergent, according to Cariani, one must elaborate a model of the system or device under observation, with its states and primitives clearly defined. This observation includes the simulation –or observation over time– of the model or device. Once this has been done, if something new appears, it will be emergent relative to this constructed model. This something new can later be incorporated into the model, and it ceases to be emergent from then on. In essence, this is the idea of emergence as generation of novelty. The details of Cariani's approach are detailed in the following sections.

5.2 The Modeling Theory of ERTM

Cariani bases his idea of modeling in what he labels the Hertzian commutation diagram (Cariani, 1992, 1998, 2003; 2011), based on (Hertz, 1899): an explication of the operational structure of the predictive scientific model. The idea is that, firstly, an observer makes measurements that produce symbols that will become the initial conditions of this formal predictive model. Then, using the model's algorithm, he or she makes a prediction of future state of the model. Finally, a second measurement on the system is made and is compared to the prediction to see if it was successful.

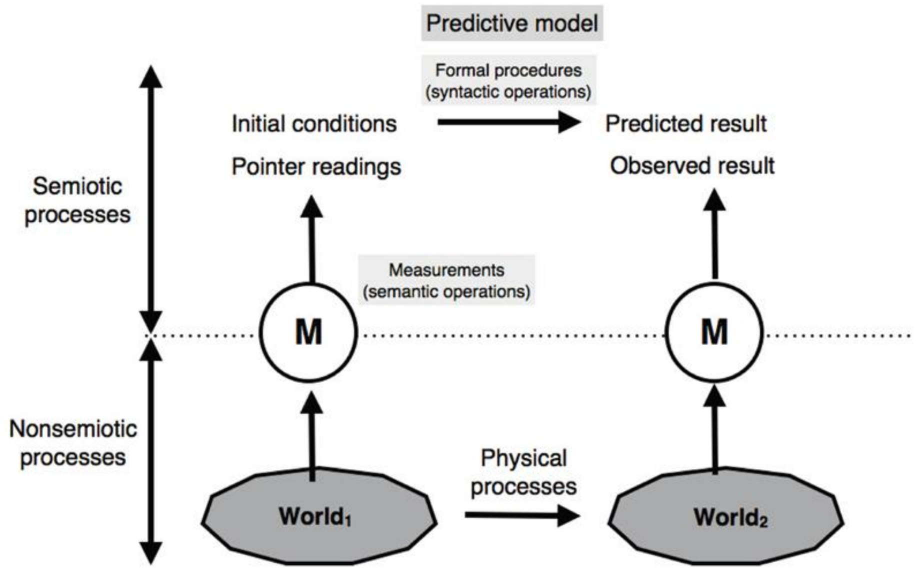


Figure 5.1: The Hertzian commutation diagram (Cariani, 2011).

An important aspect of the hertzian modeling scheme is that it operates in the semiotic realm. As noted by Cassirer, Hertz was the first modern scientist to move from a theory of knowledge based on the idea that science operates with copies and reproductions of material data, to "a purely symbolic theory" (Cassirer, in: Cariani, 2003). The formal mathematical system is made of symbols, which link to the external world through the pointer readings of the measuring devices. Then, the symbols are manipulated in order to determine a prediction, which will be confirmed with further readings.

Cariani bases here his adaptation of the hertzian idea in the semiotic triad proposed by Charles Morris: semantics (signification of signs), syntactics (combinations of signs) and pragmatics (origins, uses and effects of the signs) (Morris, 1946). This triad has its corresponding operations: measurement, computation and evaluation (Cariani, 2011). Both the three semiotic aspects and the corresponding operations are "irreducible and complementary" (Cariani, 2003; 2011), i.e. one cannot be replaced by the other: "They depend on different modes of causation. Syntactic operations are driven by logically-necessary rules on symbol types (formal causes), while semantic, measurement operations involve contingent, material interactions (material causes) and pragmatic, adjustment operations depend on the goals of the observer (final causes). If one clearly defines these semiotic relations in these terms, then one cannot replace semantics with syntactics, semantics with pragmatics, syntactics with semantics. Form, meaning, and purpose are inherently independent of each other" (Cariani, 2003).

In terms of modeling the behavior of cybernetic devices, which is the core of Cariani's interest, this is relevant because, as he notes, there are strong parallels between the structure of scientific models, the functional organization of organisms as informational systems, and the operational structure of robotic devices (Cariani, 1998a). Thus, the scientific semiotic model can be expanded to organisms and artificial devices: "The semiotics of the modeling relation in science can be extended to incorporate organisms and devices that sense the world and act on it contingent upon what they sense. While scientific models are externalizations of the structure of the individual observer, modeling relations are also embedded in the internal structure of individual organisms and devices" (Cariani, 1998b).

This is the model to which Cariani refers to in the theory of emergence-relative-to-a-model. Each individual system is to be observed, and a model to predict its behavior elaborated, within this semiotic parameters. Semantics, as the contact points of the system with its environment: the sensors and effectors of the device. Syntatics, as the internal computations that the device performs in order to decide which action to take. And pragmatics as the evaluation of the performed action in terms of achieving a specific objective in regards to the relationship of the system with its environment.

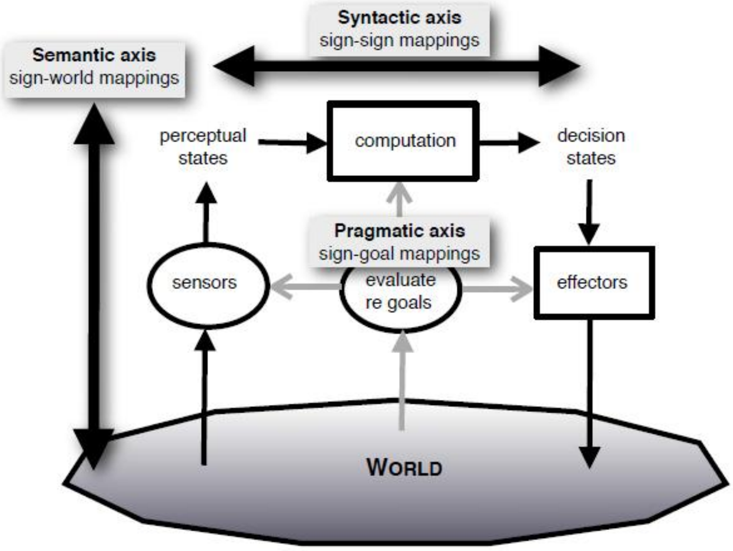


Figure 5.2: The functional organization of cybernetic devices. Relation of semiotic dimensions to functionalities and operations (Cariani, 2012).

One more fundamental aspect is needed in order to complete the model of the system or device: the ‘primitives’ with which the device operates must be annotated and described. The primitives are the most basic building blocks in the system’s computations: “an indivisible, unitary entity, atom, or element in a system that has no internal parts or structure of its own in terms of its functional role in a particular system” (Cariani, 2012).

The modeling of ERTM consists in describing system states in these terms, through the observables that the external observer can account for. Upon such observations, predictions on the future behavior of the system or device are made. When such predictions fail to account for future behaviors, two possible types of emergence occur: syntactic adaptation or combinatoric emergence when the primitives have been combined in a novel way, and semantic adaptation or creative emergence when new primitives have entered into the system’s computations. These two types of emergence are reviewed below.

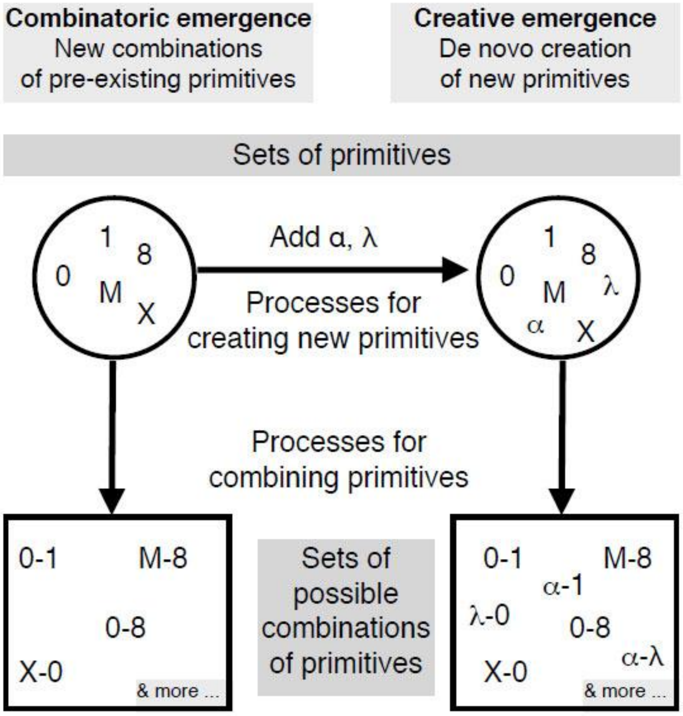


Figure 5.3: The primitives, exemplified with letters and numbers, and their possibilities of combination (Cariani, 2012).

The concepts of ‘system’ and of ‘state’ are borrowed from Ross Ashby (Cariani, 1992). In coherence with ERTM, the approach proposed by Pask is epistemological: a state is an observable distinction, a well-defined condition or property of the system that will be recognized if it occurs again. System, in this context, is a series of distinguishable states chosen by the observer. The behaviors of the system consist of the patterns of the transitions among system states over some period of time.

Cariani summarizes his ERTM approach in this (1992) passage: “Once we have fixed our observational frame we can talk precisely about emergence: whether the behavior of the physical system in question has changed with respect to the frame and in what ways it has change. If we observer the system for a period of time, we can build up a model including all of the syntactic and semantic state transitions we have previously observed and we can see if new state-transition patterns appear over time as the system subsequently transit through its states. The model thus constitutes the observer’s expectation of how the system will behave in the future. … If no new state-transtions arise, then the device or organism is ‘nonemergent’ with respect to the observer’s frame. … If the pattern of syntactic state transitions changes such that new computational transitions are formed, then the device or organism will appear to be ‘syntactically emergent’. … If the pattern of semantic state-transitions changes such that new measurement or control transitions are formed, then the device or organism will appear to be ‘semantically emergent’.” Both kinds of emergence, which are not mutually exclusive, would force the observer to change the model in order to continue tracking the device or organism.

5.3 Epistemic Autonomy

When a system or device becomes capable of choosing its own syntactic and semantic couplings, it attains epistemic autonomy. That is, it becomes able to change how it interprets the symbols with which it operates (syntactics) and how it perceives and acts upon the world. Epistemic autonomy is a central concept in Cariani's theory of emergence, as it is a fundamental prerequisite for the open-endedness necessary for emergence to occur: "emergent devices to be useful in amplifying our own creativity must have both a degree of structural autonomy relative to us as well as richness of potential structure" (Cariani, 1992b).

Both epistemic autonomy and the modeling scheme described above are related to Pattee's concept of semantic closure (Cariani, 1989a; 1991). It is also connected to concepts such as autopoiesis, self-modifying systems, self-reproducing automata or anticipatory systems. Semantic closure is a property of both natural and artificial systems (Pattee, 1982). In artificial systems, it is the condition by which a system is capable of autonomously operating according to the semiotic model described above. That is, of defining its semantic relations with the environment, the syntactic operations it performs with the symbols built from these, and the pragmatics to evaluate actions and purposes. Semantic closure is, for Pattee, a necessary condition for evolution (Pattee, 1982; 1985).

For Cariani, epistemic autonomy is a necessary condition for emergence. Only devices with a certain degree of structural autonomy, that is, with capabilities of deciding their internal and external mapping, can achieve it. This is a characteristic that can be traced back to the early cybernetic experiments: "As Ashby noted, in order to achieve better performance over its initial specification, a device must be informationally open, capable of interacting with the world independently of its designer" (Cariani, 1993). This is a key idea in emergence-relative-to-a-model: overcoming the limits of designer specification is at the heart of the objective of elaborating creative devices, capable of exhibiting novel behaviors as they adapt to their environment. It is what permits epistemic emergence (Cariani, 2009a): the creation of new points of view, and it relates to the possibilities of amplifying the conceptual landscape in minds, as reviewed in section 6.6 below.

5.4 Syntactic Adaptation or Combinatoric Emergence

Syntactic adaptation and semantic adaptation are arguably the two core concepts in ERTM. They are the two main manners in which emergence occurs, as the device under study creates new behaviors that either change how it computes the readings from the environment (syntactic adaptation) or its relationship with the environment, through changes in its sensors and effectors (semantic adaptation). In his early writings, Cariani refers to these two core concepts with the terminology that links back to the Morris' semiotic triad: syntactic and semantic adaptation[46] (Cariani, 1989a; 1990a; 1990b; 1991; 1992a, 1998). Later, the same concepts are referred to with the terms combinatoric emergence –or combinatoric novelty– and creative emergence –or creative novelty– respectively, although their meaning remains essentially the same (Cariani, 1997; 2008; 2009a; 2009b; 2011; 2012). Since they are the most recent formulations, this dissertation will use from this point on the notions of combinatoric and creative emergence.

Combinatoric emergence consists in the appearance of new structures (system states, functions or behaviors) in a system or device, generated with recombinations of the existing set of primitives with which the system operates. These structures are new in respect to the model that was build to account for the system states and behaviors of this particular system.

As a creative strategy, combinatoric novelty is dynamic, as it constantly brings new combinations of elements into being (Cariani, 2012). This happens when a device alters its computational parts contingent upon its performance in an external environment (Cariani, 1992). The recombinations constitute an alteration of the mappings between what the system perceives or reads from the environment and how it responds to such readings: "adaptive alteration of percept-action mappings is a form of combinatorial syntactic emergence –new percept-action mappings are created from existing percept and action primitives" (Cariani, 2011).

---

[46] Pattee used the terms syntactic emergence and semantic emergence in a similar way in discussing the types of emergence relevant to Artificial Life in (Pattee, 1989).

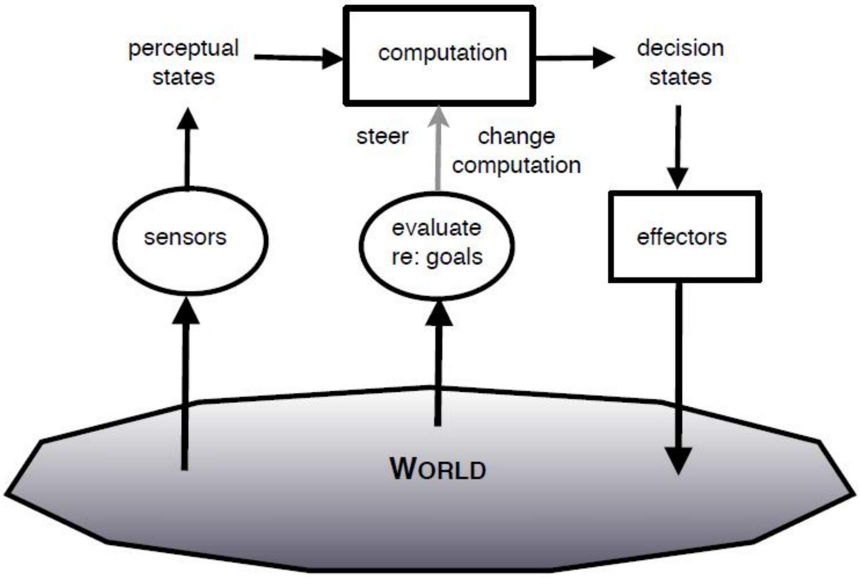


Figure 5.4: Combinatoric emergence; an adaptive robotic device that modifies its computational input-output mapping contingent on its evaluated performance (Cariani, 2012).

However, such combinations are closed. They are always based on a fixed number of primitives. If the primitives are the letters of the alphabet, combinatoric emergence consists in creating new combinations of such symbols which were not observed before, without new symbols being added into the computation pool. Likewise, in computer simulations, a researcher sets up a space of variables and their possible states. But however large in number they can be, the simulation can only recombine such initial values, not add new variables and states than those predefined in the simulation: "pure computation by itself can generate new combinations of symbol primitives, e.g. new string of existing symbols, but not new primitive symbols themselves" (Cariani, 2012).

Within the computational domain, combinatoric emergence, according to Cariani, accounts for trainable classifiers and controllers, neural nets, and genetic algorithms (Cariani, 1992), while in the biology it accounts for short time-scale microevolution based on combinations of existing gene sequences, or structures and functions arisen from novel combinations of existing molecular, cellular, and organic constituents. In psychology, primitive sensations and ideas, combined in novel ways, would be the basis for emergent mental states[47] (Cariani, 2012).

As Cariani notes, "combinatoric emergence is compatible with reductionist programs for explaining macroscopic structure through microscopic interactions" (Cariani, 2012),

---

[47] See section 6.6 below.

and here he refers to (Holland, 1998), adding latter than “in the realm of logic and mathematics, the primitives are axioms and their consequences are deduced by means of logical operations on the axioms. Digital computers are ideally suited for this task of generating combinations of symbol primitives and logical operations on them that can then be evaluated for useful, interesting, and/or unforeseen formal properties. … In virtually all trainable classifiers, the feature primitives are fixed and prespecified by the designer, contingent on the nature of the classification problem at hand. What formally distinguishes different kinds of trainable machines is the structure of the combination-space being traversed, the nature of the evaluative feedback, and the rules that steer the search processes” (Cariani, 2012).

Combinatoric emergence, therefore, connects nicely with the cases of self-organization emergence which involve some sort of novelty in the patterns observed at the system’s level. As it has been discussed, self-organization doesn’t necessarily imply novelty. For an informed observer, the Boids or the Game of Life will not show anything new on a regular simulation. Some prespecified rules generate a pattern of behavior that is read as self-organized by the external observed. Each simulation has its particularities (depending on the initial values of the variables), but the patterns are not novel in any fundamental way. The patterns are emergent in the sense specified by self-organization emergence. Only in the cases were a novel behavior is discovered, e.g. a previously undocumented pattern in Game of Life appears to a cellular automata researcher, it would be a case of combinatoric emergence.

5.5 Semantic Adaptation or Creative Emergence

Creative emergence consists in the appearance of new structures (system states, functions or behaviors) in a system or device, generated through the introduction of new primitives in the computations that direct them. According to Cariani, this is in fact the classic use of emergence, in Broad, Morgan or Bergson's accounts, which were concerned with "processes that create new primitives, i.e. properties, behaviors, or functions that are not logical consequences of preexisting ones" (Cariani, 2012).

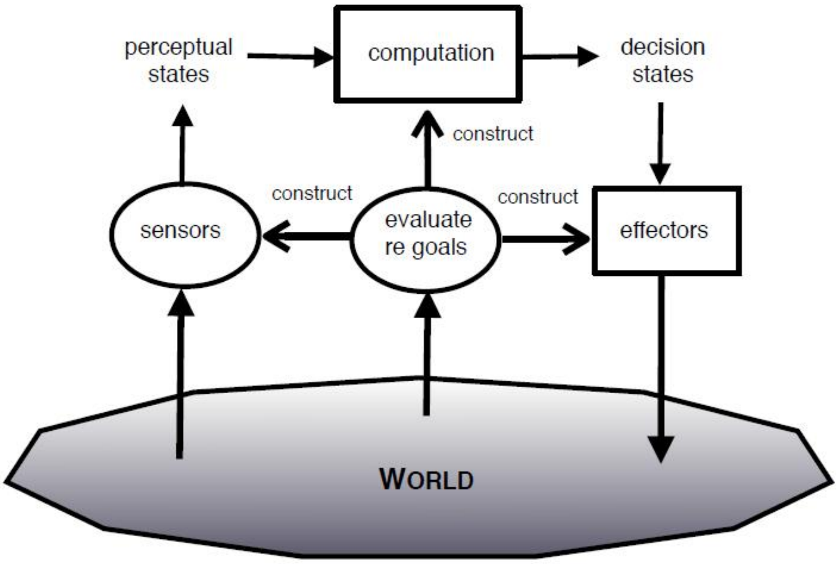


Figure 5.5: Creative emergence; a robotic device that adaptively constructs its sensing, effecting, and computational hardware contingent on its evaluated performance (Cariani, 2012).

In (2012), Cariani acknowledges the importance of the philosophical debate, and recognizes the relationship of conscious awareness to underlying material process, proposed, among others, by Jaegwon Kim. It is a debate that exists on the ontological realm, as conscious awareness is not something externally observable (it is "accessible only though private observables"). Accordingly, creative emergence could be further divided into the appearance of new primitives that require only epistemological reframing, and of those that require also ontological reframing. While the second is unavoidably problematic, the first, according to Cariani's claimed pragmatic approach, is preferred in this context: "More practical than 'de novo' creation of new forms of being is the creation of new functions, which are both verifiable and useful to us – creativity as useful novelty."

In this sense, the evolution of new sensory capabilities in biological organisms becomes the most notable example or functional creative emergence. A new sensory organ in an organism necessarily implies that a series of new primitives (those perceived through the new sensing organ) will enter the domain of computation that will inform its behavioral decisions.

When this is translated to artificial systems, it soon becomes apparent that it is not an easily implementable concept in computerized domains. Only devices capable of building their own sensors and effectors would have the degree of epistemic autonomy to generate creatively emergent behaviors: "to summarize, combinatoric creativity in percept-action systems entails an ability to switch between existing internal states (e.g. 'software'), whereas creative emergence requires the ability to physically modify material structures (e.g. 'hardware') that create entirely new sates and state-transitions (Cariani, 2012).

While, as explained above, examples of combinatoric emergence are easily found in artificial systems, creative emergence is extremely rare. Only one artificial example is discussed in Cariani's literature: Pask's electrochemical device (Cariani, 1991; 1992; 1993; 2008; 2012). As discussed in chapter 4, such device acted as an artificial ear which was capable of building its own 'relevance criteria' or, in Cariani's terms, its own external semantic meanings. I.e. the device decided what to pay attention to and how to classify what was received through the sensor which it grew. As discussed, the electrochemical ear was capable of acquiring the ability to sense the presence of sound, and even to distinguish between two different frequencies. This example is, for Cariani, the "proof that creative emergence is possible in adaptive devices" (Cariani, 2012).

A similar example, presented fifty years after Pask's, is the evolved radio developed at the University of Sussex by Jon Bird and colleagues (Bird et al., 2003). This group of researchers, inspired by Cariani's taxonomy of robotic devices, pursued the creation of epistemically autonomous hardware, in what they called the 'unconstrained intrinsic hardware evolution,' "a design method in which the fitness of electronic circuits is determined by instantiating and evaluating them in actual hardware rather than in simulation." It is an exploratory approach, which deviates from the common engineering strategies, in order to "explore unusual circuits with strange structures and dynamic behaviors beyond the scope of conventional design."

With this approach, they were able to create and evolved radio. After some experimentation with oscillators, they found that some circuits, which had achieved good fitness according to the predefined criteria, were however not oscillating in a stable manner. Upon further examination, they found that these circuits had evolved to pick up radio frequencies that were present in the physical environment were the prototypes were being tested. To do so, the circuits had evolved to use some of its components as antennas. In this particular instance, besides the transistors of the evolvable motherboard on which the circuits were constructed, the circuits also utilized the analog switches and the printed circuit boards (Bird et al., 2003).

The authors claim that this is “the second experimental system ever to construct novel sensors through a process of creative emergence”–the first being Pasks’ device–. However, such systems have a fundamental problem, they argue: since these circuits sometimes utilize environmental conditions and component properties that are very particular of a given implementation, the do not always generalize well. But if, to avoid this issue, the evolutionary process is constraint, so that the circuits are more robust, then the possible advantages of unconventional design, in terms of flexibility, are lost. The most important of these advantages being the possibility of construction of new primitives.

However, despite its difficulty to be implemented in artificial devices, creative emergence remains the most attractive idea in respect to producing fundamental novelty in adaptive systems. Understood in Cariani’s terms, the epistemological and pragmatic view emergence is liberated from much of the debate around how exactly emergence occurs when considered ontologically: “Because the epistemological approach is based on a limited set of macroscopic observables that do not claim any special ontological status, there is no necessary conflict with physical causality or reduction to microscopic variables (where possible). No new or mysterious physical processes or emergent, top-down causalities need to be invoked to explain how more complex organizations arise in physical terms or why they can cause fundamental surprise in limited observers. The novelty that is generated is partially due to internal changes in the system and partially due to the limited observer’s incomplete model of the system” (Cariani, 2012).

5.6 Emergence and Aesthetics: Creation of New Concepts in Minds

As seen above, creative emergence in artificial devices is an extremely difficult task. In contrast, if properly contextualized, it can be a strategy to understand the creation of new concepts in human minds, as Cariani's latest writings suggest (Cariani, 2009b; 2012). He proposes aesthetics as a first context to understand this idea. As he admits, there is a difficulty in distinguishing between combinatoric and creative emergence in aesthetic realms, mainly because of the indefinite spaces of generative possibilities and the ambiguities of human interpretation and expectation. However, to the extent that there are certain existent cultural conventions, it is possible to delineate what conforms to the expectations of the public in the aesthetic domain.

The idea here is that combinatorial creative works operate within a set of primitives, the stylistic or generative rules of the artistic context in which they exist, thus exploring the creation of new forms through combining these elements that exist within an established framework. Emergent creative works, instead, "break conventional, stylistic rules and may violate basic expectations related to the nature of the aesthetic experience itself. On thinks of the Dadaists and the world's reception of Duchamp's urinal as a found-art object" (Cariani, 2012).

Here, the objective would be to find the right balance between novelty and predictability. Novelty must be accommodated by the receiver of the artwork, and this can turn into unpleasantness if the newness is too much. Following the ideas of (Matchotka, 1980) and his study of nineteenth century psychologist Daniel Berlyne, Cariani explains this with the Wundt curve plotting the empirical psychological data related to the relation between arousal (as related to novelty) and experienced pleasure: "Low novelty produces boredom, low arousal, and low pleasure, while extremely high novelty produces high arousal that is experienced as unpleasant. Between the two extremes is an optimal level of novelty that engages to produce moderate levels of arousal that are experiences positively. The degree to which a new piece shocks (and its unpleasantness enrages) its audience is an indication of how many expectations have been violated. An individual's response tells us something about the novelty of the piece in relation to his or her own Wundt curve" (Cariani, 2012).

Cariani proposes two ways in which art can produce creative emergence in (2009a). The first is the creation of artistic autonomous objects capable of evolving new primitives by themselves. Such creative emergence can or cannot be intended by the artist. As a prominent example, he cites Sympathetic Sentience, a robotic sound installation by Simon Penny (Penny, 1995). In the installation, a series of devices sense the sounds in their environment (as other devices produce them), and repeat them with a certain delay, while modifying some and creating their own sounds for other devices to hear and repeat. This complexifies over time as sounds reverberate in a closed loop among the devices[48]. According to Cariani, this process is highly reminiscent of neuronal pattern-resonances in the brain, understood “as a set of interactions between sets of neuronal assemblies that are emitting their own annotative signals. A triggering sensory event thus generates a process of semantic elaboration and eventual convergence to a stable set of pattern resonances that becomes the final ‘meaning’ of the event. The signals continue to reverberate and to regenerate themselves within the network in a manner that parallels [an] autopoietic model of consciousness” (Cariani, 2009a).

The second way for art to produce creative emergence is the amplification of the audience’s conceptual landscape mentioned above. The idea is that certain art pieces can provoke new ideas, meanings and perspectives in their audience, as happened when the art audience was first presented the Duchamp’s fountain.

With this conceptualization, Cariani opens here a door for emergence in interactive art, in considering that the combination of computerized systems and human interactors can be generators of creative emergence, despite the limitations that the computers have in principle in terms of generators of emergent novelty: “As entities in and of themselves, digital computers and formal systems are therefore bounded and closed, but in collaboration with human beings, they can greatly facilitate formation of entirely novel ideas in their human collaborators. In turn their human collaborators can add new primitives to expand their state-sets. Thus human-machine combinations can be open-ended systems that generate new primitives. Enhancing human creativity using flexible human-machine interfaces and other ‘tools for creativity’ is a much more efficient route at present for generating open-ended novelty than attempting to build autonomous self-organizing systems that are creative in their own right” (Cariani, 2009b).

Starting from this idea, Cariani develops a theory of how new primitives can be generated in neural systems and, thus, in the human brain, in order to sustain the idea of how these new concepts in minds are specifically created. He argues how the brain is mostly a combinatorically creative system and that this sort of creativity is in fact achieved effortlessly in our every-day lives: concepts are constantly recombined in order to generate new ideas. In parallel, entirely new concepts can be formed, in a process which, through experience, expands the dimensionality of our concept system. The details of these processes, which can be found in (Cariani, 2009b; 2012), expand beyond the scope of the present study.

---

[48] See chapter 8 for a more detailed explanation of Sympathetic Sentience.

5.7 Related Approaches

The following is a brief account of some of the approaches to emergence that either directly or indirectly relate to emergence-relative-to-a-model.

5.7.1 Emergence as Surprise

The idea of surprise can be used to intuitively explain the idea of emergence as generation of novelty (Penny, 2009; 2010). If what is emergent is new, it follows that whomever perceives this newness is likely to be surprised. This is usually only a discursive resource to explain the idea of emergence, especially in approaches like Cariani's which revolve about novelty, but for Edmund M. A. Ronald and colleagues it became a central idea in order to describe a method for accounting for emergent phenomena.

Ronald, Sipper and Capcarrère (1999a; 1999b) developed a proposal that identifies emergence with the surprise of the observer of a system. Theirs is clearly an epistemological approach, but it could be argued that they move the epistemological focus from the observer's objective instruments to the observer's subjectivity, in a proposal that has been labeled as subjective emergence (Monro, 2009).

The authors of the idea introduce the emergence test to determine whether or not emergence is present in the behavior of a system, in an analogy to how Allan Turing proposed the Turing Test to determine if a machine was capable of thinking. If the Turing test was, in a nutshell, an attempt to see if a computer was capable of passing as a human in a text-based conversation (Turing, 1950), the emergence test attempts to surprise someone familiar with the design of a particular complex system (in the context of Artificial Life) by running the simulation of such system.

The test consists of three steps: first, the designer of a system (or someone who is 'fully aware' of its design) describes the local interactions among the system's components (the micro level). Next, the simulation is run and the same person describes the global behaviors of the system (the macro level). Finally, if the causal link between what was described in the first step and what is observed in the second is not obvious to the observer, cognitive dissonance (i.e. surprise) occurs, and thus emergence is accounted for.

The authors of the proposal consider that their approach broadens Cariani's emergence-relative-to-a-model by including the observer's internal expectations into what is taken into account to describe emergence. However, Cariani does explicitly claim that his approach aims to create a description of emergent phenomena that is valid in terms of scientific communication. That is, that whatever were considered to be emergent by a particular observer would be considered emergent by another one using the same model of the system that is being studied. In contrast, internal states, as said above, are problematic to assess, due to the ambiguities of human expectation and interpretation (Cariani, 2012). Therefore, the psychological state of surprise, while useful to intuitively describe emergence, is not a suited parameter for a theory that aims at generating scientific discourse on emergence.

Still in a strictly epistemological realm, generative artist Gordon Monro elaborates on this notion and expands the idea of surprise to a more complex concept combining it with wonder, mystery and autonomy. Monro proposes a definition of emergence from the perspective of generative art that consists of two states, both of which assume an observer with complete knowledge on the design of the observed system. Firstly, the work produces a result or behavior which is “unobvious or difficult to predict”. And secondly, this result or behavior “evokes feelings of surprise-wonder-mystery-autonomy” in the observer (Monro, 2009).

Again, this is completely dependent on the observer’s internal states and the same critique of Roland, Sipper and Capcarrère’s approach could be applied here. Monro’s surprise-wonder-mystery-autonomy is beyond any scientific discourse. Once more, the weight is moved to subjectivity, and just like the preceding theorization it falls short of concretion. Neither of these two approaches explains how the design must be specified or how the internal states they rely on shall be accounted for.

5.7.2 Steels

Luc Steels is an Artificial Intelligence and robotics researcher who participated in the early stages of Artificial Life. While AI had been traditionally inspired by mostly logics, cognitive psychology and linguistics, he defends the importance of AL as a catalyst for the interest of a subgroup of researchers within the Artificial Intelligence which embraced its approach to intelligence (Steels, 1995). An approach that is based mainly in generating self-organized (emergent) intelligent behavior. That is, behavior that is read as intelligent by an observer as multiple agents locally interact (e.g., Reynold’s boids), or parts of a robotic device (such as Walter’s tortoises) perform simple local operations.

Steels is interested in the modeling of the behavior of agents that are capable of self-preservation in a dynamically changing environment (i.e. of adaptation). In this context, a behavior system is the set of mechanisms that play a role in establishing a particular behavior. In order to properly define such notion, he proposes a distinction between functionality, behavior, mechanism and component as follows:

- Functionality. It is a task or goal that the agent needs to achieve (e.g., obstacle avoidance or signaling another agent).

- Behavior. It is a regularity in the interaction dynamics between the agent and the environment (e.g., maintaining a certain distance from the wall). Several behaviors can contribute to realize a functionality.
- Mechanism. These are the principles or techniques for establishing a particular behavior (e.g., the use of a map or a particular mapping between sensing something and the response to it).
- Component. It is the physical structure or process that is used to implement the mechanism (e.g., a sensor or an actuators, data structures or program).

When an agent is modeled, according to Steels, it shall be made with these four levels of complexity in mind. The behavior system consists of a series of dynamic and static structures. Sensors and actuators, as well as networks, temporary states, electric signals propagating through this network, and the software program that runs in it. All of this is built in order to allow the device to pursue accomplishing the defined functionalities.

Steels identifies four main guidelines in designing such systems in (Steels, 1995), derived both from the experiences of classical AI and the new approaches of Artificial Life. First, behavior systems shall be as specific as possible. Knowledge tailored to a specific task and domain will be more effective than trying to provide the device with a general scheme. Second, exploit the physics. Clever morphology designs can help to solve problems that would otherwise require strong computations (e.g. Walter's tortoises bumping when the creatures encountered an obstacle, which might eventually get it out of a dead end situation). Third, do not think of the device's sensing and acting as symbol processing. E.g. instead of making a rule that if the robot is at a certain distance of an object it should turn this much to the right or left, one can just couple the readings of the infrared sensors (which read the distances) to the motors driving the direction of the robot, in what Steels refers to as a 'truly subsymbolic' level.

Finally, a fourth guideline points directly to self-organization emergence: design simple mechanisms and rely strongly on their interactions, in order to generate complex behavior. The idea here is to avoid complex objective models of the world to guide the device's behavior. Instead of this, design simple elements in interaction and let them act on the world and see how they achieve the task. There are different approaches to doing this: through neural networks, algorithmically, with circuit design or understanding behavior systems as dynamical systems. All of these are reviewed in (Steels, 1995).

With these categorizations, guidelines and approaches, the design of a particular agent can be undertaken. And once it has been created, Steels identifies two ways in which the agent can become more complex. First, a designer might indentify a new functionality and act on the agent in order to provide it with the means to attain it. The second case is emergence: "existing behavior systems in interaction with each other and the environment can show side effects, in other words, emergent behavior. This behavior may sometime yield new useful capabilities for the agent, in which case we talk about ‘emergent functionality’” (Steels, 1995). Steels’ approach is very similar to that of Cariani. Here emergence is not about the overall self-organized patterns of behavior of the groups of agents, but the novelty that appears in each one of them: the ‘side effects’. Once functionalities and behaviors have been established (by design), emergence is the appearance of these side effects, of the new behaviors that happen though the interaction of the agent with others and with the environment.

From an engineering point of view, emergent functionality has some disadvantages. It is less predictable than the functionality of traditional design, and the side effects will not necessarily be positive. But, according to Steels, this approach is the only one with which an agent can autonomously increase its capability, an idea that is essentially the same as Cariani’s epistemic autonomy.

Also along the same lines as Cariani, Steels defines emergence from the point of view of the observer as the need to add new categories in order to describe changes in the behavior of a system. And like in Cariani (1992a), the separation of levels in neither here a necessary condition of emergence: “It is not necessary that the two descriptions (the emergent behavior and the behavior of the individual components) are at different levels, although that is not excluded” (Steels, 1995).

Finally, he introduces another viewpoint to emergence: that of the components implicated in the emergent behavior. In doing so, he distinguishes between controlled and uncontrolled variables. Controlled variables are directly influenced by the system (e.g. the own device’s forward speed). Uncontrollable variables are environment variables which the system cannot influence directly (e.g. the distance to a wall. The device can move to change it, but it cannot move the wall). On top of that, a second distinction is introduced in these uncontrollable variables: they can either be visible or invisible variables. Visible are those variables which the agent can read through its sensors (e.g. distance, temperature, obstacle collision, etc.). Invisible variables are variables which the user can measure but the system can’t. I.e. they are not modeled into the system. Emergent behavior involves either both or only invisible variables. That is, no component is directly responsible for the behavior. From this point on, if the system is capable of building up onto this, then emergent behavior – semantic emergence– occurs: the system detects, amplifies, and builds upon the emergent behavior (Steels, 1995).

This idea is equivalent to Rosen’s emergence as a deviation of the system’s actual behavior vis-à-vis the predicted behavior by the model. In Rosen, these deviations happen because of the degrees of freedom that the system, but not the model, possesses (i.e. because of the necessary abstraction of the model in respect to the modeled system). The degrees of freedom which are not accounted for in the model are Steels’ invisible variables.

5.8 Concluding Remarks

Cariani’s approach to emergence offers a unique effort to formalize the discourse around emergence. Often, this concept is used in a manner that leaves room for the intuition of the reader, and that makes scientific communication difficult. Cariani aims to eliminate this problem by defining in detail what is meant by emergence and in what concept. In doing so, he concentrates on emergence as generation of novelty, and this leaves out a big part of the discourses around emergence, which are centered on the idea of self-organization emergence.

However, these two views are compatible. If a sound definition of what is meant by self-organization is elaborated, Cariani’s method can be expanded in order to formalize a system of analysis which can account for both self-organization emergence and emergence as generation of novelty. Once this is done, this method can be applied to interactive systems to analyze whether or not they present emergent phenomena of either kind. The rest of this PhD proposal is devoted to this.

PART 2: EMERGENCE AND ARTIFICIAL LIFE ART

6. ARTIFICIAL LIFE AND ARTIFICIAL LIFE ART

The first part of this thesis has reviewed some of the scientific and philosophical discourses around emergence that are relevant to interactive media. As it has been shown, the concept is far from being a simple idea. Different disciplines present it in distinct ways, and authors' accounts can be very far apart depending on background, context and intentions. The objective here is not to end this debate, but to discern a coherent and clear way to understand the issue in order to make it useful as a tool for art and digital media creation and analysis. The second part of the thesis is dedicated to this.

This chapter presents the first part of this endeavor. First, it reviews the historical roots of Artificial Life, to move afterwards to a taxonomy of Artifical Life Art based on the techniques that the artistic branch of ALife has borrowed from the discipline. It also includes, in each of the four sections of the taxonomy, a review of some of the most relevant artwork in order to illustrate both the taxonomy and the discussion that will be presented in the following chapters. In some cases, detailed discussion on the artwork's interactivity and emergent properties is left for chapters 7 and 9.

6.1 The origins of Artificial Life

Before delving into the discussion of how ALife and ALife Art are related, it is useful to discuss how ALife originated and how concepts such as life and simulation were understood within this movement. Artificial Life as a discipline and with such name, originated in the 1980s, but it has been argued that its roots are much older than that. This section focuses on these historical roots, especially regarding the work of Jessica Riskin on eighteenth century automata, with a focus on how the technological and cultural context of ALife influenced the conceptual basis that lies at the basis of the discipline.

The history and context in which ALife originated has been reviewed in detail (Langton, 1988; Levy, 1992; Waldrop, 1992; Helmereich, 2001; Penny, 2010; Solé, 2012), in some cases with a special emphasis on its artistic ramifications (Whitelaw, 2006; Reichle, 2009; Penny, 2009). As discussed in chapter 3, Artificial Life became a formal research field in Santa Fe Institute in 1987. Christopher Langton, the head figure of the movement, defined the field in the following year: “Artificial Life is the study of man-made systems that exhibit behaviors characteristic of natural living systems. It complements the traditional biological sciences concerned with the ‘analysis’ of living organisms by attempting to ‘synthesize’ life-like behaviors within computers and other artificial media. By extending the empirical foundation upon which biology is based ‘beyond’ the carbon-chain life that has evolved on Earth, Artificial Life can contribute to theoretical biology by locating ‘life-as-we-know-it’ within the larger picture of ‘life-as-it-could-be’” (Langton, 1988).

From this idea, the differentiation between carbon-based live and silico-based live originated. A fundamental premise in Langton’s proposal is that the basic characteristics of live are independent from matter. According to this, carbon-based live (the life-as-we-know-it) is only a subset of life, and silico-based is an abstraction from matter which maintains on it what is really essential in live. Hence Langton’s claim that Artificial Life can widen the scope of theoretical biology. Therefore, Artificial Life, at least in Langton’s initial formulation, understands that (the quality of) live is independent from the material substrate on which it is sustained. The ‘essential properties of live’, he argues, are in fact better understood if we can abstract them from matter, which is not essential to live but accidental (Langton, 1988). Pattee’s and Cariani’s critiques of Platonism come to mind here[49]. Under this idea of ALife, life in a computer is not just a simulation or an abstraction of the live that exists in nature. It deals with life as much as if it studied natural living beings. Langton states that, while biology is concerned with the ‘material’ basis of live, and moves analytically from top to bottom (from the living being to its organs, etc.), Artificial Life is concerned with the ‘formal’ basis of live, and moves on the opposite direction, from the simple to the complex, synthetically. And

---

[49] See chapter 5.

here is where Langton places emergence as a central concept within the discipline, as discussed earlier, since it is the key idea in order to explain how this bottom-up process from the simple to the complex takes place.

The historical roots of ALife are the "history of attempting to map the mechanics of [mankind's] contemporary technology onto the workings of nature, trying to understand the latter in terms of the former" (Langton, 1988). Langton situates in ancient Egypt the first attempts to create behavioral machines. Their water clocks, known as Clepsydras, were technological devices that mapped the progress of time. He also cites Hero of Alexandria and some of the human and animal-like shaped gadgets described in his pneumatics treatise. But, as he points out, the first serious attempts to simulate the mechanics of life were the eighteenth century automata, which imitated both animal and human movements and functions.

In fact, as noted by Jessica Riskin (2003a), the term simulation is an anachronism when applied, in this sense, to eighteenth century technologies. At that time, simulation had a negative connotation and was thus avoided in writing about the automata. However, the goals of their efforts and how they approached the subject (modeling nature, experimenting with the models and drawing conclusions about what was being modeled) coincide with what the twentieth century understood by simulation.

Riskin moves two centuries back Owen Holland's claim that Artificial Life existed already some decades before Langton coined the term (Holland, 1997; 2003). She considers that the eighteenth century experimentation with automata is not just a far antecessor, but an effort worthy of the Artificial Life label. The long intervening period between the late-eighteenth century and the late twentieth century ALife, according to Riskin, responds to the changes in the larger cultural context. The two moments of Artificial Life, despite some clear similarities in theme's, goals and approaches, are rooted in radically different cultural (and we could add: technological) discourses. It is not by chance that both movements found their support, and build their metaphor upon, the 'new technologies' of their time: sophisticated clockwork machinery in the eighteenth century; computers on the twentieth.

Mid-eighteenth century ALife "was crucially informed by a particular philosophical development, namely a materialist, mechanist understanding of life and thought. Materialists repudiated Descartes's separation between mind and body, and insisted that all the functions that might be ascribed to mind and soul actually resided in the stuff of which living creatures were made. Mechanists argued that interaction among the body's parts, animal machinery, was directly responsible for all vital and mental processes" (Riskin, 2003a). As said above, the twentieth century ALife is, at least to some degree, informed by a platonic view of which resides precisely on the dichotomy between mind and body. A separation that, as it has been argued, has become pervasive in many computer related practices, ALife among them, in the form of the differentiation between hardware and software (Penny, 1996). In this respect, the eighteenth century movement connects to the theories of embodiment.

Despite the differences, however, the mechanistic and materialistically informed movement, like the twentieth century ALife, did transpose too the qualities of live into the devices that served as the basis of the simulation. Even though they did so without grandiose claims equivalent to Langton's, there was a sort of vification of the machine embedded in the practices of designing animal-like automata in this manner: "if live was material, then matter was alive, and to see living creatures as machines was also to vivify machinery" (Riskin, 2003a).

The automata of the second half of the eighteenth were, according to Riskin, successors of earlier work that, instead of aiming at simulating the animal and human behaviors and processes, simply represented them. An example of this was Maillard's mechanical swan, which paddled and moved its head with internal mechanisms that in no sense tried to replicate the animal's internal organs and structures. It was build with gears and other mechanisms with the sole purpose of creating an external appearance that resembled the movements of the swan. In contrast, later automata did try to imitate the inner workings of the modeled beings: piano players, scribers, talking heads, etc. tried to imitate larynges, bones and all sorts of organic structures. A paradigmatic example is the work of Ambroise Paré: prosthetic legs, arms, and especially the hand, which was full of gears and mechanisms in order to try to imitate the complexity of the human hand (Solé, 2009). Some of the automata were programmable, like Vaucanson's flute player, which could play different tunes when the pinned cylinders were changed.

In fact, the physiological simulation, characteristic of the eighteenth century automata, was not followed by the second ALive movement in the twentieth century, were the focus was never on physiology but on behavior. From Grey Walter's tortoises to Simon Penny's Petit Mal or Ken Rinaldo's Autopoiesis installation (both discussed below), the interest has not been in the interior of the devices but on behaviors: on how they interacted with their environment and, except in the first case, with the interactors.

In contrast, the inner works were such a concern in the earlier automata that most of them had a way to show the inside of the devices, mostly in order to defeat the (not always gratuitous) accusations of trickery. The first largely famous automata was Jacques de Vaucanson's Duck, which could cavort with its wings and bill, stretch and bend its neck, flex its feet, and even drink water. But more remarkably, it had a simulated digestive system with which it could eat grain and, a little later, expel excrement. Vaucanson claimed that the duck was indeed capable of digesting the food, and this is what made it famous all over Europe. The feathers of the duck were made with a semi-transparent material so the public could examine (up to some degree) its inner workings. As said, the duck became very famous, but it was not too long until it was discovered that, in fact, it was doing nothing at all with the ingested grain. Instead of following through the simulated digestive system it was stored. And the excrement, which had been previously loaded, was taken from another container (Riskin, 2003a; 2003b). Another very famous automata, perhaps the only one to be more so than the duck, was again a deceptive construction. The Mechanical Turk was an incredibly talented chess player automaton, which went from hand to hand for more than a century. During each show, the interior of table that contained the mechanisms was shown to the audience, one compartment after another. What was inside, in fact, was a hidden person that moved around as the compartments as they opened and that operated the Turk's hand during the games (Standage, 2002).

The discovery and publication of these deceptions contributed to the decline of this first golden era of Artificial Life, and obscured the legitimate achievements that had taken place in it. But this decline has deeper reasons in the larger cultural context, as changes in how machines and animals were understood, in part motivated by this very same research, modified the relevance that was given to such endeavors. Eighteenth century automata were a reflection of a moment when animals were starting to be understood in terms of machines, and the same time machines were starting to be understood as animal-like. It is, according to Riskin, "one moment in an ongoing dialectical engagement between our understanding of life and of machinery, in which living creatures and machines have continually redefined each other, both by being identified with each other and by being opposed." And as said above, it was the same practices that helped modify the context in which they originated: "it was the articulation of certain differences between natural and artificial life that triggered the invention of machines that undermined those differences. But these machines in turn led people to rethink what constituted life, and to redefine natural life by contrast with artificial life. … And so people's assumptions about what is essential to life and what is within the purview of machinery have continually transformed each other" (Riskin, 2003a).

Riskin's account of the following development of the dialectics of the separation of the animal and the machine are illustrative of the loss of momentum and latter recovery of the Artificial Life efforts. In the early nineteenth century, the critics of this first ALife phase reclaimed the Aristotelian principle that animal life differs from machines in that the former are capable of producing their own source of power for moving, while the latter are not. All the automata had to be powered externally, in one way or another. But the development of the concepts of energy and its conservation eliminated this argument.

Another claim against early ALife was that only animals, and not machines, were able to respond to changes to their environments, to adapt and continue to operate their basic functions. This difference would again be undermined, this time by the first impulses of the second wave of Artificial Life: the responsive machines of early Cybernetics, such as Walter's tortoises, that were discussed in chapter 4 and which have been presented in the literature as predecessors of Artificial Life (Owen, 1997; 2003; Penny, 2009; 2010).

However, Cybernetics is not the only predecessor of contemporary ALife. The twentieth century ALife was again a moment when machines and nature could be understood under the same reading glass. This was, arguably, the case of John Von Neumann, another agreed-upon predecessor of ALife. During the later part of his short but remarkably fruitful career, Von Neumann, among other things, concentrated on the problem of creating a self-reproducing machine. The machines he was interested in were the automata: self-operating machines whose behavior can be defined algorithmically. According to Steven Levy (1992), Von Neumann "saw no reason why organisms, from bacteria to human beings, could not be viewed as machines –and vice versa." Indeed, Von Neumann would mention both natural and artificial automata (Von Neumann, 1966) assuming that one could be understood in terms of the other. Therefore, his efforts on the problem of self-replication automata can be understood as the problem of self-replication in general. In solving it along with his colleague Stanislaw Ulam, Von Neumann invented what would become known as the first cellular automata: a mathematical abstraction of automata and of computation which were to become one of the basic building blocks of Artificial Life some decades later. Langton saw in cellular automata a perfect example of what ALife was: "Cellular automata are good examples of the kind of computational paradigm sought after by Artificial Life: bottom-up, parallel, local-determination of behavior" (Langton, 1988). The work on cellular automata by Langton and Wolfram was discussed in chapter 3. Its influence on Artificial Life Art is reviewed below.

Regarding the similarities between twentieth and eighteenth century ALife, Riskin points out the sensationist doctrine of materialist philosopher La Mettrie, which constitutes, in contemporary terminology, an embodied approach to knowledge. La Mettrie, and along him the materialism that was informing the development of eighteenth century ALife, considered, against the Cartesian dualism, that "ideas were not innately implanted in the mind, but were created by the action of the senses and the nervous system and, therefore, could not be abstracted from the body" (Riskin, 2003a). This resonates, she argues, with twentieth century ALife's "principle that intelligence must be 'physically grounded' and 'embodied,'" and with "Brooks' claim that the software of the mind cannot be abstracted from its hardware." In a word, as said above, of embodiment which is, for instance, at the heart of Brooks' situated robotics[50].

Despite the parallelisms, twentieth century ALife (from now on, again, ALife) does have an enormous difference of scale with its eighteenth century predecessor. As mentioned above, the cultural context was a decisive factor, but it also was the available technologies at the time. The automata can be largely regarded as amusements that probably, as Riskin claims, influenced up to some degree the philosophical and scientific developments of their time, but they did not, by far, establish a scientific field

---

[50] See section 6.6.

of their own, nor were they particularly influential in the advancement of technology. When Langtons' ALife appeared, it did so in parallel, and because of, the availability of computation. As Simon Penny has noted in (2010) "Artificial Life is 'native' to computing in the sense that large scale iterative process is crucial to the procedures which generate (most) artificial life phenomena." The field, which is certainly not mainstream in academia, does have some degree of autonomy and recognition, and its developments have had some influence in various fields, mainly in some aspects of computer science (e.g. the use of genetic algorithms as a computing strategy) and especially in robotics. And, of course, it has allowed for the appearance of ALife Art, to which the rest of this chapter is dedicated to.

6.2 A Taxonomy of Artificial Life Art

Since the early stages of ALife, artists borrowed the discipline’s techniques and adapted them in their work, both in methods and themes. Because ALife flourished while (and because of) computing was becoming widely available, it coincided and merged with the initial stages of digital art. The term ‘artificial life art’ was, according to Ingeborg Reichle (2009), introduced by Kenneth Rinaldo in 1998 on a special printed issue of Leonardo. In it, Rinaldo states that “Artificial Life artworks could be considered as a subgroup of Artificial Life research in that most artists are more concerned with creation of an aesthetic as opposed to testing theoretical biology” (Rinaldo, 1998). This section reviews the most relevant techniques that artists have borrowed from ALife and exemplifies it with ALife Artwork. The following table should help illustrate the issue.

The taxonomy presented here does not claim to cover every single instance of Artificial Life Art nor to allow for an impermeable classification of the existing artwork. Several projects do fall into more than one category. For example, virtual ecosystems are often programmed using evolutionary algorithms and consist of somewhat autonomous agents. The classification is intended to stress the most relevant aspect of each piece, and to be inclusive.

Thus, the basis of this taxonomy is the general technique employed in the creation of the work. The proposed classification is based on those previously elaborated by Mitchell Whitelaw (2006), Simon Penny (2009; 2010) and Ingeborg Reichle (2003). It combines elements of the three, and in occasions the categorizations fully coincide with these authors. The second row in Table 6.1 makes these coincidences explicit.

Category	**Coincidences with previous classifications**	**ALife examples and authors**	**ALife Art examples**
Abstract Generative Systems	Abstract machines (Whitelaw)	·Von Neumann's self-reproducing machine ·Game of Life ·Wolfram	·Breed ·Rule 30 ·Human Cellular Automata ·Sand Lines ·Substrate ·Propagaciones ·The Conversation
Evolutionary programming and genetic algorithms	Evolved painting, sculpture and animation (Penny) Breeders (Whitelaw) Evolutionary image process (Reichle)	·John Holland's Classifiers	·Genetic Images ·Evolved Virtual Creatures ·A-Volve ·Autopoiesis ·Performative Ecologies
Simulated ecologies	Virtual ecologies (Penny) Cybernatures (Whitelaw) Artificial worlds (Reichle)	·Tierra	·Tecnhosphere ·Iconica ·Digital Babylon
Autonomous agents	Physically instantiated ALife systems (Penny) Hardware (Whitelaw) Living sculptures and robotics (Reichle)	·Brooks ·Steels	·Petit Mal ·Sympathetic Sentience ·El Ball del Fanalet /Lightpools ·Tickle Robots ·Sniff

Table 6.1: ALife techniques and their correspondences in ALife Art.

The examples appearing in the rest of the chapter will be used as case studies for the remaining chapters of the thesis, and referred to with the assumption of familiarity that its presentation here permits. Many of these art pieces have been awarded in VIDA, a festival on ALife Art organized by Fundación Telefónica in Madrid since 1999, which has become a major reference point in this artistic discipline. The choice of examples does not attempt to include all relevant ALife Art work in each category, but to show a broad spectrum of them. The intention is that some of these examples, too, will serve as the basis of the discussion of the model for analyzing the presence of emergence, which will be presented in chapter 9.

6.3 Abstract Generative Systems

Abstract generative systems is a descriptor that includes artwork that is based on cellular automata (CA), fractals, L-Systems and recursive systems in general and similar techniques. These systems attempt to model functions that are present in living systems and build up the artworks from there. Simulating or emulating these functions as they are found in nature, or mimicking the external appearance of the modeled living systems is not necessarily a goal in these pieces.

As said above, the origins of CA are found in the work on mathematician John Von Neumann, whom in the in the 1940s set up to create a self-replicating machine with the intention of understanding the logic of the process, much in consonance with Langton's latter claims that the fundamental processes of life could be abstracted from matter. Neumann's question, as formulated by his colleague Arthur W. Burks, was: "what kind of logical organization is sufficient for an automaton to control itself in such a manner that it reproduces itself?" (Burks, 1970). After ruling out the possibility of building it with the available technology of the time, Von Neumann decided to create a model of it, in order to concentrate on the logic behind the process (Von Neumann, 1966). He first attempted to do this with a kinematic model: a robot that would fetch and assemble pieces in order to build what it was told to, including a copy of itself. But the kinematic nature of the modeled device created the problem: either too much had to be taken for granted (how does the soldering works, how do the sensors operate and communicate with the arm, etc.) or too much details had to be explained. If the first was the case, then the problem with this model was that it had "too many black boxes" (Levy, 1992: 29). Burks, instead, assumes that nothing could have been left as a black box, but this created the problem that too much had to be explained in order to maintain the model within some reasonable boundaries of simplicity: "the motional capability of the kinematic system is the source of complexities which, in the present context, are not worth their cost and might better be eliminated" (Burks, 1970). Hence, the kinematic model was abandoned. But Von Neumann didn't abandon the project. Instead, he moved on to propose a more formal one. This second attempt, which would turn out to be a success, constituted the creation of cellular automata.

Cellular automata are one type of finite state automata: "In brief, a CA model consists of a regular lattice of 'finite automata,' which are the simplest formal models of machines. A finite automaton can be in only one of a finite number of states at any given time, and its transitions between states from one time step to the next are governed by a 'state-transition table': given a certain input and a certain internal state, the state-transition table specifies the state to be adopted by the finite automaton at the next time step. In a CA, the necessary input is derived from the states of the automata at neighboring lattice points" (Langton, 1988). The most well-known example of cellular automata is John Conway's Game of Life, which was presented in chapter 1. As it was reviewed in chapter 3, CA were a basic building block in the work of Christopher Langton and Stephen Wolfram. Besides these and other theoretical implementations, CA have been used in other computation domains such as computer graphics. E.g. to simulate the growth of plants in a 3D implementation (Greenem 1989) utilizing the capacity of the CA for simulating branching patterns (Prusinkiewicz and Lindenmayer,2004).

Along with CA, fractals are a second basic building block for artistic abstract generative systems. Fractals were introduced by Benoit Mandelbrot in 1967 with his famous paper *How Long Is The Coast of Britain?* in which he described how many phenomena in nature can be explained with the idea of recursive self-similarity. The coastline can be described with the same mathematical formulas at any scale: from the drawing on a map to contour of a grain of sand. Hence the rhetorical question of Mandelbrot's title. The coast of Britain's length depends on the scale that is used for the measure: the finer the features taken into account, the greater the length of the coast (Mandelbrot, 1967). Mandelbrot's work, especially the book he published 10 years later (Mandelbrot, 1977), pushed fractals into popularity. While the term was coined by Mandelbrot at the time, the basic mathematics of symmetry across scale was by then already a century old, and had been quite well-known in the early twentieth century (Penny, 2009).

Fractals have influenced several disciplines and Artificial Life is among them. The principle of self-similarity was used in animation from the 1980s onwards to algorithmically create trees, coast lines, river meanderings and mountains in 3D animation and movie scenes. Fractal art is the use of fractals for aesthetics purposes – often with quite psychedelic results. To the extent that fractals and recursive systems can be used to model natural structures, fractals are consistently considered to be among the basic techniques of Artificial Life.

The proposed taxonomy includes in the Abstract Generative Systems label the use of either cellular automata or fractals, or the combination of both. There are mainly two ways in which artists have approached CA and fractals. The first is a more or less literal representation of them, such as visual art based on creating fractal geometries, or the cases of physical instantiations of cellular automata like the sculptures of Erwin Driessens and Maria Verstappen, Kristoffer Myskja's Rule 30 or Matthew Fuller's happenings based on Game of Life. The second are reinterpretations or projects that build on the general idea of the abstract machines that cellular automata represent.

6.3.1 Erwin Driessens and Maria Verstappen: *Breed* (1995-2007)

The Dutch artist duet Driessens & Verstappen has produced some significant Artificial Life Art pieces. Among them, Breed is a series of sculptures that combines cellular automata and evolutionary programming. The first instances of the sculptures were handmade with plywood, and later models were made from nylon or metal, with more sophisticated building techniques. The construction of Breed is based on two principles, inspired in CA and genetic algorithms respectively. The combination of both controls the formation of the sculpture.

Initially, the material is a solid cubic block. Then, a series of steps take place in which the sculpture’s computer model is shaped. In each of these steps, the cube is divided into eight smaller cubes. During the successive divisions, a morphogenetic rule determines, depending on neighboring cells or empty spaces, whether or not each of the eight the newly created spaces will be filled. These are cellular automata rules applied in a 3D space in a process of recursive division that will go as far as the program specifies it. Along with this process, the divisions are checked in order to create a viable physical sculpture. That is, since the result must be not a computer model but a physical instantiation of it, it is necessary for it to comply with certain rules: being in one piece, distributed weight so it does not break, etc. This is done with genetic algorithms: iterations are checked against the defined fitness criteria and only those which are successful will be implemented. As noted by Whitelaw (2006), this reengineering of both cellular automata and genetic algorithms is in fact deterministic: each genome (the rules governing how the successive iterations will be generated) will necessarily produce the same results. As it will be explained later, a common approach in order to avoid this deterministic constrain in genetic algorithms is mutation, the random generation of variation that opens the door for non pre-specified changes to occur.

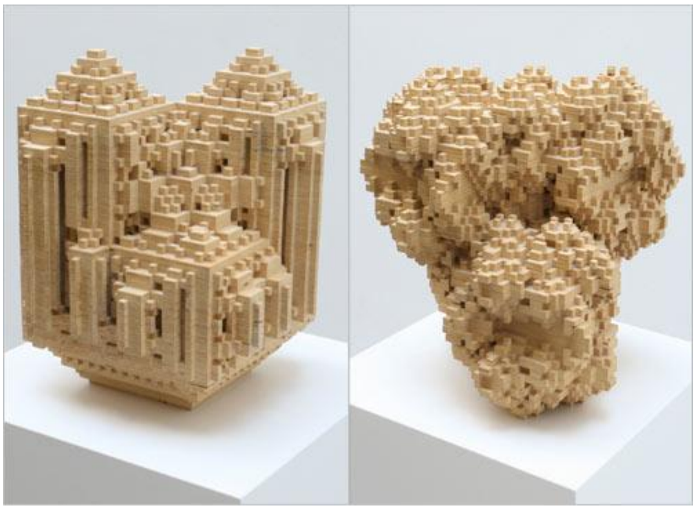


Figure 6.1: A plywood instance of Breed (source: http://www.dataisnature.com/?p=565).

The result is a series of sculptures developed into very different shapes but with a common basic structure that is recognizable. If one is familiar with cellular automata it is easy to guess, in some of the final results, that the underlying rules of the construction are inspired by them. Pau Prudence remarks the similarities of the latter metal models to bismuth crystals, which grow in very straight lines and squared turns due to the relation of electrical charge and growth; a process which is essentially the inverse of that proposed by Driessens and Verstappen (Prudence, 2010). In any case, Breed is a visible and palpable example of one of ALife’s most repeated claims: complexity can be generated from very simple local rules. Both techniques are combined exclusively locally, voxel by voxel, and the resulting sculpture, which was indeed encoded in the initial genome, only reveals itself as the computer simulation advances.

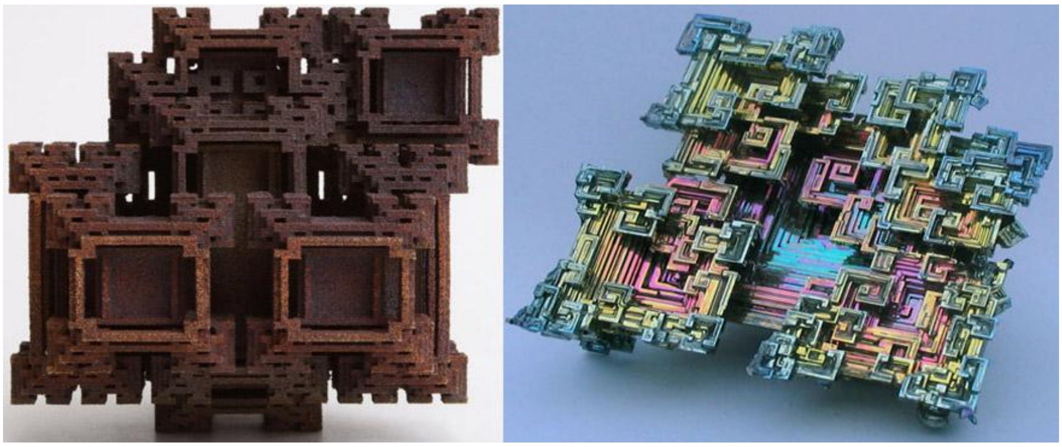


Figure 6.2: A metal instance of Breed (left) and a Bismuth crystal (right) (sources: ttp://notnot.home.xs4all.nl/breed/e234-c.html and http://www.dataisnature.com/?p=565).

Like with some structures on Game of Life such as the R-Pentomino[51], the number of iterations needed for the process to come to a conclusion is too great to anticipate it. And only through the simulation the form appears. As it was discussed, this is enough for some authors to consider the resulting sculptural form as emergent. However, this poses the problem of where to draw the line between where anticipation is possible and where does the need for a simulation start.

6.3.2 Kristoffer Myskja: *Rule 30* (2008)

Kristoffer Myskja is a Norwegian artist who has been producing kinetic sculpture since 2005, inspired in some cases by Cybernetics or Artificial Life. In 2008 he created Rule 30, which received a special mention in VIDA 13. This kinetic sculpture is a mechanical device that prints out the resulting iterations of Wolfram’s one-dimensional automata known as Rule 30. This particular case of CA was a prototypical example of what Wolfram was interested in: a CA that would result in an attractive visual pattern, as chaotic structures repeat in different scales in an apparently random manner.

---

[51] See section 2.7.2.

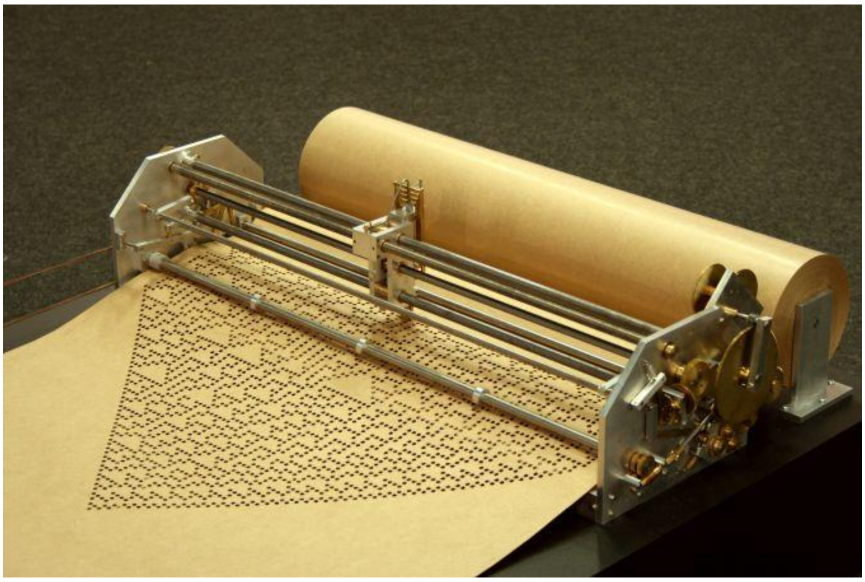


Figure 6.3: Rule 30 (source: http://kristoffermyskja.com/work.php#work-display).

Rule 30 is, and intends to be, a very literal interpretation of cellular automata. What makes it remarkable is the analog, mechanistic approach to an issue which has always been related to computers and represented either in computer screens or in printed final forms (such as in Wolfram's books and papers), or presented with mathematical formulae. Myskja's approach, instead, focuses on the process, as the slow paced machine moves the perforating head to either leave blank or puncture each of the spaces in the paper. The construction of the machine is also reminiscent of early computer systems, which operated with punched tape, and even of the Turing Machine, since Rule 30's rolled tape, which continuously unrolls as the CA's visual representation is punched into it, could theoretically be infinite.

6.3.3 Matthew Fuller: *Human Cellular Automata* (2000-2001)

Human Cellular Automata is a group action that was organized by Matthew Fuller in London (2000) and Berlin (2001). The action is reminiscent of the 1960s and 1970s rule-based happenings and performances and latter reinterpretations, such as Heath Bunting's King's Cross Phone-In, from 1994, where a group of people were instructed, via internet, to either call or receive calls in a group of public telephones on a specific London location. The author recognizes the influences of Fluxus and the work of Vito Acconci (Fuller, 2001).

Fuller's happening is poorly documented, but the idea is easy to imagine from what can be found in the author's website. Human Cellular automata consisted in recreating a simplified version of Game of Life with people standing on a grid with either a paper on their head (the cell is 'on') or without it (the cell is 'off'). On each iteration cycle, each person checks the cells at his or her right, left, front and rear (unlike Game of Life were all 8 neighboring cells are checked). Then, if either two or three neighboring cells are ‘on’, the person stays or changes to ‘on’ state. If not, the person stays or changes to ‘off’. Once everyone has checked his or her state, the next cycle shall be executed. No further instructions are specifically given: “the game shall go on as long as it does go on” (Fuller, 2001).

The experience was presented within a larger event under the name of Software as Culture’ and Fuller presented it as “a program and a game” that shall constitute “an entertaining and surprising example of mass collaboration made possible by a set of simple rules” (Fuller, 2001). The documentation does not account in detail for the outcome of the performances, although Fuller states that when it was first presented in London they “took over a small city square and turned it into the slowest, most high-energy, processor for miles.” In any case, the author had expected them to be entertaining and to develop quite intuitively and fast, thanks to the simplicity of the rules. Hence the slogan: “The human cellular automaton is like a Mexican Wave in two-dimensions.” A Mexican Wave is a phenomenon created by sport stadiums’ fans that, like cellular automata, generates a group pattern from simple local rules. The rule is simple: when the person on your right (or left) stands up and raises his or her hands, do the same, and then sit down again. The effect is a ‘wave’ that moves sideways around the stadium as people stand up and sit down. The name refers to the world-wide popularization of the phenomenon during the Football World Cup of Mexico 1986. Fuller’s comparison, however, leaves out the most important element in the Mexican Wave: spontaneity. The rules of the wave are as simple as they can be and this makes it very easy for it to propagate. This spontaneity is hardly reproducible in the setup proposed by Fuller, which involved a much more complicated rhythm (it doesn’t make sense to move to the next cycle unless everyone is ready) and more complex rules, which implied checking on four persons instead of one52.

In any case, Human Cellular Automata is a good example of the physical instantiation of CA and of making explicit the local character of the rules. It is, like the case of Rule 30, a very literal interpretation; a staging, in fact, of a version of Game of Life. Each participant is a person operating as a cell in a computerized system, and therefore the computation each one does is neatly local. Since it is visible for everybody that there is no other person telling everyone what to do, the overall effect can only be attributed to the accumulated effect of the local interactions. For an external observer, provided that the resulting patterns were fluid enough, this could be read as an emergent phenomenon

---

[52] On January, 2015, I replicated the experience with 42 students in a grid of 6x7. Despite the fact that it proved the idea of generating a game-like experience solely through local rules, and that everyone was playing all the time (as opposed to what it would have been to simulate a Chess game) the results were in general disappointing for those playing the role of a cell. Students within the grid had a hard time keeping track of what was going on around them and errors accumulated, resulting in a truly stochastic system. Even in the moments were it seemed to work, the perception of the players was nothing like the point of view of the observer of a simulation of Game of Life.

in the form of abstract modeling of self organization, as it can be in any cellular automata system.

6.3.4 Paul Brown: *Sand Lines* (2000-2001)

Sand Lines is a 'kinetic painting' by Australian artist Paul Brown. It is one of several works of his based on cellular automata, which he produced from 1973 onwards (Whitelaw, 2006). In Sand Lines, Brown presents a noisy canvas with a series of white curvy lines that slowly but constantly reconfigure their trace and their connections with other lines. As the process happens, simple 8-bit-like noises indicate these new connections.

The name of the piece offers quite a literal interpretation. The lines are drawn onto a surface that resembles sand. A material made out of an accumulation of small stones that in the piece are represented by the grayscale pixels that look like the noise on an old TV. The lines move and reconfigure themselves in a very organic way. There's always a subtle general movement on the background texture, and a small number of lines being reconfigured, while the others remain still for a little while.

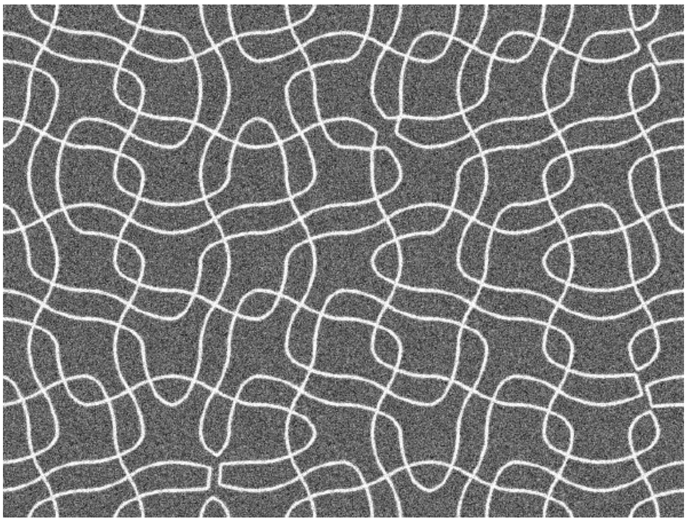


Figure 6.4: Sand Lines (source: http://www.paul-brown.com/GALLERY/TIMEBASE/SANDLINE/INDEX.HTM).

Behind the scenes, what is operating is a cellular automata system. But the system is not shown directly as it was in the previously presented examples. Brown, instead of literally translating cell states into visual representation, maps the transitions of the cells onto the visual canvas. The cells are seen as being born, dying or continuing in the dead or alive state. In the particular case of Sand Lines, also, he smoothed the transitions with an animation, so instead of flickering, the cells slowly transit from one state to another, and this is what produces the animation of the permanently changing lines.

The interest of Sand Lines lies mainly in this indirect representation of cellular automata. It is not any more about showing the graphical output of applying whatever rules to the grid-distributed system, but about using the logic of it as a motor for image creation and, thus, for animation. In contraposition to the previous examples, the final result will not necessarily be immediately identified as being based on cellular automata by a connoisseur of the theme.

6.3.5 Jared Tarbell: *Substrate* (2003)

Iteration of computational processes at different scales might serve as a descriptor of many generative art pieces which are based on the fractal principle of symmetry across scale. Among them, Jared Tarbell's 'living works,' a series of web-based generative pieces published by this New Mexico programmer. Substrate is representative of the philosophy behind his work. It starts as an empty canvas, and then the drawing starts to appear. Black thick lines advance until they hit another line or the edge of the screen. While they do so, they leave a trace of perpendicular semitransparent colored lines with random lengths. Eventually, another black thin line might start with a 90 degree inclination in respect to the main line at a random point. The same rules apply to this.

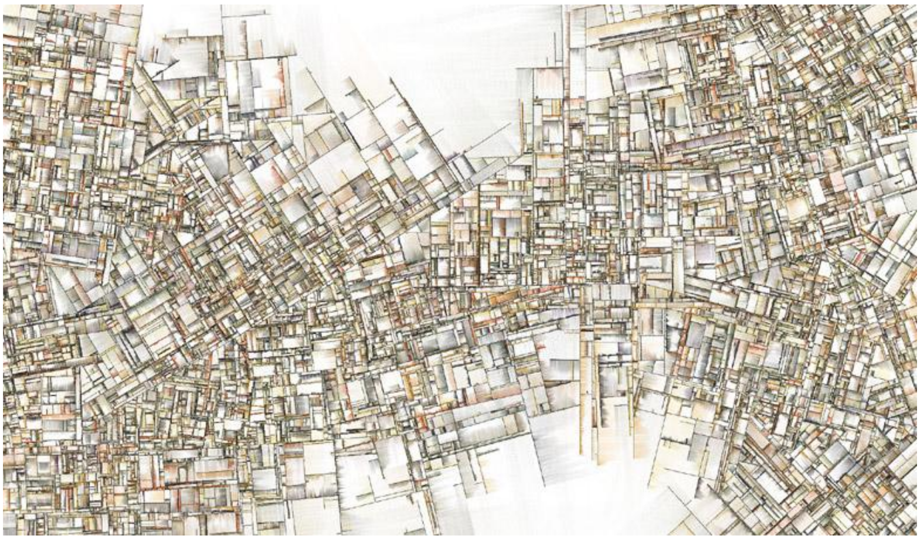


Figure 6.5: Substrate (source: http://www.complexification.net/gallery/machines/substrate/index.php).

What starts off as a series of simple abstract forms being drawn on the screen, slowly complexifies as more and more lines, smaller every time, reiterate the process. The final result is infallibly interpreted as a city map that keeps getting denser and denser up to a point of saturation, were one sees hardly anything but noise. Tarbell shows several screen shots of this and similar works in his website, always taken before the saturation point, in which one can guess the rules which have generated the images. But the ‘living’ of the pieces is the generative process of auto-creation. This is a point where generative art and ALife Art intersect. The work, like others which would also mostly be qualified as generative art, presents a free interpretation of the fractal principle of symmetry across scale, of recursive processes that reiterates in smaller and smaller pixel spaces. And because of this use of fractals they can be studied too from the point of view of ALife Art. Indeed, other works from Tarbell and other generative art creators such as Casey Reas present more organic forms than Substrate. However, the principles utilized are the same.

6.3.6 Leo Nuñez: *Propagaciones; Game of Life* (2007-2008)

Argentine Leo Nuñez is an installation artist with several works related to ALife themes. In some of them he investigates the possibilities of interactivity in celluar automata based work. Both Game of Life and Propagaciones, which in 2007 won the VIDA 10.0 third prize, are such cases.

Game of Life is yet another literal interpretation of cellular automata. As the name indicates, it is a representation of Conway’s proposal. The cells are presented in a table with lights that represent living cells when they are on and dead cells when they are off. A sound indicates a change of state. The installation uses the original Conway’s rules, but it adds the possibility for the visitors to interact with the cells. By passing their hand on top of the cells they can change their states, thus triggering new processes when the system has come to a stop or is stuck in a loop.

Propagaciones is an installation work made out of 50 autonomous robots. Each of these is functionally independent from the others, but does react to the neighboring units. Every robot has two sensors and one effector in the form of a light. The only movement they can do is to spin on their own axis. These robotic cells react on each other and on the proximity of the users. Like in the Game of Life installation, the system has its own dynamics. The patterns of behavior propagate as robots spin and trigger the movements of their neighboring units. In the event of the system falling into a static state, visitors can approach the robots and provoke their reaction, which will in term will propagate into a new chain reaction.

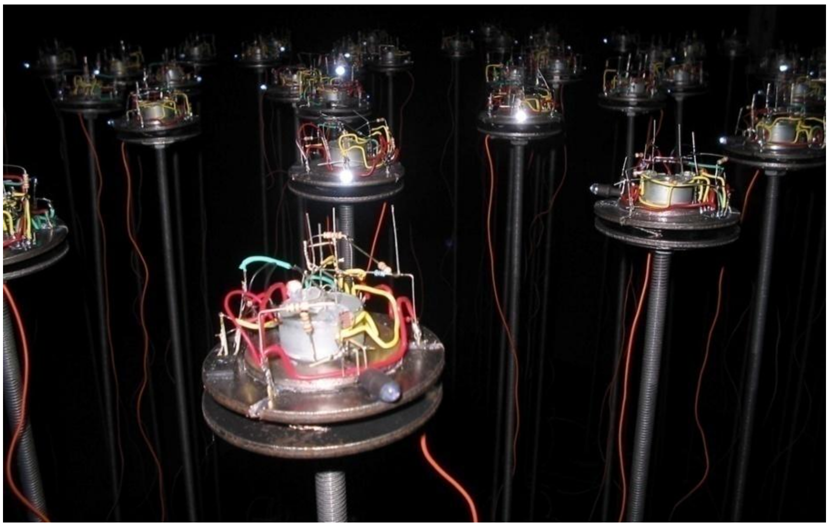


Figure 6.6: Propagaciones (source: http://www.leonunez.com.ar/propagaciones.html).

Like some of the previously discussed examples, Propagaciones utilizes the basic idea of cellular automata; local interactions among local agents trigger patterns of behavior that can be observed at the group level. In this case, the actual patterns are hard to perceive due to how the work was presented (with the visitors being among or by the side of the robots). Unlike with Breed or Sand Lines, the result is not a visually recognizable pattern but a seemingly chaotic dance among the different robots that spin and react to each other in a way that is difficult to perceive as orderly by the visitor of the installation. A different case is Game of Life were, especially if one is familiar to Conway's original formulation, patterns are easier to recognize. In both cases, the systems are actually more reactive than interactive, especially in Propagaciones, since control of what comes out of one's actions is close to impossible. Approaching the robots triggers a reactive response by generating a chain of actions and reactions that disrupt the equilibrium of the system. In this respect, and regardless of the aesthetic interest of the installations, Nuñez's declared objective of investigating the possibilities of interactivity in CA-based systems is arguably not taken too far.

6.3.7 Ralf Baecker: *The Conversation* (2009)

German artist Ralf Baecker developed The Conversation in 2009 as a sculptural work clearly reminiscent of the idea of homeostasis. The system is composed of 99 solenoids placed in circle and all facing inwards. Each of the solenoids is connected to its two neighbors through a wire, and every three units stand between another wire that is connected to three rubber bands that sit in the middle of the sculpture. Once the system has started, its goal is to maintain stability. Each of the magnets works autonomously, and tries to maintain its position while adapting to the forces of the network. The polyphonic sound produced by the solenoids in the piece amplifies the sense of tension of the multiple individual units working together.

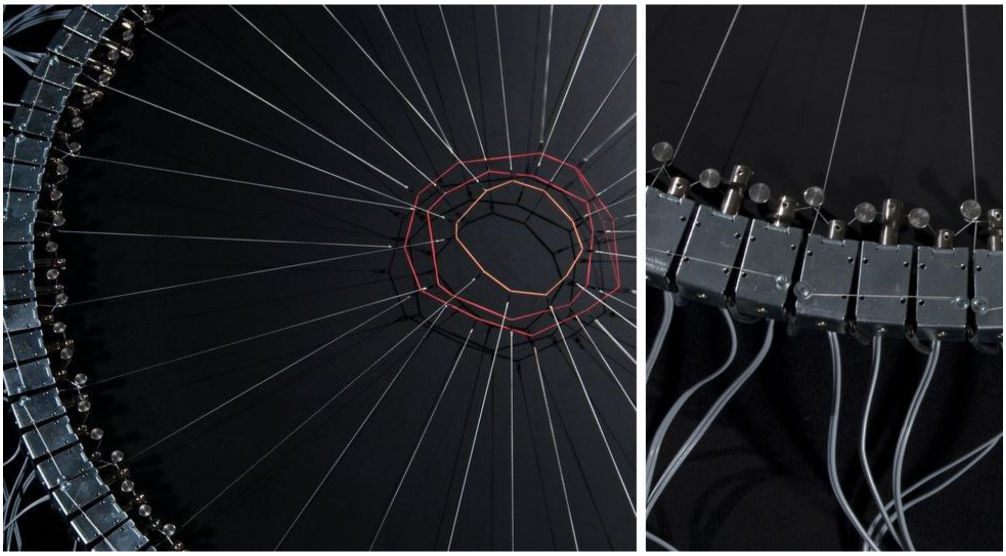


Figure 6.7: The Conversation (source: http://www.rlfbckr.org/work/the_conversation).

The result is a sculptural device which is not designed to respond to anything (or anyone) but to its own dynamics. The author presents it as “pataphysical processing environment” (Baecker, 2009) in an ironic gesture to its unpurposefulness, which is more literally referred to in (Watz, 2010). The Conversation is in this respect reminiscent of Ashby’s original Homeostat, and of Norbert Wiener’s remark on it being a machine without a purpose. The piece is part of a series of works that aim at “deconstruct[ing] the fundamentals of symbolic processes” (Baecker, 2009). It intends to investigate the very fundamentals of computing –of proto-computing– and work up from there in what the author refers to as an emergent process (Whitelaw, 2013), a strategy that fits the primary principles of Artificial Life: simple rules generate complex results that depend on connections among the elements that are often hidden or, at least, not obvious at first.

The conversation works as a sculptural installation and is very successful in connecting with the idea of homeostasis. In terms of emergent behavior, since the magnets work autonomously and there is no central computer controlling the process, it is a clear example of emergence as self-organization, in the same lines as was Ashby’s creation: the individual units work to achieve local equilibrium and, through their interconnected relations, add up to an overall process which is responsible for the equilibrium of the system as a whole.

6.4 Evolutionary Art and Genetic Algorithms

The term ‘evolutionary,’ when applied to art, refers to a design strategy that mimics the process of natural evolution, understood as a blind process of trial and error were new implementations or design iterations are judged a posteriori, instead of planned a priori. Adrian Thomson defines evolution (in the context of design) as consisting of “selection acting repeatedly on heritable variation, where that variation is essentially blind, rather than incorporating detailed heuristics,” and its process as follows: “typically, the object is embedded in some sort of environment, to which it responds, which it influences, or in which it behaves. The evolutionary algorithm designer devises a fitness evaluation procedure that monitors and possibly manipulates the environment and abject, and returns objective function values” (Thompson, 2002).

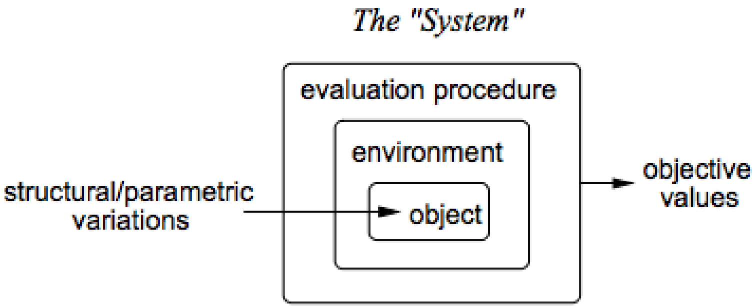


Figure 6.8: A view of an evolutionary algorithm (Thompson, 2002).

Thompson differentiates three types of design problems. According to him, both the traditional and the evolutionary approach are valid options in the first two, while the evolutionary design is the only way to approach the third kind. The premise is that all designable object or system has a set of objective values that we can test in order to know whether or not the design has improved. From this, design can either proceed with an inverse or with a reverse model, and this creates Thompson’s three types of design problems:

- Type A: The inverse model is that which allows the designer to work in advance the sequence of variations that shall bring about the desired improvements (as measured with the objective values).

- Type B: The reverse model implies an iterative approach. One can predict the influence of variations upon the objective values, but not how they should be exactly applied to de designed artifact. Therefore, a trial-and-error procedure is required, with modifications that approximate estimated improvements.

- Type C: Neither of the previous approaches applies. Nor the influence of variations upon he objective values, much less the necessary modifications in order to influence them can be known in advance.

As said above, in types A and B the approach can be either traditional design (and usually that is the case) or evolutionary design. But in type C problems the evolutionary approach, according to Thompson, is the only possibility. Before evolutionary computation, Thompson argues, all design problems had to be simplified or constrained to fit the type A and B problems. But there were some exceptions of blind trial and error design. One such example is the design of golf balls, in which the relevant iterations of the design didn’t have anything to do with anticipation of possible improvements, but discovered by chance as materials were changed for economic reasons or the textures changed due to deterioration by use. See Thompson (2002) for details.

Evolutionary design has been successfully used in fields such as computer science or electronics, and its open-endedness has made it into an attractive form of design from the point of view of creativity as a strategy for AI (and ALife) systems in order to exhibit ‘truly transformational creativity’ (Boden, 2009) or as the basis for a metadesign general strategy (Giaccardi and Fisher, 2008).

In Artificial Life, most of the efforts involving evolutionary strategies have been realized thought genetic algorithms, a technique initially developed by John Holland in the 1970s. A prominent exception is Tom Ray’s Tierra, a work discussed in the first chapter, which develops a complex simulation of an evolutionary system with his own evolutionary rules. These rules, since they simulate DNA-based evolutionary processes, are indeed very similar to Holland’s genetic algorithms. In fact, the code was based on an early computer game called Corewar (Penny, 2009), a programming game about programs competing for computer resources, onto which Ray added the evolutionary strategy for his simulation among other modifications.

Tierra has become an iconic ALife work, since it represents a very solid example of what scientific ALife was about. Ray, a biologist, simulated evolution in order to speed up the possibilities of observing the process. In doing so, he created, from very simple rules, a virtual system in which two species in the form of algorithms cohabited and evolved, while breeding and competing for CPU resources. What was remarkable in the system was that non-initially programmed entities appeared, such as parasites and their hosts, and cycles of aggressive and defensive behaviors. Another remarkable aspect was that Tierra was the first model for punctuated evolution (Boden, 2009). This process, as shown in Tierra, starts with the accumulation of small changes in the genotype of a species without any manifestation on the phenotype. Then, an accumulated change triggers a phenotype manifestation that is the result of all these unseen accumulated changes, resulting in an abrupt change in the species external aspect. According to Boden, until Tierra when a case of punctuated evolution appeared (usually a new biological species appearing in fossil record with no similar fossil ancestor) the only argument for the Darwinians was that the ancestor species must have existed, but never fossilized or was never found. With Tierra, although in an artificial model, the process was first demonstrated to be coherent within the theory of evolution.

Genetic algorithms (GA), the details of which were discussed in chapter 3, are a simulation of the process of evolution that were developed as a computational strategy in order to solve complex open-ended problems, like those that would fall under Thompson's type C design problems. GA have been used with different purposes: experimental simulation, problem optimization or artwork creation among them, often with a diffuse line between one and the other. The range of areas of use is also very wide: evolved image creation (Sims, 1991; Todd and Latham, 1992), evolved animation (Sims, 1994a; 1994b; Ventrella, 1994), evolved radio antennas (Solé, 2009) or the creation of camouflage textures for digital environments (Reynolds, 2011).

Genetic algorithms (GA) are capable of generating emergent phenomena –particularly, of combinatoric emergence[53]. The recombination of elements typical of these systems tend, as intended by design, to create results that escape what was previously specified. The key element for novelty in genetic algorithms (and other similar evolutionary programming techniques) is mutation. In GA, when a new generation is created, the variables that configure the successful individuals of the preceding generation are recombined (e.g. the offspring will get half of the variables of one individual and half of another). But mutation allows for these values to differ from the initially pre-specified thresholds and, therefore, to escape the designer's predispositions. That is, a certain variable may have a random initial value, but this value will always be between a minimum and a maximum that the designer pre-specifies. Mutation, which will only happen every so often, alters a value once it has been set for the offspring and can, therefore, force the value of a variable out of its initial limits.

Evolutionary programming and genetic algorithms have been widely used in ALife Art. The best-known examples are the work of Christa Sommerer and Laureant Mignoeau, Karl Sims and William Latham, but there are many other examples of artists who have used, mainly, genetic algorithms in order to create their artwork. Evolutionary art and its relations to generative art and aesthetics have been reviewed in (McCormack, 2005; 2008; Lewis, 2008).

Philip Galanter identifies three main problems of evolutionary art in (2010). First, the extensive use of genetic algorithms and similar strategies presents some issues regarding aesthetic evaluation. In systems in which humans select the pieces that have to serve as the basis for successive evolutionary steps, problems of exhaustion and the search for novelty over quality appear. And when the judgment is machine made, a whole set of other problematics appear –Galanter discusses this issue in great detail in (Galanter, 2012). The second problem is innovation. Evolutionary artwork, according to Galanter, shares a certain 'cast or sameness' and, beyond the rapidly incremental changes that each work shows, there has not been real innovation within the field in the last years. Finally, evolutionary art has not yet articulated, says Galanter, the theoretical

---

[53] The notion of combinatoric emergence was discussed in chapter 5.

framework that can accommodate it within art theory. Notions such as Whitelaw's 'metacreation' –the idea that the artist shifts role from creating pieces to creating systems that create pieces– are in fact neither new, nor intrinsically digital. In contrast, he has proposed the notion of 'complexism' (Galanter, 2008a, 2008b), with which articulates a framework for a critique of evolutionary art which "rehabilitates formalism as being significant and meaningful, [and] it also reintroduces dynamism and the aesthetics of process and motion" (Galanter, 2010).

6.4.1 Karl Sims: *Genetic images* (1991)

Karl Sims' Genetic Images is one of several projects, developed mostly during the 1990s, that investigate the possibilities of creating computer generated images using genetic algorithms. Among them is the work of John McCormack, Steven Rooke, Stephen Tood and William Latham or Richard Dawkins.

The originality of Sims' proposal was that he did not present the final results of the evolution of the images, nor selected according to his aesthetic criteria the ones that should be used in each successive generation. Instead, he proposed an installation where the fitness criteria are the visitors' choices. Genetic Images presents sixteen images in sixteen monitors equipped with sensors to track the position of the visitors. Standing in front of one of the images is understood as a sign of approval, and used consequently to select the most viewed images in each generation to serve as the 'genetic' basis for the next. The process reiterates, and therefore the resulting images are both a results of Sims' predesigned gene pools and the process of evolution guided by the user's choices.

These resulting images, like those of many similar projects, are indeed prototypical of imagery produced through evolutionary programming. However, Sims's interest in Genetic Images is not only on the result but in the process: the interactivity between the users selecting the images and the system. Whether users know it or not, by standing in front of a particular picture and not of another, they are interacting with the system. Or maybe it would be more precise to say that the system is interacting with them, as the visitors might not necessarily be aware of the processes they are triggering. In any case, ignored pictures are eliminated from the monitors and the new images are made out of combinations from the selected ones.

An important aspect of Sims' approach is that, despite the simplicity of the interactivity proposed, Genetic Images represents a participatory experience. Each image is the result of the intervention of many users that collectively –and perhaps unknowingly– participate in the selection. Thus, albeit in a very different manner than those which are in mainstream use today, Genetic Images is a participatory environment. Each user interacts with it in the same way, but what they find depends on what other users have done with it, on how they have contributed (with choosing some images over others in this case). The idea of participation in relation to interactivity is discussed in chapter 7.

Another feature of this and other systems that use human choices is that it subverts the original genetic algorithms' approach in two ways. First, it slows down and necessarily reduces the number of iterations of a process which was designed as an accelerated decision-making system with the goal of arriving at an optimum result after a large number of iterations. Tom Ray, for instance, designed Tierra precisely because he wanted to accelerate evolution. The second subversion is that the objective of the evolutionary search is the search itself, the process. There's no goal to arrive to. No solution to a problem that the search is aiming at solving.

On presenting his piece, Sims alludes to the concept of metacreation that Whitelaw, Galanter and McCormack would discuss several years later, by questioning the role of the (human) designer in his evolved images: "a designer seems absent in this process, and yet very complex and interesting results can still arise. If enough selections are made by the user and the number of possibilities is large enough, is the user actually being creative, or is the presence of a purposeful designer necessary?" (Sims, 1993). The idea parallels, he admits, Richard Dawkins' main argument in 'The Blind Watchmaker,' published five years before and which represented an enormous contribution to the popularization of genetic algorithms among what would become the Artificial Life Art community. The first editions of the book were released with the software Biomorph, which implemented GA in the growth of simple biologically inspired forms. The software fascinated many of the artists that at the time were beginning to get in touch with computers and also a great number of computer science and computer graphics researchers like Karl Sims.

6.4.2 Karl Sims: Evolved Virtual Creatures (1994)

Three years after Genetic Images, Sims presented Evolved Virtual Creatures; a series of animations created using genetic algorithms to generate animated oddly looking characters which operated in very unconventional ways. Sims set up a series of very simple rules in order to create 3D entities that would be placed in a simulated environment and given tasks to accomplish. The entities consisted of blocks (rectangular prisms) –the nodes, in Sims' terms– connected and articulated at joints –the connections. The number, size and position of the blocks, the placement of the articulations, the control system and all what constituted the entity and determined its possibilities of movement was generated randomly from a specified pool of possibilities (see Sims (1994a; 1994b; 1994c) for details).

On Evolved Virtual Creatures, the entities are placed in the virtual environment in order to pursue the accomplishment of a certain task. As individuals are tested towards these tasks, only the more successful are selected as parents for the next generation of entities. The process, which is in fact textbook genetic algorithms, was applied by Sims in a very successful way, producing very compelling 3D animation sequences. The initial tasks assigned to the entities were swimming, walking, jumping and following a light source (Sims, 1994b). There were also some competitive tasks like being the first to grab an object (Sims, 1994a; 1994c). In the resulting video documentation, the entities are seen crawling, contorting, fighting, etc. in order to achieve their goals. What is most interesting about the work is that none of these behaviors, nor the morphology of the 3D creatures, were actually designed by Sims. What he did was to set up the initial conditions, the environment and the fitness criteria, and then just let the simulation run and hope for the best. The result is probably the best known use of genetic algorithms in animation.

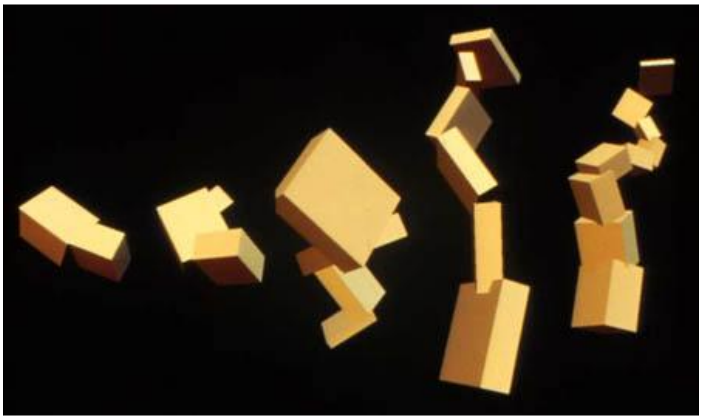


Figure 6.9: Evolved virtual creatures (source: http://www.ugotrade.com/wordpress/wp-content/uploads/2008/02/karlsimspost.jpg).

With Evolved Virtual Creatures, Sims achieved one of the milestones of ALife: to create a very organic feel without organic forms. The constitution of the evolved creatures is at the opposite side of the organic visual simulation: rectangular solid blocks oddly connected among them. No textures, nor anything that simulates limbs or sensory organs. And precisely because of this aspect, and the fact that no organism in particular is simulated, the focus of the work goes exclusively to the seemingly organic movements of the virtual creatures. In this respect the fact that the creatures have been generated through genetic algorithms is, in a way, secondary. The main aesthetic interest of the work is on the surprising reminiscences of human and animal movement –often not so much of walking or swimming as of limping or moving awkwardly– that the blocks movements suggest to the viewer of the resulting animations. From this point of view the work is very successful. Then, Galanter's argument that evolutionary art reached a top that has not been overcome since is probably true here too, as none of the reinterpretations of Sims' Evolved Virtual Creatures seem to have innovated in any significant way.

6.4.3 Christa Sommerer and Laurent Mignonneau: *A-Volve* (1994-97)

A-Volve is an installation that represents a small virtual ecosystem by artists Christa Sommerer and Laurent Mignonneau that has become iconic of ALife Art and the artistic use of genetic algorithms. It was awarded the prestigious digital art award Prix Arts Electronica in 1994 and, remarkably enough for a digital art installation, it has been exhibited, with some interruptions, for at least during eighteen years (Sommerer and Mignonneau 2012).

The installation consists of two modules: one is used to generate the creatures that will inhabit the ecosystem by drawing them in a tactile screen, and the other is the pool where the creatures are shown and interact among each other and with the visitors of the installation. The first way an interactor can participate in the piece is, therefore, to draw a creature. The drawing interface is a screen that is separated from the main part of the piece (the pool with the virtual ecosystem). In it, visitors can draw two-dimensional forms, which are then interpreted by the system and converted to 3D forms, with abilities like speed or aggressiveness dependant on certain features of the form. The 3D creatures will not appear on the drawing screen but on the pool.

This second interface is a large horizontal screen, a virtual pool that the 3D creatures inhabit and were they interact among them and with the visitors. A remarkably distinct feature of A-Volve's interface was that, not only was the pool a tactile interface long before tactile interfaces became popular, but it had a thin layer of water on top of the horizontal screen. A straight-forward and literal interpretation of a pool, reminiscent of other Sommerer and Mignonneau very literal interfaces such as the 'tactile' plants in Interactive Plant Growing (1993), which was nonetheless very effective.

Figure 6.10: A-Volve (Sommerer and Mignonneau, 2009).

In the pool, the creatures interact with each other, and can be affected too by the visitors. The interaction among creatures is based on the predator-prey model and on reproduction. Some of the creatures will hunt on others and kill them if they get to them. Others will pair and produce offspring. It is in this generation of offspring where genetic algorithms appear. The child that appears will inherit some of the characteristics of one of the mates and some of the other, and mutation can occur. In order to interact with the creatures, the visitors have to place their hand over them in the water. By doing so, the interactor provokes that the creatures remain under his or her hands. Then, the interactor can choose to protect the creatures from their predators or have them approach another creature in order to attempt to help them start the mating process.

A-Volve is interesting for a number of things. First, like other interactive artwork in the 1990s, the interfaces were quite an amusement on their own for the audience. In fact, this particular piece retains this characteristic still nowadays, at least in the pool. While the touch-screen where the creatures are drawn is nothing new for tablet or smartphone users, the water-filled table still surprises the visitors. The engagement of the visitors with the piece is, like in other ecosystem-inspired ALife work, constructed in part by the creator/creature relationship. The interactor draws his or her child, which will then be thrown into the pool to try to survive as much as it can and, if possible, to reproduce. Visitors to the installation with enough patience can create several beings and see how they develop within the virtual environment, while affecting this development by interacting directly with them if they so desire. In turn, this creates the context for social interaction among the installation interactors. In this respect: Sommerer and Mignonneau talk about three ways in which interactivity is found in A-Volve; between visitors and creatures, among creatures and among visitors (Sommerer and Mignonneau, 2009).

Figure 6.11: A-Volve (source: http://www.upf.edu/hipertextnet/en/numero-8/a-life_art.html).

A-Volve, according to its authors, is a complex system. Indeed, the creatures that populate the virtual water, and their relation to the others and to the environment is quite complex. Despite the fact that users only need to draw simple 2D forms in order to generate a creature, the interpretation of this form by the system is a process of several steps. First, the flat form is extruded and segmented in 20 parts. From this, the body shape and, more importantly, the body dynamics will be created; depending on the form, a creature will be better or worse at moving around the pool. Once on it, energy is computed so there are strong and weak creatures. The energy is affected by parameters such as stress, and in turn affects who becomes a predator and who becomes a prey, or which creatures are, at a certain point, ready to mate. When they do mate, the complex process of genetic algorithms is applied. Muscle articulations, vision fields and affectation by water pressure add up to the complexity of the virtual creatures.

The problematic with all this complexities is: is this level of complexity (or of complication) relevant to the interactor? And if it is not, is it relevant at all, or is this a case of over-complication? The same question could be posed to another ALife artwork by Sommerer and Mignonneau: Life Writer (2006). This piece is a simpler installation in respect to A-Volve. It presents a typewriter with a fixed projector screen where the paper to write in would be placed. Once a user types on the keyboard, type characters appear as if the machine was writing normally. However, once the carriage return is pushed, the letters transform into insects that start flying and populating the paper. The authors describe how these creatures are created using genetic algorithms: “the creatures are based on genetic algorithms where text is used as the genetic code that determines the behaviour and movements of the creatures. … Here the text functions as genetic code for the creation of artifical life creatures. As in the Life Spacies system the artificial creatures created by the act of typing can be faster or slower depending on their genetic code and body shape. All of the artificial life creatures also need to eat text in order to stay alive and when users type a new text the creatures will quickly try to snap up these characters from the paper in order to get energy. Once creatures have eaten enough text they can also reproduce and have off-spring so eventuelly the screen can become very full when creatures a fed well” (Sommerer and Mignonneau, 2006). What the user sees, however, is that virtual insects populate the virtual paper once the letters are transformed. The space gets crowded quickly and there is no way to discern the differences among the shapes of the creatures or among their behaviors.

Figure 6.12: Life Writer (source: http://www.interface.ufg.ac.at/~christa-laurent/WORKS/IMAGES/LIFE_WRITER_PICTURES/LifeWriter.html).

This is a good example to discuss over-complication. If not even an interactor who is informed on evolutionary programming and genetic algorithms can appreciate how these programming techniques improve the work, when comparing it to what it would be if the virtual insects are simply randomly generated, is it relevant that the authors are using genetic algorithms? Or is it just a layer of complexity which becomes unnecessary? If we evaluate the piece from an interactive art perspective, where interactivity –the relationship between the system and the interactor– is the key for the aesthetic analysis, the answer is no. The use of genetic algorithms is irrelevant and, thus, an unnecessary artifact. If the evaluation is from the ALife point of view, their use is important in terms of design. And, arguably, the use of genetic algorithms is as much an aesthetic resource as any technique can be, regardless of the fact that the viewer of the artwork appreciates it or not. But since the piece is an interactive installation, interactivity must at least be one of the central focuses of analysis.

In the case of A-Volve the same question can be asked, although the case is less clear. Regarding the complexity of the forms and their relation to the virtual environment, it is clear that the visitors of the installation will not appreciate all the subtleties and details of the movement of the forms, but there is indeed a certain feel that some creatures move different from others, and since the mappings are quite straight-forward in terms of aerodynamics, some appreciation of it can indeed be possible. In terms of observing the effects of evolutionary programming, which is impossible in Life Writer, a patient visitor in A-Volve can see how, when offspring appears from the mating of two creatures, this new creature shares some traits with both parents. Therefore, although some details are difficult to appreciate, the complexity is generally justified in terms of interaction. Also, unlike Life Writer, in A-Volve the ecosystem is at least as important as the fact that the system is interactive, and thus the use of sophisticated algorithms and ALife techniques are justified precisely to avoid an overly simplistic virtual ecosystem that is difficult to differentiate from one that is based on random parameters. On a final note on Life Writer, and looking back on the trajectory of the authors, a certain suspicion of redundancy, of an exhausted model, arises. It might not be by chance that this is the last of Sommerer and Mignonneau's genetic algorithm based interactive art pieces.

6.4.4 Ken Rinaldo: *Autopoiesis* (2000)

Ken Rinaldo's autopoiesis is another icon of Artificial Life Art. The work extends a previous one, The Flock (1993), an installation piece which consists of three large articulated robotic arms hanging from the ceiling. The arms are made of grapevine sections, joined with pine and tensed with steel wire. Electric cable, cords and pulleys are placed within the structure, and a head at the top contained the main circuitry and sensors. The Flock's arms respond to proximity (with infrared sensors) and to sound (with microphones). Each arm interacts with the visitors of the installation and with the other arms. The first objective is to avoid collisions with any of them, but each arm also tends to approach sound sources –sounds that, due to the installation setup are likely to be made by visitors– but only until a certain threshold is reached. At that moment, the arms retreat and move away from the source. The arms also have an infrared sensor at the tip, in order to perceive the visitors and avoid colliding with them. A second level of interaction was among the three arms. By emitting telephone touch-tones (easily recognizable by the fellow robots) they communicate the presence and location of a found source of sound to the other arms, so they can also try to approach it. The effect is that the arms' movements appear to the visitors as being coordinated. Also, the telephonic sounds invaded the spaces as if the arms were singing in a private language that the visitors are not able to understand (Whitelaw, 2006: 112).

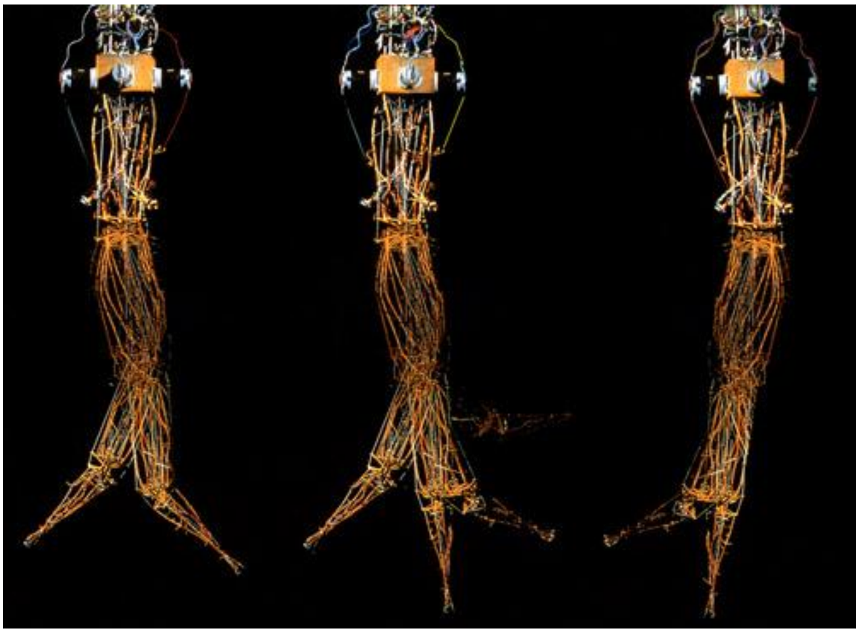


Figure 6.13: The Flock (source: http://www.ylem.org/artists/krinaldo/works/flock/theflock.html).

As said, Autopoiesis expands this idea. First, by augmenting the number of arms from three to fifteen; an augmentation of scale that contributes to amplify significantly the immersive qualities of the work. Two of the arms, apart from the sensors that were already present in The Flock, have also video cameras on their tips, the images of which are projected onto the walls of the installation space, to enhance the visitors sensation of being observed. The self-organization principle of The Flock is present too in Autopoiesis: the arms behave locally but communicate in order to generate the overall behavior. However, in Autopoiesis there's also a central control that communicates to all the arms.

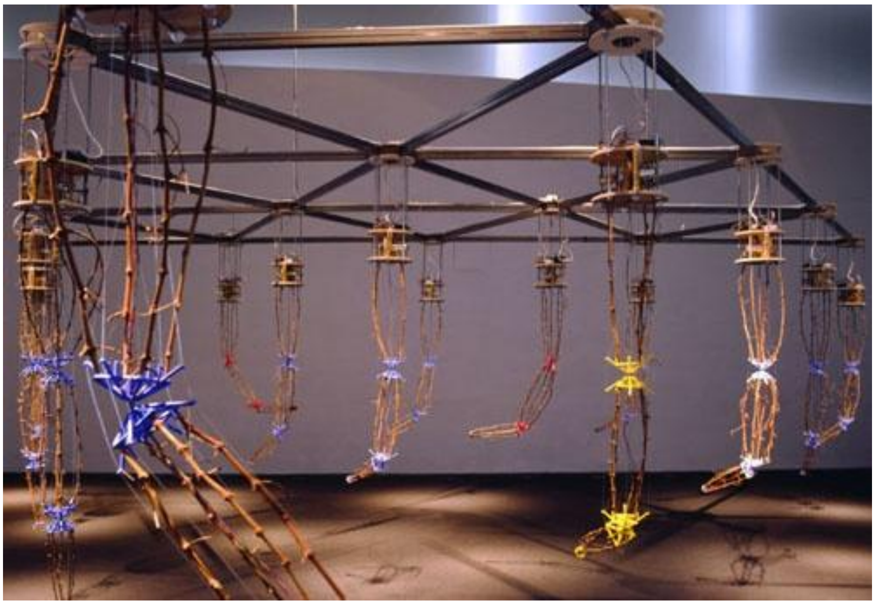


Figure 6.14: Autopoiesis (source: http://www.ylem.org/artists/krinaldo/works/autopoiesis/atp01.html).

But what is most relevant here is that Autopoiesis, according to Rinaldo, incorporates evolutionary programming and exhibits emergence. Rinaldo insists, in the presentation of the piece (Rinaldo, 2000) in that evolution, along with emergence, is the central moto for it; the piece is constantly evolving. However, he does not explain how this evolution happens. In none of the available documentation on the piece is there any explanation of how evolutionary processes are implemented or manifested in it (Rinaldo, 2000; Fundación Telefónica, 2000; Medien Kunst Netz, 2000; Wilson, 2003; Balkjo and Tenhaff, 2008).

In terms of emergence, while The Flock was a case of self-organization, Autopoiesis incorporates a central controller, which in principle would invalidate it as a candidate for self-organization. But again there’s nothing on the documentation that explains what it does exactly. Baljko and Tenhaff (2008) note that “the entire group sends its data to a central state controller for coordination of group behavior,” and Rinaldo (2000) states that “the robotic sensors compare their sensor data through a central-state controller, so the viewer is able to walk through the sculptural installation and have the arms interact both individually and as a group.” In both texts it is stated that the individual behavior supersedes the group behavior, but no further details are given.

As an interactive installation, a group of sculptures that respond and interact with their visitors, Autopoiesis is a success. The lifelines of the piece, sense of immersion, aesthetic value, etc. are all documented in the literature. However, the claims of evolutionary processes and emergence are unclear, at least in the available documentation.

6.4.5 Ruairi Glynn: *Performative Ecologies* (2007-2010)

Performative ecologies is a ‘conversational kinetic environment’ inspired in Gordon Pask’s writings and artistic work, especially the Colloquy of Mobiles (Glynn, 2008a; 2008b). Like Autopoiesis, it is an installation that consists of a group of robotic devices hanged from the ceiling, although their aspect is radically different than Rinaldo’s robots. Performative Ecologies presents simpler and more stream-lined devices, which consist essentially of a long pole with a head on one side, which contains a camera, and a tail on the other; a long light that can rotate on its axis and change its color. The robots ‘dance’ in front of the visitors of the installation. They rotate and move their tails, which can rotate at the same time in a colorful choreography.

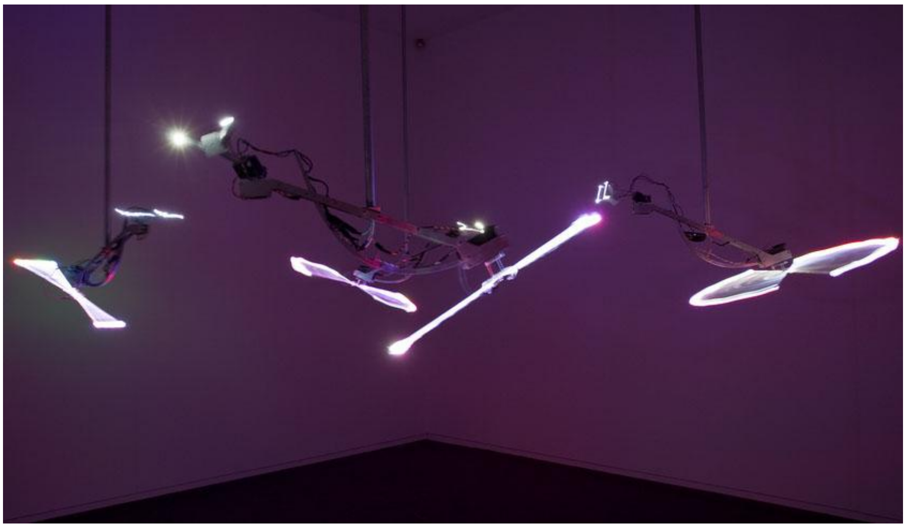


Figure 6.15: Performative Ecologies (source: http://www.ylem.org/artists/krinaldo/works/autopoiesis/atp01.html).

But the choreographies of Performative Ecologies are not predesigned. They start with random movements, to later evolve using genetic algorithms. Each of the robots starts moving randomly as soon as their camera detects that a person is looking at them using face recognition software. Then, the robots will evaluate each of the improvised dances in terms of visitor attention. The more a dance is capable of maintaining a user looking at the robot, the more successful it is. And in order to create successive dances, this successfulness is used as the fitness criteria in order to evolve new dance movements. That is, the dances that have not maintained the gaze of the visitors are discarded, and those which have are used as the basis for the evolution of new movements.

Finally, the piece has yet another feature worth mentioning, reminiscent of the robot-to-robot communication of Autopoiesis. In Performative Ecologies, when the robots find no visitors to dance to, they use the time to teach the successful dances to each other, hence establishing a sort of social communication among them that will, theoretically, improve the overall group’s success in the successive dances, as the ‘database’ of successful movements of each individual robot grows with this communication.

Performative Ecologies is successful in the immediate interaction with the visitors –the robots react to people’s presence and this establishes a first layer of interactivity– and at the same time if offers another example of ‘cumulative interaction,’ a concept introduced in (Soler-Adillon, 2010a), which will be discussed in detail in the next chapter. The basic idea is that, besides the immediate interactivity of the system all visitors’ actions accumulate in configuring the piece, which will keep changing over time according to the accumulated history. This is the case in Performative Ecologies, where through the use of genetic algorithms each interaction with a visitor is accumulated in the installation as a piece that will be used to evolve into the new behaviors of the robots. Arguably, like with many uses of genetic algorithms, combinatoric emergence is likely to occur, too, since the basis of this programming technique is precisely the constant recombinations of the elements with which the systems operate.

6.5 Virtual Worlds and Ecosystems

Virtual worlds and ecosystems have been a recurrent theme in ALife Art since the early days of the discipline. This is a category that overlaps with the preceding (evolutionary programming) and the following (autonomous agents), as often the simulated inhabitants of the ecosystems are programmed using evolutionary techniques, and in one way or another they can be understood as autonomous agents. However, the category is relevant as the metaphor of the virtual ecosystem that is found at this intersection has been, as said, widely used by artists approaching the ALife themes and techniques. The models based on the metaphor of the ecosystem have become increasingly popular too in the field of generative art (Kowaliw, McCormack and Dorin, 2011).

Artificial Life examples of simulated ecosystems include from abstract approaches likes Tom Ray's Tierra to the simulation of insect colonies that aim to reproduce and better understand how colony form and operate (Kluegl et al., 1998). The ecosystem approach implies, in one way or another, an idea of closeness in the sense that the cycles within the simulated environments, no matter how simplified they are, can go full circle. This is the reason why projects such as the Boids, which does simulate a specific group of animals, are considered here not as ecosystem but as a case of a system of autonomous agents.

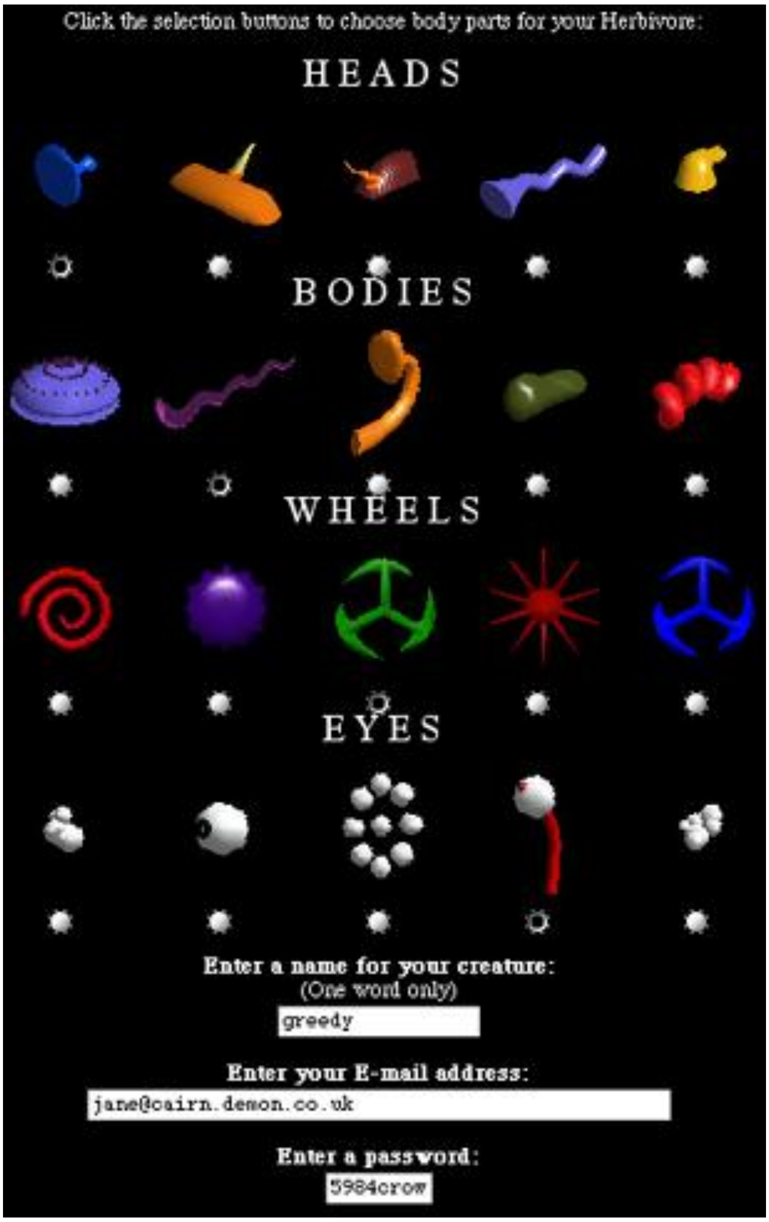
ALife Art ecosystems are typically presented as virtual environments where the users are offered the possibility of creating its inhabitants and/or interact with them. These are the examples where metacreation is not only relevant to the artist's role but also to the itneractor's. As said, often the ecosystemic artwork uses evolutionary programming in one way or another, in order to incorporate evolution to (and hoping for unexpected change in) the system. Four examples are discussed below. Other relevant examples, such as Biota.org's Nerve Garden, are left out of the discussion, because those characteristics that are relevant here are shared with one or more of the pieces discussed. Another important example that could fall into this category is A-Volve, which was discussed above due to its importance, too, as an example of evolutionary programming.

6.5.1 Jane Prophet and Gordon Selley: *TechnoSphere* (1995-2002)

TechnoSphere was an active online virtual ecosystem for 7 years. It was created by British artists Jane Prophet and Gordon Selley along with a group of developers. The first version was released on September, 1995 and was accessible to users through the web. The project consisted of a 3D generated terrain that was populated by virtual ALife creatures. The homepage of the project introduced is as follows: "TechnoSphere is a 3D model world inhabited by artificial lifeforms created by www users. There are thousands of creatures in the world all competing to survive. They eat, fight, mate and create offspring which evolve and adapt to their environment. When you make a creature it will email you to let you know what is has been getting up to in this world. Using the creature tools you can find out how your creature is surviving, what it is doing at any time, and where it is in the terrain” (Prophet and Selley, 1995).

The terrain of TechnoSphere is divers: it has rivers, forests or deserted areas that creatures inhabit. The main goal of the creatures is survival and a secondary goal is to procreate. Creatures in TechnoSphere are created as herbivores or carnivores, and if they have enough energy on a particular moment they can mate and create offspring, which will be generated using evolutionary programming: the child will inherit part of both of the parent’s characteristics. Creatures can be killed by other creatures or die of old age. Initially, 30.000 creatures were randomly created so that the virtual 16 square kilometers of TechnoSphere weren’t inhabited, but the major strength of the project was user participation. Over the seven years it was online, TechnoSphere had over 100.000 users which created over a million creatures. Typically, 20.000 were active at the same time (Prophet and Selley, 1995).

When users visited the TechnoSphere website they were invited to design a creature. First, they would chose to generate either a carnivore or an herbivore, and then they would choose from a series of predesigned features to be combined into the new virtual animal. An important part of the process was that they would be asked to give a name to their creation (the system would also assign them an ID to avoid duplicates) and their email address, so the system could keep them up to date about the endeavors of their virtual child.



Click the selection buttons to choose body parts for your Herbivore:

Enter a name for your creature:
(One word only)
greedy

Enter your E-mail address:
jane@cairn.demon.co.uk

Enter a password:
5984crow

Figure 6.16: TechnoSphere's selection menu for creature creation (source: http://www.mindatplay.co.uk/images/technospehereheadsbodies.jpg).

Once the TechnoSphere's new inhabitant was created, it would be released into the virtual world. From this point on, the creator of the creature would not directly interact with it anymore. Like a parent of a child who has left from home, he or she would only get news from it every once in a while. One way to know about one's creation was to ask the system for a postcard. This resulted in receiving a 2D rendering of the creature in a scene showing the environment where it was in the moment that the image was created. Another way was though the update emails that the system generated for the users, which were sent whenever the creatures fought, mated, generated offspring or died.

There are several examples of the communications that TechnoSphere sent to its users posted online. Some are long details of the encounters of the creatures with others (mainly fights and matings). In the project's website there one of such examples:

***************TechnoSphere Telegram ***************
These are the highlights of The Banker's (ID 815178) life in TechnoSphere recently:
The Banker mated with I HATE POKeMON (ID 793404) but conception didn't take place.
The Banker mated with wee craig (ID 818519) but conception didn't take place.
The Banker successfully defended itself from an attack by Bittron (ID 808695) who lived to fight another day.
The Banker attacked Shiba (ID 812379) but it was too strong and got away.
The Banker mated with destroy 2 (ID 793975) but conception didn't take place.
The Banker attacked destroy 2 (ID 793975) but it was too strong and got away.
******************** End Transmission *********************

Figure 6.17: ATechnoSphere creature and its offspring (source: http://dwork.com/DWLPages/TechnoSphere.htm).

Further developments on the system led to TechnoSphere III, which incorporated a real-time 3D visualization of the environment. However, this version was presented in a different context, a museum installation, which forced the authors to change the 'narrative' of TechnoSphere (Prophet, 2001). In this real-time version, the experience was accelerated. First, the creatures were introduced in the world as "adolescents with raging hormones" instead of children in order to accelerate the experience. If in the web a user could be getting update emails for several days, in the museum visitors needed to experience the whole range of possibilities of the project in a much shorter time. Thus, in a matter of minutes, creatures would play, mate and die.

TechnoSphere is a prototypical example of an online interactive ecosystem. In it, users acquire a god-like role of creators, and then relate to then are expected to engage with their creations emotionally, as those autonomous virtual creatures develop into the environment until they eventually perish. Like in other examples such as A-Volve, the authors present it as offering various levels of interaction. First of all, there’s the interaction of the users with the system (the interface used to design the creatures), and then among the creatures within the environment. Authors also offered the possibility for users to reach each other after their creatures had interacted. Finally, Prophet mentions the interaction between users and developers as an important part of the process, too (Prophet, 1996). TechnoSphere II, released in 1996, incorporated new features based on user suggestions.

Besides its value as an ALife project, Technosphere is also a good example of the division between the real and the virtual that was present in the discourses of the 1990s decade; a divide which in recent years has blurred in a move towards ideas such as ubiquity (Penny, 2013b). The project existed as a 3D environment, but the early web technologies of the time didn’t afford for a real-time visualization of the environment. So the TechnoSphere world was virtual in every sense. There was no visualization possible of one’s creature’s developments. The users in the real world would release virtual creatures which existed independently. The metaphor of the telegram is illuminating in this respect. It represents the news that arrive from a far-away land, a one-way asynchronic communication message that doesn’t expect to be responded, which stresses the separation of the two worlds.

6.5.2 Troy Innocent: *Iconica* (1998)

Despite the bizarre look of the creatures and its humorous approach to portraying both mating and fighting in the 3D version (Prophet, 2001), TechnoSphere is a rather literal approach to the ecosystem. In contrast, Iconica, offers a completely abstract ALife based ecosystem. In consonance to Troy Innocent’s interest in the “explorations and deformations of the native aesthetic language of computer culture” (Whitelaw, 2006: 86), Iconica uses icons instead of natural representations. What the user sees and can interact with is not the naturalistic-looking result of some computation but, rather, an iconographic representation of this very computation.

Figure 6.18: Iconica’s entitiy library (source: http://iconica.org/iconica/entitylib/entityLib.html).

Iconica was presented to the public as a two screen installation. In one of the screens users can interact with the creatures that inhabit the virtual world. On the other, a map is projected representing the whole system. The entities that populate Iconica are made out of the combinations of a series of simple entities: elements, forms, entities (formed by bodies, moods and behaviors) and spaces, each of them with six related categories, combined to form a complex world that is represented by non realistic humanoid and animal-like figures. The user can interact with those with yet another layer of complexity; the language that is originated from the icons that constitute the system.

Figure 6.19: The constituting elements of the entities in Iconica (source: http://iconica.org/iconica/language/ic3.htm).

In order to interact with the entities, then, the user has to learn the language of the icons, which can be perceived as being very confusing at the beginning. Other interfaces offer alternative possibilities for viewing and interacting with the world, including the creation of new forms. And the system has its own rules for evolution and mutation of the creatures. Thus, in some respects Iconica is very similar to other ecosystemic ALife Art work, such as A-Volve or TechnoSphere. Users create creatures, interact with them and observe the virtual world as it evolves. The difference, as said above, is that Innocent was interested in using ALife within a larger personal discourse on language and semiotics. His main concern were the symbols that constituted the language with which the creatures communicated, and which in turn represented their constituting code. ALife was, in a way, accessory to this discourse. However, the result was rich both in his main intentions of a language experiment and as a ALife Art system. According to Whitelaw, the piece "achieves a quite remarkable richness. While it users the standard devices of artificial genetics, mutation, and reproduction, it allows for a wide range fo complex interaction and behavior, culminating in its iconic language" (Whitelaw, 2006: 85-86).

6.5.3 Joan Soler-Adillon: *Digital Babylon* (2005)

Digital Babylon is an installation that presents the user with a virtual 2D real-time generated ecosystem on a projected screen. As opposed to the previously discussed examples, here the ecosystem is visible in its totality at all times. Also, here the visitor does not create individual creatures, he or she can only interact with some of them. Graphically, it is a move towards simplicity, hence displacing the weight of the work towards the observation of behavior. The ecosystem consists of a black background populated by one species represented by triangles and another one by blue tailed ellipses. Green plants appear in the form of a series of dots generated by a ‘random walk’[54].

The system works autonomously, without the intervention of the interactor. It contains two animal species (from here on S1 and S2) and one vegetal. The two animal species relate as predator (S2) and prey (S1). They both reproduce using genetic algorithms, and therefore are evolutive, although the dynamics of the process are different in each one. On starting the system, there are a few randomly generated individuals of S1 and of S2 and some plants. From here on, the ecosystem is left to find its equilibrium, which might eventually be disrupted by the human visitors.

S1 creatures have two states: they are either looking for food or looking to mate. They are drawn as small triangles and their color changes according to their state: if they are looking for food the color is red when there is food around and orange when not. They are yellow if they want to mate and purple if they are dead. The creatures have a certain level of energy that increases as they eat and decreases as they move. Like in A-Volve or TechnoSphere, if the creatures reach a certain high level of energy they will try to mate. If they succeed, offspring is created using genetic algorithms, so the child inherits some characteristics of either parent, and mutation will occur eventually.

S2 creatures prey on S1. They move in group (in a sort of slow flocking) and will attack the preys if they get close enough. If the group gains enough energy, all the S2 that are hunting go back to the lower part of the screen where they have a nest. There, genetic algorithms are applied to create new eggs, from where new groups will appear in the future. Only the individuals that come back the strongest stay to nest. The rest goes out to die. Finally, plants appear wherever an S1 or S2 individual dies, as if the dead body fertilized the ground. Each plant has a series of dots which is equal to a fixed amount of energy. Once a dot is eaten once, it disappears.

---

[54] The random walk is a computational process, typically a programming exercise, were an element in a particular position on the screen moves at simple regular steps always either up, down, left or right at random.

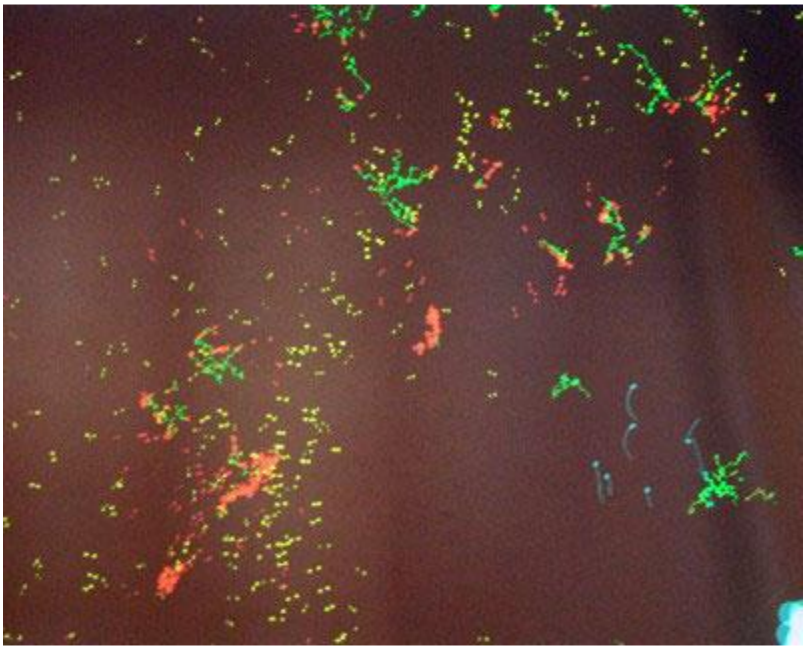


Figure 6.20: A view of Digital Babylon’s ecosystem.

The overall dynamics of equilibrium in the system provoke cycles of abundance and scarcity of aliment for S1, resulting in high and low peaks of population. The cycles generate as follows: when there is abundance of plants, individuals obtain energy and mate often, so the population increases rapidly. But this augments the consumption of plants, which therefore become scarce and eventually inexistent. Then, the weakest individuals start to die massively of starvation, thus decreasing the population. As a result, several plants appear in a short time, and the cycle starts again. Meanwhile, S2 is at the same time affected. When there’s a lot of S1 population running around either looking for food or to mate, hunting is easier, so S2 birthing cycles accelerate. When there’s a low peak of S1 population, or they are all preoccupied in eating, hunting becomes harder, and so does survival and reproduction.

In front of the projection screen where the ecosystem is displayed, there’s a designated area for the interactors. In stepping onto it, a small mark will appear on the screen, as a direct mapping of the visitor’s position in the ecosystem. From this point on, the S1 creatures will have a certain tendency to move towards the position of the visitor. Depending on how their individual digital DNA is encoded, each creature’s behavior will range from completely ignoring the visitor to going straight to his or her mapped position. Most of the creatures will approach the vicinity of the user, with a stronger o or weaker pull towards him or her.

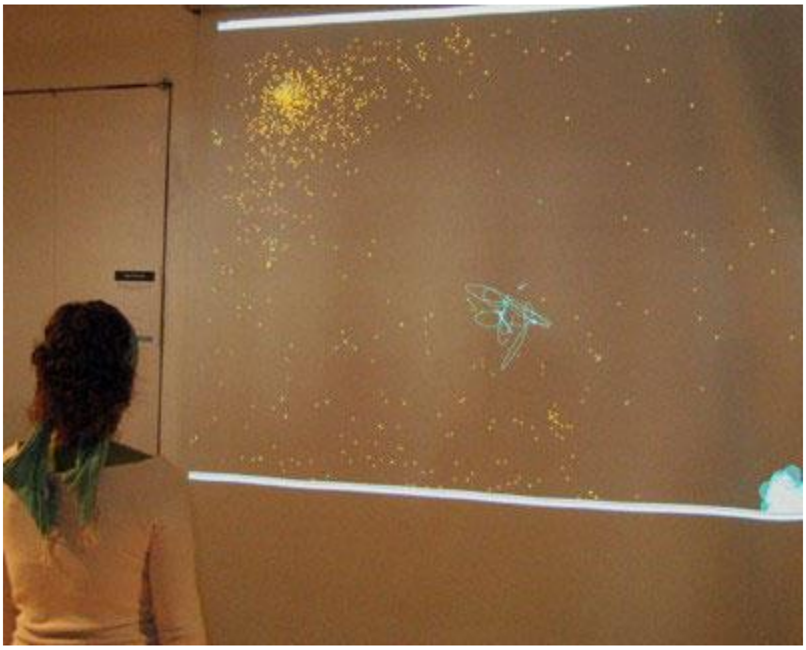


Figure 6.21: A visitor of Digital Babylon.

This is all the direct interactivity that a visitor finds in the installation. However, this interactivity can affect the system as a whole and, more importantly, the successive interactions of other visitors. Since the S1 creatures that approach the user are those that, within the spectrum of behaviors, have a stronger tendency to it, the visitor may choose to protect these friendly creatures from their predators by sanding far from where the group of hunters is. By doing so, the S1 creatures which will have a bigger chance of being hunted –and therefore die and not generate offspring– are those which ignore the user. Consequently, new generations will be friendlier, as they inherit their parents' DNA. Thus, a persistency of this behavior by the users will drive the evolution of S1 towards friendliness, and successive visitors will find creatures with a stronger pull to go towards them –in a word, with a stronger tendency to be interactive. If, on the contrary, users tend to go close to the predators and get the friendly S1 creatures killed, the species as a whole will tend to ignore more and more the user. In this case the S1 creatures, and therefore the piece, will actually become less interactive.

The fact that in Digital Babylon one can see the whole system at all times make it easier to see its global dynamics. This is valid for the cycles mentioned above, but also to appreciate the evolution. In A-Volve creatures may generate offspring, and with careful observation one can see that the child resembles the parent, but it would take a very successful family line to appreciate improvements (such as grandchildren being smarter at avoiding predators than their antecessors). In TechnoSphere this would be impossible. Mostly, this is due to the differences in the individual specimens. The fact that behaviors are already divers make it difficult to appreciate improvements at the level of the species. In Digital Babylon all the individuals of S1 look the same and behave similarly. So when after some hours with the system in motion the individuals have become generally faster, or more difficult to get haunted, or stay together much more than initially, like it happened in some of the simulations, the overall effect can easily be appreciated.

Cummulative interaction can also be attributed to this piece. In fact, it is a general characteristic of genetic algorithm based interactive works. However, in this installation the effects of the accumulation directly affect the possibilities of interactivity for successive interactors. It is not only that a visitor will find the system to be different depending on what previous visitors have done with it, but that the system will actually be more or less responsive, i.e. more or less prone to interactivity, depending on how the responsive individuals within the system (the S1 creatures) have been treated by previous users.

6.6 Autonomous Agents

Autonomous agents is a broad descriptor that encompasses both the physically instantiated works of ALife related practices and the virtual autonomous agents which can't be defined in the terms of the previous categories. It includes both isolated agents and systems of agents. A prominent example of the latter are Craig Reynolds' Boids: each of the simulated birds or fish moves autonomously, according solely to local rules, individually, although the result is that the observer recognizes in the system a simulation of group behavior. Some approaches to collective robotics also fall into this classification. Some have similarities to the ecosystem approach, since the metaphor used to model the system is the behavior of a particular group of animals. Such a case is the simulation of an ant colony in (Deneubourg et al., 1990). As said above, however, the ecosystem approach implies an idea of closeness that these projects and others that deal with groups of robots (Beckers, Holland and Deneubourg, 2000) are not necessarily interested in.

When individually considered, autonomous agents are exactly what cybernetic devices were about: agents, robots, systems, etc. that can operate within their environment without the aid of an external guiding entity. Walter's tortoises or Pask's Musicolour are early examples of such autonomous systems. In art, there have been numerous instantiations of this idea, with very different outcomes, some of which are reviewed below. Regarding ALife in general, this is arguably the field in which the discipline has had a bigger success, both in research and, latter, commercially.

Prominent among this is the work of Luc Steels and especially of Rodney Brooks, whose influence has been very significant in robotics research from the late 1980s onwards. Brooks developed an approach to robotics that was very critical of 'classic' Artificial Intelligence and that was very much in consonance with Artificial Life related ideas such as flocking or emergence, which resulted in the development, among other things, of 'swarm' and 'multi-agent' robotics (Penny, 2009). In most cases, the emergence that can be attributed to these systems can be identified with self-organization, although the generation of novel and non previously specified behaviors is often sought, too.

Brooks' critique of AI is drawn from an antireductionist position –or, at least, from a 'soft' antireductionist point of view, as explained below. He considers that the initial objectives of the field, replicating human intelligence in machines, was misguided from the beginning, and is the origin of the frustration that arose within the discipline in its first decades; a frustration that resulted in slow progress and hyper-specialization (Brooks, 1999). The initial objective was abandoned in favor of the resolution of very concrete problems. This failed approach is what Brooks labeled as 'top-down'.

In contraposition, he proposed a 'bottom-up' approach that is generally a general strategy of ALife, as Langton had been proclaiming since the founding years of the field. In Brooks' understanding, intelligence cannot be, with the current knowledge of the time, correctly decomposed into sub-problems, as AI had aimed at doing from the beginning. Hence the previous mention at a ‘soft’ antireductionism. Brooks does not say that intelligence is in principle too complex, but that it is so in respect to the current knowledge, and that his approach is precisely a way to start advancing towards a solution to that problem: “human level intelligence is too complex and little understood to be correctly decomposed into the right subpieces at the moment and that even if we knew the subpieces we still wouldn't know the right interfaces between them. Furthermore, we will never understand how to decompose human level intelligence until we've had a lot of practice with simpler level intelligences” (Brooks, 1999).

In any case, the result of the argument is the same: the AI approach has not been successful because it could never be. Instead, he proposes a bottom-up approach that is modeled after simple but effective intelligences. The basic idea coincides fully with Grey Walter’s approach in building his robot tortoises, as discussed in chapter 4: since the bigger system is too complex to model –it’s an exceedingly complex system–, start with simpler devices that perform simple actions, but that are capable of autonomously performing their required tasks.

In order to do so, according to Brooks, intelligence needed to be redefined and abandon the central notion upon which AI had been so unsuccessful with: the notion of representation. That is, robots should not necessarily be built with a ‘brain’ that mapped their environment, i.e. that made a representation of it in order to operate in it. Instead, it was sufficient that robots had enough sensors and effectors in order to move around successfully, regardless of whether or not they had fully mapped their environment. Again, Walter’s robots come to mind: they had a very simple circuitry, and by no means it was intended that they had a representation of their environment in the sense of elaborating a map or simulating how a human inspects his or her surroundings before moving around. But the result was very successful in terms of the pre-specified tasks. This is the context for Brooks’ famous claim that modeling the environment (the world) is unnecessary, and the world itself is what shall be used: “when we examine very simple level intelligence we find that explicit representations and models of the world simply get in the way. It turns out to be better to use the world as its own model” (Brooks, 1991).

There are other examples of successful bottom-up robotics, and some of them have combined this approach with evolutionary programming. A remarkable example in terms of emergence as generation of novelty is the case of Italian researcher Dario Floreanos’ robots. Within a small squared space, a series of robotic agents with sensors, motors, a light source and a neural network as a brain, were placed to move around with the objective of maintaining their energy. In the space there were also two cylindrical objects. One of them charged the robots and the other discharged them. The robots could sense the lights that neighboring robots emitted. With this setup, Floreano and colleagues expected some group behavior to appear, and as described in (Solé, 2009) this is what happened. At first, Robots self-organized in a cooperative manner. On finding light sources they would emit light signals that would become messages to other agents, which then in turn would approach the source and recharge. Also, some messages emitted by the discharging unit could be read by the observers as warning signals among the robots.

However, in the long run this cooperative strategy turned out to be harmful individually, because the energy source would quickly become overcrowded. And here is where emergent novelty, in the form of the appearance of a behavior that was not contemplated in the observational frame of the researchers, appeared: eventually, some robots learned how to lie in order to free the space for their own recharge. One instance of this was a robot whom, after being blocked out of the recharge unit, would signal as it would have when neighboring the harmful unit, provoking the nearby robots to move away –and therefore free the recharging space or, at least, not overcrowd it any more. The second instance was a robot whom, next to the discharge station, would signal as if it were next to the charging one, so nearby agents would go to the wrong place, only to move then itself to the recharging unit.

These are two cases of combinatoric emergence. The robots didn't acquire new sensors or develop new artificial neurons to expand their processing possibilities. What happens is that the elements that were existent within the system recombined in a manner that allowed for newly observed behaviors to appear. Consistently with the emergence-relative-to-a-model theory, once this behaviors are observed they can be incorporated into the observational frame (the model) and they cease to be emergent in that sense. In parallel, the overall cooperative or competitive behaviors of the robots are instances of self-organization emergence.

6.6.1 Simon Penny: *Petit Mal* (1989-2005)

Petit Mal is a robotic interactive installation inspired in the small epileptic seizures known by the same name. These seizures consist in very short-term (often only a few seconds long) and apparently harmless sudden losses of consciousness. Typically the subject suffering them interrupts whatever activity he or she is doing for a few moments, to resume it later as if no interruption had taken place. Simon Penny's robot behaves in a way that resembles these lapses of consciousness. It is a radically reactive machine with essentially no memory. Built intentionally with very low processor power and a few low-tech sensors, Petit Mal was designed to live in the threshold between reliability and mal-function or, rather, it is a machine that "malfunctions reliably" (Penny, 2000).

The appearance of Petit Mal avoids anthropomorphism or zoomorphism. The author's intention is that the artifact be not disguised, but shown as crudely as possible as the machine it is. Similarly, the behavior doesn't imitate anything in particular. The intention was not modeling intelligence in any sense, but to "build a device that gave the impression of being sentient, while employing the absolute minimum of mechanical hardware, sensors, code and computational power" (Penny, 2000). The idea was to present the viewer with this sentient machine as an interactive actor contextualized, physically and culturally, in the real world: "Petit Mal is an attempt to explore the aesthetic of machine behavior and interactive behavior in a real world setting." It was left to the viewers to project their interpretations to the device, whom, as Penny notes, would draw from their previous experience with babies, animals and moving artifacts in order to project to Petit Mal behaviors much more complex than those that were actually attributable to the robot in terms of its actual construction and code.

Figure 6.22: Petit Mal (source: http://simonpenny.net/works/petitmal.html) ©Simon Penny, 1993.

The actual behavioral rules of Petit Mal are simple: when it perceives that something is approaching it, it moves away. But when that something stands still, then it is the robot who initiates the approach, only up to a certain distance before it moves away again. It will also spin on its axis to face the moving target and wonder around when it is left alone. The sensors on its 'head' allow it to perceive the interactors, while the two big bicycle-like wheels it possesses, along with the penduling body and counter-weights allow it to move in a rather strange manner. From this simplicity, the overall behavior and the 'personality' of the robot build up.

The approach resonates with those of the cybernetic artifacts discussed in chapter 4, especially with Walter's tortoises. However, while Walter was using simple elements to explicitly model a much complex system (the brain) Penny is interested in what the artifact is on itself, not in anything that it is modeled after. Petit Mal is an experiment on interactivity as a cultural practice, on the agency of the artifact within a cultural context, "an actor in social space" (Penny, 2000).

The piece was not explicitly framed within the Artificial Life Art context. Rather, it is the discourses on ALife Art that have claimed that Petit Mal should be a part of the physically instantiated systems, the robotic works that are an important part of this artistic movement. In fact, Petit Mal explicitly avoids some of the basic characteristic that constitutes ALife Art. First, it does not model any animal or human behavior. What it does, it does because of a careful design process and choice of minimal hardware, which was done in order to achieve a certain navigation and interaction with the humans. It doesn't try to be intelligent, either. It is as dumb and short-termed as it was intended to be. Second, as said, it does not try to look like anything but a robot. The only explicit objective of the artist which can be linked to ALife is that Petit Mal aims to have, among other things, 'personality'. The design approach mentioned above is undertaken "in order to induce the maximum of personality" to the device.

Figure 6.23: An interactor with Petit Mal (source: http://simonpenny.net/works/petitmal.html) ©Simon Penny, 2006.

The result is a robot that very successfully engages with the visitors, either individually or in groups. Its movements are reminiscent of curiosity and playfulness. The awkwardness of its looks combined with its seemingly inconsistent movements help creating an image of a sort of good ugly character, a dance partner in the interaction space. The robot is, in this respect, successful in becoming an agentic entity.

6.6.2 Simon Penny: *Sympathetic Sentience* (1995-1996)

Sympathetic Sentience in an interactive sound installation that consists of a group of twelve robots that communicate to one another sequentially. Like with Petit Mal, in this piece Simon Penny investigates the possibilities of creating complex behavior out of simple system mechanics and low-tech hardware. Each of the robots is a relatively simple electronic device, capable of producing one chirp each minute with a particular rhythm. Situated in the walls of a darkened room, besides producing the sounds each unit passes along to the next an infrared signal indicating its rhythm. Then, this receiving unit combines its own rhythm with the received information, and passes the new rhythmic pattern along. As the process advances, the complexity of the rhythms increases as they cycle around the group. In Sympathetic Sentience II, which used programmable microchips instead of logic hardware circuitry, changes in pitch were also involved in the process.

The system is self-organizing or, in Penny's words, self-governing (Penny, 2000). The sound pattern that the visitors to the installation perceive is neither predesigned, nor directed by any one of the robots or an external entity. As the robots communicate locally –each one to the next in circle– the overall process is formed. In this respect, it is an example of emergence as self-organization.

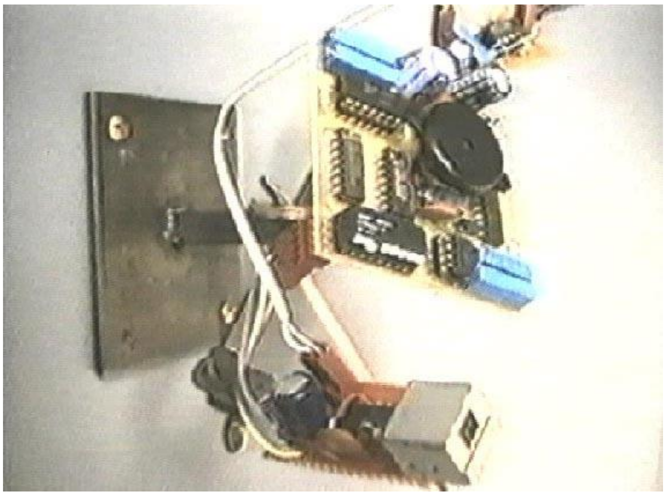


Figure 6.24: A Sympathetic Sentience unit (source: http://www.fundacion.telefonica.com/es/at/vida/vida10/paginas/v2/penny.html).

Interactivity is included in the piece in an unconventional way. After an initial build-up period, the system works on its own as rhythmic patterns continually evolve. It never becomes fully saturated, nor does it become fully silent if it is not interfered with. However, this equilibrium can be disrupted, often unintentionally, by the presence of the installation visitors whom, by moving around the space, interrupt the infrared transmissions of the units. If these interruptions are short in time, lapses of silence will infiltrate the rhythm patterns of Sympathetic Sentience. If the interruption is long, the whole system can be forced into complete silence. If this happens, once the interruption of transmissions ends, the build-up process will have to start all over again (Penny, 2008).

More so than interactivity, emergence is an important idea in the conceptualization of the piece. The goal of the installation is to generate "complex patterns of rhythmic sound through the phenomenon of 'emergent complexity'. Sympathetic Sentience is an attempt to build a physically real model of emergent complex behavior amongst independent units" (Penny, 2000). As said above, if we accept the premise that identifies self-organization as one kind of emergence, this system's sonic patterns are indeed emergent, since they result solely from local interactions and are not predesigned. However, in presenting the work Penny refers too to emergence as related to (epistemological) novelty, in a formulation that resonates with emergence-relative-to-a-model: "whether [Sympathetic Sentience's] behavior is deemed to be 'emergent' is a matter of previous experience. Most visitors find it reminiscent of the sound of communities of frogs, crickets or cicadas. But to at least one rather dry observer, it was simply a chaotic system of a certain numerical order. To another it was a demonstration of one model of neural propagation. Here emergence would seem to be 'in the eye of the beholder'" (Penny, 2000).

In fact, Peter Cariani, refers to this piece when commenting on the possibilities of creative emergence in art. However, he does not explicitly state that it is in fact an example of "art that creates autonomous objects that themselves independently evolve new primitives." It seems, rather, that in any case the association would be a better fit to combinatoric emergence, as Sympathetic Sentience's units do not actually create the primitives or evolve new sensoric devices, but they do constantly instigate the search spaces and recombine these primitives to generate novelty. This doesn't impede Cariani to associate the installation with a model of the brain activity that is indeed capable of creative novelty: "for me Sympathetic Sentience is highly reminiscent of how I think about neuronal pattern resonances in the brain, as a set of unfolding interactions between sets of neuronal assemblies that are emitting the own annotative signals. A triggering sensory event thus generates a process of semantic elaboration and eventual convergence to a stable set of pattern resonances that becomes the final 'meaning' of the event. The signals continue to reverberate and to regenerate themselves within the network in a manner that parallels the autopoietic model of consciousness" (Cariani, 2009a).

6.6.3 Erwin Driessens and Maria Verstappen: *Tickle robots* (1993-2006)

The tickle robots from Driessens and Verstappen are a series of robotic pieces that deal with physically instantiated behavioral devices and their relation with the human body. In particular, as the name suggests, a relation that is driven by pleasure, in what in principle would seem to be a very innocent form: tickling. The robots could be described either are radically interactive or as not interactive at all, in the sense that the user here does not control them in any way. In fact, here the visitor of the installation is expected to lay face down as in a massage parlor, and let the robots act on their body.

The first tickle robot, Spear (1993), consists of a spear of grass suspended from a metallic chain which is controlled by a solenoid. This electromechanical device “produces a chaotic oscillation, resulting in unpredictable, jerky motion patterns” (Driessens and Verstapen, 2006), which is exactly what the authors were looking for in order to simulate tickling. As Verstappen points out, “an important aspect of good tickling is that it has to be unpredictable” (Verstappen, cited in: Whitelaw, 2006: 131). In fact, although the image might suggest otherwise, the movements of spear aren’t soft and slow phased. The solenoid produces fast and oddly paced jumps of the spear of grass, but the softness of the tip of the device compensates this and resulted in the pursued effect.

Ticke (1996-2006) is a more sophisticated robotic device. The size of about a pack of cigarettes, this robot has two rubber tracks at each end which spin independently so the device can move and turn at will. Its objective is to move on top of the body and not fall, for which it has a dual axis tilt sensor that will avoid the fall. It crawls over the body where it is placed by generating paths at random, as an automated massaging device. In its last version the piece incorporated a microcontroller that enhanced its moving possibilities. It has sensors that allow it to perceive when it is getting close too close to an edge, thus informing it to move away and avoid falling from the user’s body.

Figure 6.25: Spear, Tickle and Tickle Salon (source: http://notnot.home.xs4all.nl/ticklerobots/spear/spear.html).

Finally, Ticke Salon (2002) is yet another reiteration of this idea. This time, a ceiling mounted device holds a soft white brush that is suspended with wire, close to the massage bed below it. When someone lies on the bed, the robot will begin exploring the space and will encounter the body. At this point, it will perceive a loss of tension in the wires, and here a mapping of the body begins. As the robot moves, it generates an image of the visitor's body is shown, as it grows, in a monitor by the side of the bed. There's an interesting collision here between the robot's practical objective of moving around the body with the user's pleasurable, probably in some cases erotic, sensation of the body (all its parts) being visited by the stimulating device. The difference between Spear and Tickle Salon' setups infer a very different experience. The first is a simple device that quirks and moves histerically, in a close analogy to what a five year-old child would understand by tickling. The second moves towards erotism, as the device now moves slow and softly, covering all the skin it encounters. Finally, the goal of mapping the body, which is only present in this third and last example, changes the experience from the random movements to the exhaustive exploration "by means of touch" (Driessens and Verstapen, 2006).

6.6.4 Narcís Parés, Roc Parés and Perry Hoberman: *El Ball del Fanalet / Lighpools* (1998)

El Ball del Fanalet is an interactive installation that takes its name from, and is partially inspired by, a dance that is typical of Catalan culture. In it, the dancing partners hold a candle lighted fanalet (a paper lantern) as they dance. Eventually, the paper-made fanalets will burn out if the dancers are clumsy or just not careful enough. The couple that holds the fanalet the longest is praised at the end as having won the game. In the interactive and augmented version, the setup of the installation is a rounded floor were up to four users can walk in holding a device that incorporates location sensors and that works as the paper lantern. Once inside the space, each lantern will illumine the floor right below it. This is the first step in the discovery of quite a complex artificial life system.

Figure 6.26: Users interacting with El Ball del Fanalet / Lightpools (source: http://www.dtic.upf.edu/~gvirtual/lghtpls/lpinst_c.htm).

The first reaction of the system is the appearance of proto-objects: colored tetrahedrons that suddenly appear under the light of a lantern and that will fade away if they are not illuminated by the user that has the lantern with the matching color. If the connection does happen the proto-object will metamorphose into an articulated being, and grow if the light remains above it. Now the object is under the user's control, who shall learn how to make the creature grow only up to a certain point: "if the user fails to decrease the intensity of the light (by raising his or her fanalet), the object grows until it bursts, scattering a new crop of colored protos onto the floor. Other users can then nurture these protos, and the cycle starts again" (Parés and Parés, 2001). This is the first and simplest possible ways of interacting with the system. But the dance here has not yet begun.

If, instead of making it burst, the user stays with the creature, he or she can teach it how to dance in quite an elaborated way: "any sufficiently rapid sequence of movement of the fanalet is interpreted as a start of this training, which makes the partner learn every new movement of the fanalet. The partner follows and memorizes this movement until the fanalet remains still, and then starts to repeat it continuously on its own (until it is taught a different sequence). Slower movements of the fanalet are interpreted as leading movements, which the trained partner follows as it dances around the lightpool. By alternating rapid and slow movements of the fanalet, users can teach, dance, and interact with their partner" (Parés and Parés, 2001).

At this point, in a behavioral addition that resonates with Gordon Pask's work on the idea of interactive system's boredom due to overstimulation (e.g. the Musicolour), the virtual dance partners can get bored if they are trained for too long and leave the human user behind. However, before this happens two users who have trained dance partners can join their fanalets in an invitation for the two virtual partners to dance together, which they will do ignoring the human partners again. All of these phases, with the subtleties and details they have, generate an interactive system with a remarkable degree of complexity in terms of the chain of actions and reactions that develop as the users interact with the environment.

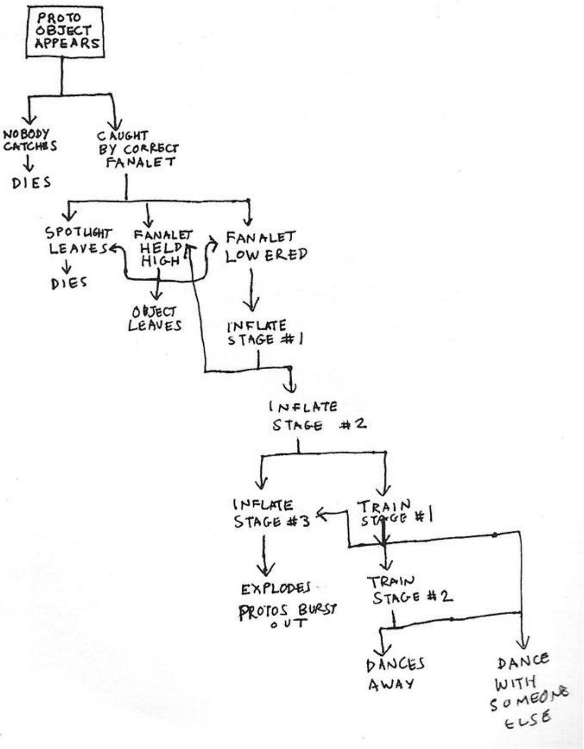


Figure 6.27: A diagram of the scripted interaction in El Ball del Fanalet (Viveros-Fauné, 2001).

According to the authors, this work is a step towards research in open-ended interactivity. In a moment were interactivity was dominated by applications created for single-users being physically confined to keyboards and mouses, the authors were "trying to create the conditions under which an interactive, immersive image can be experienced as a casual, open, social space." In fact, El Ball del Fanalet combines this openness (the dances that users improvise with the virtual partners) and the social aspect (interactions among users) with what could be labeled as scripted interaction. Different successive phases of interactive possibilities are clearly separated as the experience advances: discovery, growth, dance and final conclusion. Within each phase, the possibilities of interactivity are open to exploration. Despite the fact that El Ball del Fanalet is indeed an elaborate piece aesthetically, the authors' goals are on interactivity, and in fact used it as an example of their proposal to an interaction-driven approach to design (Parés and Parés, 2001).

In this respect, the fact that the piece is related to Artificial Life is secondary (and, in fact, never claimed by the authors as far as the documentation goes). But because of the theme of the piece, the mutating creatures and their complex and adaptive behavior, it is safe to state that the piece fits into the ALife Art category, as noted for instance in (Penny, 2011). Because of the piece's complexity and atypical approach, it is perhaps the most difficult piece to fit into the proposed taxonomy. Up to some degree, and while not presented as such, El Ball del Fanalet could arguably be considered to be a simulated ecology. In fact, what is found below the users, and illuminated by the fanalets, is a hidden world to be discovered. However, the user interacts directly with the individuals and has no means of perceiving the 'world' as a whole. What he or she relates to in the piece are autonomous agents: a series of abstract mutant creatures that, if treated properly, will come out of their virtual confinement and live and dance with their visitors until they get bored, or dance with another virtual creature, before fading away for good.

6.6.5 Karolina Sobecka and Jim George: *Sniff* (2009)

Sniff is a screen-based interactive computer generated dog; "a behavioral portrait of a lively pup" (Penny, 2010). It is shown at the scale of a real dog and senses the visitors to the installation through computer vision. With this, it establishes a playful interaction with them by following and responding to their gestures, while changing its behavior based on its state of engagement with them (Sobecka, 2009). The work is conceived to be placed in a storefront-like setting, where people pass by. The virtual dog is programmed to try to attract people's attention and, from this point on, engage in an action-reaction playful loop, along the lines of the simulation of a real dog's reaction when an unknown person approaches it. Sniff will stare at the visitor, jump towards or away from him or her reacting to approaching or backing movements, or stand on two legs in order to reach up to his or her hand.

Figure 6.28: Sniff interacting with a visitor (source: http://www.gravitytrap.com/artwork/sniff).

Visually, the 3D dog avoids realism by presenting a non textured animated 3D model, placed in a black environment. Therefore, nothing but the dog exists in the screen, and the dog itself is a simplified polygonal model. Once again, the focus of the viewers and the interest of the piece is not on the visual representation but on the behavioral processes that develops between the virtual pup and the visitors: “the contrast between the verisimilitude of Sniffs behavior and the abstraction of its visual representation heightens the persuasiveness of its physiological and behavioral modeling. Sniff is a rather subtle dog-portrait, it will identify with one of several visitors a primary interactor, and has a sophisticated behavioral memory” (Penny, 2010).

6.7 Concluding Remarks

The range and topics covered by Artificial Life Art are divers, and the selection presented here does note claim to cover them all. However, it does cover the most relevant themes and techniques, both in relation to interactivity and to emergence, in the two formulations that are developed in the thesis. The next two chapters are dedicated to analyzing these two aspects, interactivity and emergence, which are present in some of the discussed works.

In general terms, these works present an effort to model, in one way or another, some basic aspect of life or of a life-like organization. In general, this is done with a special interest in the creation of life-like behaviors, and not necessarily through aesthetic representation that resembles organic forms in any way.

The modeling efforts here are to be understood in a very broad sense. Differently from the scientific modeling, which aims at simulating some natural phenomena in order to understand it, the artistic ALife Art modeling is freed from these constraints and seeks to engage aesthetically either with the user through interaction, or with the piece itself, usually as a complex system –or a system with complex behaviors– that arises from relatively simple basic rules.

7. ARTIFICIAL LIFE ART AND INTERACTIVITY

The concept of interactivity is used in diverse contexts and often with different meanings and intentions. Like with the case of emergence, this is done, in some occasions, without making explicit a clear definition of the term in order to support the discussion. This chapter proposes a definition of interactivity and an aesthetic framework in order to contextualize the discussion on interactive art and emergence and the analysis of interactive ALife artworks. The notion of interactivity is presented and discussed in relation to ALife Art. It is important to note that interactive ALife Art is only a subset of interactive art and that not all ALife Art is interactive, as discussed below. Many pieces do not aim at engaging with the user according to what here will be described as interactivity. This chapter assumes that the reader is familiar with the artwork examples presented in Chapter 6.

7.1 Historical Notes on Computer Interaction

Historically, interaction as it is understood now in the context of computation developed in parallel to the advances of computational media in the 1960s and onwards. Before that decade, the available computers were incapable of performing in real-time; that is, of responding fast enough to the received commands for the user to perceive the responses as immediate. Some of the ideas that would constitute the basis of modern computation had been developing during the preceding decades. Alan Turing's universal machine (Turing, 1936) or Vannevar Bush's Memex (Bush, 1945) were particularly influential in pre-configuring the technological developments to follow some years later (Gere, 2008).

Turing anticipated the idea of a machine capable of performing a potentially infinite number of tasks: he envisioned a sophisticated programmable typewriting machine with infinite tape, a clear intellectual antecessor of the modern computer. Almost a decade later, in 1945, Bush wrote *As We May Think*, an article in which he imagined a desk with access to a great number of high definition microfilm files with controls that would allow the researchers to establish connections (links) between pages in these files. This desk, which he called the Memex, would be connected to others with an equivalent functionality. With this thought experiment, Bush anticipated the multimedia computer and the Internet. His concern when imagining the Memex as the technology that we might have eventually was exactly the same that was behind Tim Berners-Lee's invention of the actual web: to help scientists share their work, know the work of others (so research would not happen in parallel) and establish connections between their respective documents and those of others.

In 1960, J. C. R. Licklider published *Man-Computer Symbiosis*, a paper in which he presented the idea of the computer and the human complementing each other in real-time (Licklider, 1960). His research was made in the context of the US Defense funded Semi-Automated Ground Environment (SAGE), which fostered the use of computers as a tool for defense. Also within this context, Ivan Sutherland created Sketchpad, a remarkable application that had an immense influence on latter technological developments (Gere, 2008: 68-69). With a pointing device that SAGE operators were using to obtain information from planes in radar screens, Sutherland created a system with which users could write directly on the screen, creating both 2D and 3D figures. Once created, these figures could be instantiated multiple times, in what made Sketchpad into the first Object Oriented Programming application. It also represented the inauguration of Computer Aided Design.

However, whilst Sketchpad made way for the development of commercial and academic Computer Generated Images, it didn't have an immediate influence in what would become personal computing. It was another defense related agency, the Advance Research Projects Agency (ARPA), where these developments would take place by the hand of Douglas Engelbart in the Stanford Research Institute (SRI). Following

Vannevar Bush and Licklider, Engelbart’s driving idea was that computers could help human creativity and thinking. He used the concept of ‘augmentation of human intellect’ to present the work that he was doing (Engelbart, 2001).

During the 1960s, the researchers at the SRI were strongly funded while enjoying an important degree of freedom. This allowed for a situation in which, actually, the researchers were not creating computers to be used specifically in Defense related tasks but, rather, creating the basis for the future general-purpose personal computer. In fact, by the end of the decade, Engelbart and his colleagues at the SFI had put together a very advanced system, capable of text processing, hyperlink implementation, computer networking and with the mouse as a pointing device among other things. However, a sudden Defense funding cut in 1970, which was in part a response to the Vietnam War protests (Gere, 2008: 71), led to the disintegration of the SRI. At that moment, many of the institute’s key components would move to Xerox Corporation, to fund the Xerox Palo Alto Research Center, and they took many of the ideas that had been developing under the leadership of Engelbert with them (Almirón, 2001).

The Xerox PARC would become a key player in the history of personal computing. There the Xerox Alto was developed as the first fully functional desktop computer. It would never be commercialized, but many of what it implemented became the embryo for latter projects such as its commercially failed successor, the Xerox Star, or the Apple Lisa and Apple Macintosh, which would bring to the commercial arena the developments which can be neatly traced back to Engelbart’s work on the SRI.

It was at that moment, when in the mid 1980s and early 1990s personal computing became available to the general public, when a significant number of artists started to use computers as a creative tool and that digital art gained traction as a distinct movement. Before that, the use of computers had been limited, and often combined with analog systems. As seen in chapter 4, early Cybernetics pioneered the experimentation with programmable devices, although these were often not computer based.

Animation artists were among those experimenting with computers. The brothers John and James Whitney also had build an analog mechanical custom-made animation system in the late 1950s, and a key-frame system was developed in the 1970s at the National Film Board of Canada. In parallel, computer researchers were creating computer-generated images as they experimented with, and enhanced the possibilities of, image creation with computers. Ed Catmull in University of Utah or Charles Csuri in Ohio State University and some researchers of the Bell Labs researchers were among those responsible for the most relevant advances in this area.

In terms of interactive art, there were some remarkable milestones, such as the Cybernetic Serendipity and the Software exhibitions that were discussed in chapter 4. Another important project was the organization of the Experiments in Art and Technology (E.A.T.) in 1966 by Billy Klüver with the collaboration of Fred Waldhauer, Robert Rauschenberg and Robert Whitman among others. E.A.T. was “a first instance of the complex collaboration between artists, engineers, programmers, researchers, and scientists that would become characteristic of digital art” (Paul, 2003). It consisted mostly of technologically mediated performances, and some of them involved interactive systems. One such example was Open Score, a performance by Robert Rauschenberg that was presented in 1966. In it, during the course of a nighttime tennis game the lights illuminating the court were turned off each time the players hit the ball, up to the point where there was almost complete darkness. Then the audience would see themselves projected in the walls as infrared cameras were filming them (Daniel Langlois Foundation, 2014).

However, as technologically complex as it was to create such a performance in the mid 1960s, the interaction that can be attributed to this system is simple: the mapping between the ball hitting the racquets and the lights turning off. Indeed, the interactivity that the computer systems offered at the time was minimal and merely functional when seen from a contemporary point of view. The focus was on making advances on the capabilities of computation (image creation, text processing, networking, etc.) and user control, the type of interactivity that was emerging at the time, was only a part of it. As said above, it wasn’t until personal computers became available to the general public that artists could begin to experiment with interactive art besides the scarce occasions in which, before that, possibilities arose for such experiments.

7.2 A Definition of Interactivity

Interaction is a very broad concept that refers to any set of actions performed by at least two agents that are mutually influential. The term is useful in many areas, from the description of how chemical components react to one another to how several sociological groups relate in a particular social context. According to Jensen (1998), interactivity is a derivative of interaction that usually means ‘exchange’, ‘interplay’ or ‘mutual influence’. The concepts are often used interchangeably. Strictly speaking, however, interaction is a temporal process, while interactivity is the capacity to partake in that process.

As said, interaction and interactivity are used in diverse contexts, often with different meanings in each of them. The use of these terms is in many occasions not accompanied by a clear definition of what is meant by them and, even when definitions are given, they can be contradictory if compared (Downes and McMillan, 2000). It is in this respect that Jens F. Jensen labeled interactivity as a multi-discursive concept (Jensen, 1998), following Tim O’Sullivan and colleagues’ definition, which states that this kind of concepts are, precisely, those the meaning of which changes depending on the context they are used (O’Sullivan et al., 1994).

Thus, before delving into the considerations of what one understands as interaction or interactivity, it is necessary to define the scope of the approach. If the context is that of interactive art, it is useful to dedicate some effort in clarifying what is understood by interactive system, user or interactor, interface and other related concepts. This chapter is dedicated to this.

The first condition to describe interactivity as it will be understood here –and from this point onwards– is that it involves one of these two following cases: it is either an interaction where at least one of the agents is an artificial system or an interaction that is mediated by an artificial system. Here artificial system means simply not biological, i.e. a human made system (generally a computer). Therefore, interactivity as understood here is the domain of disciplines such as Human Computer Interaction (HCI) or Interaction Design. That is, the main focus of interest is in the relation between (artificially conceived) systems that are interactive and their potential human users (the interactors).

There is a body of literature that has addressed the definition of interaction and interactivity in relation to both general and computer media, mostly since the mid-late 1990s. The general consensus is that there are three main approaches to defining the concept from the media studies point of view, depending on whether the defining efforts situate the focus of the definition on the structure of media, on the user, or on the process of communication that develops between them (Downes and McMillan, 2000; Kiousis, 2002; Quiring, 2009; Mechant, 2012; Weber, Behr and DeMartino, 2014).

One group of authors is that which focuses on interactivity as a characteristic of media (Durlak, 1987; Baecker and Buxton, 1988; Wilson, 1994; Jensen, 1998; 1999; 2008; Sundar, 2004; Lee, Park and Jin, 2006). The center of interest within these approaches is on how media is structured in order to afford interactivity: what it offers to the potential user and how, so that interaction can take place. Interactivity is often presented in these cases as a continuum, and different media are related to different points in that continuum (from less interactive to more interactive media). The types of activities that each of these offers to users is read within these parameters, too. Jensen's diagrams and 3D tables in (1998) are quite a literal illustration of the complexity of this approach (see figure 7.1). On the other end, some authors place the focus of study on the user, mainly around the concept of perceived interaction (Wu 1999; Liu and Shrum, 2002; Leiner and Quiring, 2008; Quiring 2009). Here what is most important is what interactivity means for the person that uses the interactive system and how he or she relates to media in its terms. Interactivity here is an information-based process that is relevant in terms of individual perception (Newhagen, 2004). Finally, there is a third group of authors that center their approach on the communicational process that interactivity represents (Kiousis, 2002; Rafaeli, 1988; Rafaeli and Sundweeks, 1997; Crawford, 2003; Noble 2009; Penny 2000; 2011; Green, 2010). The focus in these cases is on the process that develops as the interactive system and the user act and react to one another. The definition that is proposed here fits into this third category.

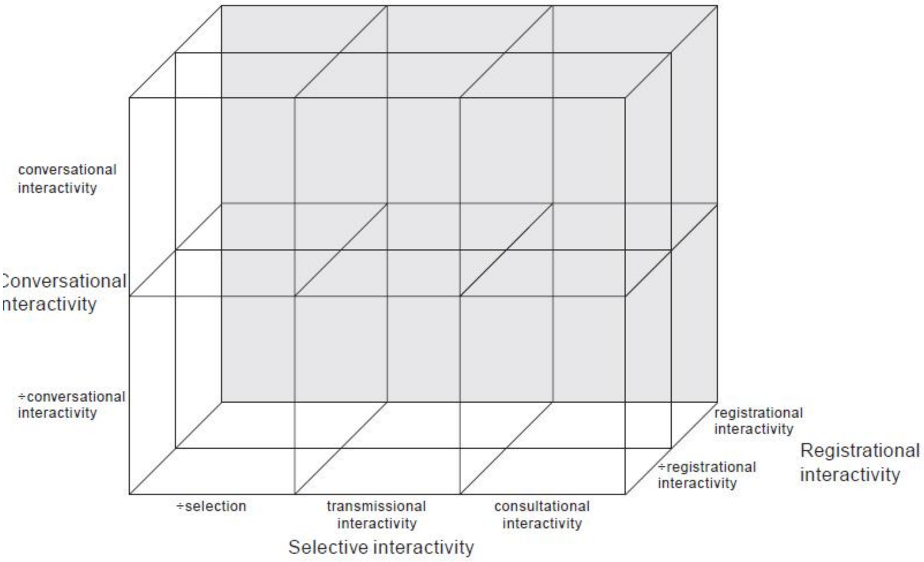


Figure 7.1: The 'Cube of Interactivity': a 3-dimensional Representation of the Dimensions of Interactivity, in (Jensen, 1998).

Despite the mentioned efforts, and mainly because of the variety of approaches, there is no canonical formula to define interactivity, even within the media-related disciplines. Some definitions strive for simplicity, like game designer Chris Crawford's, who defines interactivity as "a cyclic process in which two active agents alternately (and metaphorically) listen, think, and speak" (Crawford, 2003: 76). Others try to cover what simplicity leaves out, resulting in awkwardly complex descriptions of the term. Such a case is Kiousis' definition in his 'concept explication' of interactivity: "Interactivity can be defined as the degree to which a communication technology can create a mediated environment in which participants can communicate (one-to-one, one-to-many, and many-to-many), both synchronously and asynchronously, and participate in reciprocal message exchanges (third-order dependency). With regard to human users, it additionally refers to their ability to perceive the experience as a simulation of interpersonal communication and increase their awareness of telepresence" (Kiousis, 2002).

Within these efforts, the metaphor of the conversation is useful: the dialogue between the man and the machine that inspired Cybernetics and that is the foundation of HCI. After all, the idea is that computers exchange information with users, as the back and forth of messages advances. User does, computer reacts, then user does again in response to this reaction, and so on and so forth. This conversation can be understood in a very simplistic manner, in terms of control of systems where the artificial system responds to the commands of a human user, or with a more ambitious interpretation of the metaphor, which aims at reproducing, through the interactive system, a series of meaningful exchanges.

Gordon Pask's Conversation Theory is an important effort to theorize this (Pask, 1975; 1976; 1996), although it is a general theory not specifically centered in man/machine conversation. Pask's writings on the subject find its context in his interest in learning devices. Conversation is here presented –and interpreted by latter authors– as an advanced form of interaction by systems that have the ability to learn. More recently, Paul Pangaro has elaborated on Conversation Theory in relation to social media and design (Pangaro, 2009; Dubberly and Pangaro, 2009) and to relevant models on interactivity (Dubberly, Pangaro and Haque, 2009). Following Pask, Pangaro and colleagues understand that conversing systems are a sophisticated assemblage of (second-order cybernetic) learning systems that feed on and react to each other. This is, according to them, the most elaborated form of interaction.

Here the systemic approach is paramount. Pangaro and colleagues are, like the cybernetitians were, theorizing about the idea of system, which is yet another pervasive concept in the literature, especially since the development of Systems Theory in parallel to Cybernetics and Information Theory. Input, output, feedback, and other Cybernetic concepts are crucial to understand systems. They can be defined as "a set of things – people, cells, molecules or whatever– interconnected in such a way that they produce their own pattern of behavior over time" (Meadows, 2009) or, more simply, "complexes of elements standing in interaction" (Bertalanffy, 1969).

In this respect, before delving into the discussion on interactivity and esthetics, it is relevant to state the importance of properly defining the subject of study at each time we want to assess whether or not something is interactive. That is, it is important that the actual system under analysis is specified; one must state what the level of analysis is (e.g. a particular piece of software, and entire computer system, a network, etc.). Since in general terms all systems can be placed within larger systems, and at the same time they consist of aggregates of subsystem, this annotation of terms can be important. In this respect, it is worth noting that the user and the computer can be understood here as forming a system: the two agents (one human, one artificial) are components that stand in interaction through an interface.

Here it is useful to recall John Haugeland's distinction among components, systems and interfaces, discussed earlier in the Cybernetics chapter. Haugeland elaborates this distinction in the context of the discussion with Herbert Simon on the principle under which systems are decomposable into subsystems. He presents these three notions as the essential pieces to understand this issue: within a system, a component is a subsystem that is replaceable by a functional equivalent. A transistor in a computer or a teacher in the school system would be examples of that. An interface is the point of contact among components within a system.

Everything that happens inside the component is irrelevant from the system's point of view, as is whatever happens outside the components that does not affect either them or their interfaces. These three notions, system, component and interface, "should all be understood together and in terms of one another. A 'component' is a relatively independent and self-contained portion of a system in the sense that it relevantly interacts with other components only through interfaces between them (and contains no internal interfaces at the same level). An 'interface' is a point of interactive 'contact' between components such that the relevant interactions are well-defined, reliable, and relatively simple. A 'system' is a relatively independent and self-contained composite of components interacting at interfaces" (Haugeland, 1998).

Haugeland is here proposing to understand systems in a way that has two important implications. First, it is strongly dependant on what one decides to focus on and on the point of view taken. The same set of objects or beings can be seen as forming a component, or a system with interfaces among each one depending on whether one is looking at one kind of activities of others. Second, it is necessarily an oversimplification when it is used to model the relation between a human (or an animal) and his or her environment. He notes that the human interaction with the world (which implies the nervous system, immune system, etc.) is far too complex to be accounted for by these parameters. The concept of interface between the brain and the nervous system, for instance, doesn't really work: there's too much complexity and too much bandwidth involved in it. Nonetheless, and accounting for these limitations, the system-component-interface framework is still useful to discuss interactivity.

With this framework in mind, the system under analysis can be specified. And once this has been done, assuming it is a system that relates to a human user or interactor, the question is whether or not the behavior it exhibits affords interactivity or not. In terms of establishing whether or not a system is interactive, the first step is to discern if the responses that the system under analysis produces towards a potential user are interactive or merely reactive. Along these lines, Dubberly, Panagaro and Haque in (2009) differentiate among static, dynamic-reactive and dynamic-interactive systems. The first are systems that “cannot act and thus have little or no meaningful effect on their environment (a chair, for example).” Dynamic systems do have the ability to act and relate to their environment. Within those, they distinguish between systems that react (they have a linear relation between system activation and response) and systems that interact. The categorization of interactive systems is further elaborated and, as said above, they present conversation as the most sophisticated form of interaction (Dubberly, Pangaro and Haque, 2009).

Therefore, a reactive system is one in which the system responses are always the same when the same controls are activated regardless of previous actions by the user. E.g. a light switch connected to a bulb will always light up the bulb if it’s off, and turn it off if it’s on. If, instead, the responses are processed and given according to previous actions and context, and are thus not always the same responses to the same control activations, neither are they random, then the system is interactive. Previous context and processed responses are, thus, key to understand interactivity.

Previous context means taking into account of the previously occurred events. As said, this is key to understand the difference between reactive and interactive systems. This idea was proposed by Sheizaf Rafaeli in his definition of interactivity as “an expression of the extent that in a given series of communication exchanges, any third (or later) transmission (or message) is related to the degree to which previous exchanges referred to even earlier transmissions" (Rafaeli, 1988). The idea is that an interactive system does not simply respond to the controls activated by the user according to a (linear) one-control-one-response mapping. Instead, the response is processed before and varies according to the particular context of the action and the feedback received. In cybernetic terms this means that this system’s behavior is designed as either an open or a close loop that takes the feedback of the user into account.

In addition, besides this taking into account the particular context of each interaction (i.e. what has happened up until that point in relation to the user) the system’s response will not depend on linear one-to-one mappings but on system behavior: on how the system has been designed and programmed to respond at every given time. This means different things for different systems: a computer will respond to a keystroke depending on what program the user has activated last and where the cursor is at that particular moment; a button push on a videogame controller will result in the avatar shooting a gun or throwing a knife depending on what is the last weapon that the gamer has chosen; and artistic interactive installation will show a particular frame of video or another depending on whether the visitor is either approaching or moving away from the screen.

Thus, as noted by Haugeland, the perspective from which the system is being analyzed is here very relevant. For instance, in a videogame like Tetris, from a certain point of view the same controls will always produce the same results: a piece turns around or move sideways when certain buttons are pushed. There are no surprises there. But what piece will be moved each time, whether or not it actually moves, where it goes exactly and how fast is moving downwards depends on the context of the particular gameplay. If the piece is all the way to the right, pushing the right button will not move the piece at all, and it will not turn around if it doesn't have the space to do so. Thus, from the perspective of the controller/game within the context of a particular play, the interaction is not particularly sophisticated, but neither can Tetris be considered to be a merely reactive system.

Finally, and with this general context in mind, interactivity is defined here as follows:

A series of related actions between two agents where (1) at least one of them is an artificial system that (2) processes its responses according to a behavior specified by design and (3) takes into account some of the previous actions executed by both agents.

It is important, here, to analyze interactivity in terms of systems design and functionality, not of subjective experience by the user or interactor. That is, the point of view for the analysis is on how the systems are designed in order respond to the users' actions in a way that is read as interactive. Thus, in the terms of the three approaches defined at the beginning of the chapter, this proposal fits into the third category: that which focuses mainly on interactivity as a process. This doesn't mean to neglect the user's perception or the design of the interfaces. As said above, the system here is formed by a user interacting with a technological artifact, so all the components are relevant in this respect. However, while other approaches understand interactivity either as a property of the technology or as a characteristic of psychological perception, here the approach aims at being inclusive and, while these two aspects are relevant, the central concern is in interactivity as the process that develops when the two agents are in contact and acting and reacting to each other.

The approach presented here focuses on the process of dialogue that is created between the artificially behaving agent and the human interactor. Not on the subjective interpretation of the latter. As noted by Simon Penny, "a human viewer can have varying experiences [in watching a photograph] due to personal associations, varying proximity or lighting conditions, but there is no 'interaction' in the sense of an ongoing sequence of mutually determining actions between two systems processing agency, or as interacting components (user and machine) in the larger user/machine system. The bifurcation in such conversations is whether the critique addresses the experience of the ‘user’ or the behavior of the system” (Penny, 2011). The type of responses of the system in this ongoing sequence of actions constitutes its behavior, which in turn becomes the basis of its agency.

Finally, it is useful to distinguish between interactivity and participation. While the first always implies the latter, the distinction is helpful in order to frame the different kinds of activity that can take place in technological systems that afford the participation of users. The distinction is especially relevant in the context of the web 2.0 and social media, where participation of users is a central concern. The proposal here is to understand, in this context, participation as contribution. That is, the creating content by the user of the system as a different act than that of interacting with this content.

As discussed, interactivity is the participation in a back and forth of actions and reactions with a system. It implies an active engagement with an interface during a certain period of time, which is when the process of interactivity develops. It is, indeed, a participatory process: the interactor participates in the dialog with the machine while interactivity is happening. But participation is also a broader concept, which includes any form of contribution of a user to a (generally online) system. Uploading media to a social network is an act of participation, and it does involve interactivity when the actual uploading is happening. But it is a trivial and mechanistic kind of interactivity. Thus, the interest in such cases does not reside the interactivity as described above. It is on this very act of contributive participation that, in turn, constitutes a form of social interaction.

7.3 Poetic Interaction: The aesthetics of Behavior

The proposed definition is useful for interactivity in general, either instrumental or artistic. However, it is useful to distinguish between these two approaches to interaction. Instrumental or functional interactivity is the domain of Human-Computer Interaction and Usability: how to create interactive experiences that are, above everything else, effective in terms of users achieving the pursued goals when interacting with the system (looking up a file on a database, buying a plane ticket online, selecting a text on a word processor and changing its color, etc.). Alternatively, artistic interaction aims at engaging the interactor in an experience that may have nothing to do with achieving specific goals, just as it can be the case in any artistic experience in general. Functional and artistic interactivity have fundamentally different motivations and, while the interaction design strategies may often coincide, it is useful to elaborate on their differences in order to understand the implications of these two approaches both in the creation of interactive artifacts and when developing a discourse on its design and implications.

A characterization of these two types of interaction has been elaborated by Ignasi Ribas, who writes from the point of view of the design of the experiences. His interest resides in researching the specificities of the interactive discourse (as opposed to the discourse of linear narratives) as it developed from the early 1990s CD-Roms to the current smartphone and tablet Apps. His main focus is placed on cultural dissemination products (from interactive documentaries to museum Apps), but the approach coincides with that of artistic interaction, since the interest is not on creating efficient but motivational interfaces and interactive experiences (Ribas, 2010). Ribas differentiates between 'design for efficiency' and 'design for motivation'[55]. The first is the equivalent to functional interaction. Here the transparency of the artifact is desired, and what the designer aims for is for a presentation of interfaces and controls that allow the user to perform the desired task as fast and efficiently as possible. Within this approach the links from what the user does to how the system responds have to be clear and consistent, so the user can anticipate the outcome of his or her actions. On the other hand, design for motivation seeks to engage the user into exploring the possibilities of the system he or she is interacting with: "it is in the indirect solutions, which playfully try to get, maintain and augment the user's interest while the knowledge transmission

---

[55] The original term is 'estímul' in Catalan, which would be literally translated by 'stimulus'. However, a less literary translation such as 'motivation' is probably a better fit. The idea is that the design should encourage the users to experiment with the piece, try out options and discover hidden features.

process is happening, where we can find the most innovative approaches [to interaction design] from the point of view of interactive communication" [56] (Ribas, 2001).

More directly related to Art, Simon Penny has proposed to understand this approach as poetic or aesthetic interaction: the aesthetics of behavior (Penny, 2000; 2011; 2013). The idea is that, when designing interactive art artifacts, the type of interactivity that one is seeking to create is fundamentally different that in the case of functional interaction. Poetic interaction is the kind of interactivity that develops within a system formed by an interactor and a behaving system that has been designed with artistic purposes. The aesthetics of behavior refer to the analysis of such type of devices and the experiences they afford.

It is a type of interaction that relates to that of videogame design. In game design the objective is not to facilitate as much as possible that the interactor will reach his or her goals as fast and clearly as possible. Systems are designed, instead, so that goals are difficult enough to get so that the experience is engaging, but not too difficult as to be discouraging. However, these motivational design strategies do not necessarily affect interactivity directly in terms of the user being able to exert control over the system. In fact, in this respect interactivity is functional. More often than not the user of the videogame is expected to be familiarized with the controls of the game (keyboard, mouse, joystick, gamepad, etc.) and even to be an expert in controlling his or her avatar. It is in how the game develops that this control becomes not an aspect that affords rapid goal achievement, but one that creates another type of experience.

In poetic interaction, the poetics comes from the behavior of the system and from the way the interactor relates to it. Behavior and interface are the two key elements here. Differently from game design, the interfaces of interactive art are often designed to surprise the user, offering either custom made devices or unconventional modes of interaction. Usually, the discovery of how these work is part of the aesthetic experience. Section 7.4 discusses some aspects of interface design.

In terms of behavior, the artistically behaving systems are different from functionally interactive and gaming systems. The first are meant to be predictable and easy to use. The design for efficiency is just that: interfaces must be clear and informative, so that users can achieve their goals with ease and the behavior of the system must be as predictable as possible. Gaming systems, in respect to that, move one step towards motivation. Interfaces are not necessarily as informative, and often users must explore and try out options in order to move forward in their gaming experience. The behavior of the system becomes less predictable in some occasions. In others, it remains

---

[56] The original text, in Catalan, is: "És en les solucions indirectes, que juguen a captar, mantenir i augmentar l'interès de l'usuari mentre es produeix el procés de transmissió de coneixements, on es poden trobar les propostes més innovadores des del punt de vista de la comunicació audiovisual interactiva."

predictable but its control becomes more difficult (i.e. the acceleration of a Tetris game).

Aesthetically behaving systems move yet one more step towards unpredictability. Systems are neither necessarily easy to control in terms of how the interface is presented, nor are they necessarily predictable in their behavior. The idea is that they do so without following into what is perceived by the interactor as complete random behaviors. Otherwise they become uninteresting in terms of interactivity, since when that is the case all sense of dialog or control is lost[57]. Thus, artistic interaction is situated in the search for a point of equilibrium between what is predictable and what is not. According to Simon Penny it “should not be predictably instrumental, but should generate behavior which exists in the liminal territory between perceived predictability and perceived randomness, a zone of surprise, of poetry” (Penny, 2011).

And in this liminal territory it seeks to create conversation. Moving away from the linear responsive model of interaction, as described by Dubberly, Pangaro and Haque in (2009), conversation is here again a more elaborated mode of relation to the artificial system: “artistic interaction [can be conceived] as an ongoing conversation between system and user rather than the conventional (Pavlovian) stimulus and response model” (Penny, 2000).

This equilibrium between randomness and predictability resonates with Cariani’s account on the use of the Wundt curve to theorize about novelty and experienced pleasure discussed in chapter 5. The basic idea is that, on one end, the absence of novelty (in an artistic experience) means no arousal and, thus, boredom. But on the other end a high level of novelty generates high levels of arousal, and this is experienced as unpleasant. The art audience has a series of expectations, and novelty is about violating some of these expectations: “The degree to which a new piece shocks (and its unpleasantness enrages) its audience is an indication of how many expectations have been violated” (Cariani, 2012). The goal, then, is to try to find the right balance between what is its expected and what is not. To find “an optimal level of novelty that engages to produce moderate levels of arousal that are experiences positively.” This very same idea is applicable to poetic interaction. The interest resides in finding the right balance between what is controllable and what is not: “a point of equilibrium

---

[57] An example of such a system is The Synthetic Oracle by the research group SPECS from Universitat Pompeu Fabra (SPECS, 2008). It is an installation formed by a series of 3 meters (9 feet) tall light tubes that change colors as users walk by the space. It is described as “a computer-controlled matrix of sensitive and reactive pillars, cameras and microphones that interact with its visitors and establishes a dialog with them.” However, whilst the complexity of the behaviors is quite remarkable, in terms of user experience the experience is no different than that of moving among a series of randomly illuminated light poles. The system interacts with visitor’s movements and generates light patterns that are too difficult to perceive from within the installation space.

between randomness and predictability that can result in a behavior with which one can establish a conversation" (Soler-Adillon, 2010b).

Finally, it is worth noting that the notion of 'goal,' which is crucial when defining functional and gaming interaction, acquires a very different connotation within the paradigm of poetic interaction. The functional goal of getting a task done, or the goal of achieving a certain milestone in a game, while they are not banished from the design of artistically conceived interactive experiences, are often irrelevant. In the same sense that looking at a painting or reading a poem is not usually done with a functional goal in mind, interactive art is designed to afford an experience and not as a tool to get something done.

Poetic interaction lives in the realm of art and experimentation. In it, interactivity and behavior becomes an aesthetic concern and, as discussed in section 7.5, this has a series of implications in this type of art creation and on the discourses around it. In terms of understanding how interactivity is designed and understood, as said above this approach is to be interpreted as differentiated from functional interaction or videogame design. The objective of poetic interaction is to experiment with the possibilities of interactivity. As noted by Krueger when describing his early video tracking experiments, interactivity becomes here a central aesthetic concern (Krueger, 1989) and, as such, becomes part of the artistic interest in these pieces, although not necessarily in an exclusive manner.

The methods of poetic interaction may coincide with those of functional interaction, and some of the design principles that apply to the latter apply to the former. But since the objectives when designing such systems are fundamentally different, their interpretation has to be made from a different point of view. These experiences, the processes that develop as the interactor and the system act and react to each other, shall be read as any other aesthetic experiences: a dance between expectations and surprise where the interactive artifact seeks to engage the user into exploring and experimenting with it; a dance of agency, to use Pickering's terms. The interactive system is not conceived as a tool. Instead of being a system to help a user perform a task efficiently, in this paradigm it exists on its own right, as an entity –a behaving agent– that will drive the experience of the visitor to the piece.

7.4 The Design of Interfaces

An interface, as it is generally understood, is what connects either two systems or two components within a system. Interfaces can be physical (a lever, a knob, a USB port, etc.) or digital (a software protocol, an option on a menu, a button on the screen, etc.). Typically, interfaces in computer systems have been studied within the parameters of Human-Computer Interaction (Shneidermann, 1993; Dix et at., 2004 Moggridge, 2007; Sharp et al., 2007). Classic HCI is interested in designing functional interfaces –the 'design for efficiency' of (Ribas, 2001). In the late 1980s, Donald Norman published *The Design of Everyday Things* (Norman, 1988) which has become a classic design textbook. In it, he defines the principles for the design of objects that are easy and efficiently usable in the terms of what later would be known, when applied to digital interfaces, as usability. Norman's basic proposal, although it was formulated in the late 1980s and focused on analog and electronic interfaces, can be used as a guide for the design of digital and physical interfaces for interactive systems. If the aim of a technological system is to have interfaces that inform the user how to use it efficiently, then following Norman's guidelines on visibility, feedback and navigation is a very sound starting point. If the approach is that of poetic interaction, then these guidelines are the rules to bend when designing the experience (Soler-Adillon, 2012).

As discussed above, the concern of interactive art is not the design for efficiency but poetic interaction. And this affects the design of the interfaces that have to afford this type of interactivity. This does not mean that the usability principles do not apply to interface design in the artistic context. But since the final objective is fundamentally different, the design rules are often applied in a different manner, thus generating less efficient but more motivational approaches that include experimentation, discovery, surprise and similar notions.

Because digital art appeared in parallel to the availability of computers, during its early years interactive art was faced with the challenge of designing interfaces for generally untrained users (Penny, 2011). This did have its advantages, since interaction modalities that are nowadays a commonality such as full-body interaction, touch screens, light tracking or the use of external electronics were then novel enough to attract the public and represent an interesting experience on their own, which added to the value of the interactive artwork. Jeffrey Shaw used a bicycle as an input interface in his 1989's Legible City (Shaw, 1989), Jim Campbell the user's body (through video tracking) in Interactive Hallucination, presented in 1989 (Campbell, 1999), and Christa Sommerer and Laurent Mignonneau used a flashlight with which the user pointed at a projection screen where a virtual light unveiled the illuminated part of a virtual world in the 1994 installation Phototropy (Sommerer and Mignonneau, 1995). These are only three of many examples of interfaces that were, on their own right and independently from the piece they were part of, quite amusing to an audience that was not accustomed to dealing with these types interactive technology at that time.

The successive commodification of interactive technology, first in the form of the keyboard-mouse desktop with digital menus and buttons and latter in multiple modalities, has not liberated the artists of the need to develop interfaces for the untrained user. As noted by Simon Penny (2011), this is due to the fact that, with some exceptions, interactive art is about presenting the interactors with novel or unexpected modalities of interactivity. This, affirms Penny, leaves out of the realm of interactive art an entire modality of interaction design: the design of interactions for the virtuosic user, the equivalent of an experienced videogame player in a particular game or saga, or the skilled user of a particular commercial software. Interactive art is not about these virtuosic systems but about lay interaction.

This means that, generally, interaction in this context is designed for a user that approaches the system for the first time. One possibility is that it presents a novel interface. An object that works as a kind of magic wand to explore a virtual space, a video that responds to the position and movements of the visitor, a table with unlabeled buttons, Shaw's bicycle and other unusual devices can be used as control devices in these installations. In these cases, part of the experience is learning how to establish a meaningful relation to the interactive system, in terms of determining what the user can control and what he or she can expect from the system. Naturally, as these types of interfaces enter the commercial realm, users become accustomed to it (they use Roombas's at home, or play kinect-based videogames). Thus, surprising them with novel interfaces becomes a question of using systems that have not yet been commodified.

However, lay interaction can also consist in presenting systems that behave unexpectedly, and this can be done both with novel and with known interfaces. TechnoSphere, for instance, is a piece that is designed to be used through a web browser and the email, and with a mouse and keyboard by users who already know how to operate these devices and software.

As a result, although the artists main focus is on poetic interaction (or design for motivation), they do need to take into account the principles of the design for efficiency. And while these systems are not designed only and mainly from the point of view of usability, precisely because they are created to be used by untrained users they need to efficiently convey what they need to in order to become interesting. Artists want the visitors to their work to engage in it as soon as they step on the installation space or grab hold of the controlling interfaces. And they have to be able to do so intuitively: "no one wants to do a tutorial for an artwork" (Penny, 2011).

In any case, regardless of the technological sophistication (or the lack of it) of the designed interfaces, the general idea of Penny's poetic interaction is that the interactive work of art should provide responses that a user can intuitively understand and relate to. This is the liminal territory between perceived randomness and predictability that seeks to create a conversation.

7.5 The Aesthetics of Interactive Artificial Life Art

Artificial Life Art is not always about interactivity. Works involving cellular automata or autonomous agents, either robotic or in simulations, are often not designed to interact with the public. Breed, Rule 30, The Conversation or Evolved Virtual Creatures are examples of this. Works such as A-Volve, Petit Mal, Autopoiesis and many others, however, do have interactivity among their central concerns. This section deals with the intersection of the aesthetics of interactive art in general and ALife Art.

As a subset of interactive art, interactive ALife Art faces its same aesthetic concerns. As discussed by Simon Penny (2011; 2013), it can be argued that the traditional aesthetic discourses fail to properly cover the practices originated around the concept of interactivity. The notion of an artistic behaving system arose with those practices, and the discourses have yet to fully develop a comprehensive discourse. The interactive art system expands the artistic practice to a realm that had not been covered by any of the previous practices. Notions such as behavior, agency or real-time interaction are novel in respect to these preceding creative efforts: “The behaving artwork is an agentic system which changes its nature, its expression, in real time, as a result of changes in its environment” (Penny, 2013). These agentic artworks are behaving systems, and this behavior constitutes how they interact with their visitors (the interactors) and their environment. In order to properly engage in an aesthetic discussion regarding this interaction, new aesthetic questions about the meaning of interaction appear, which were inexistent (or irrelevant) in the discourses on traditional art objects (Penny, 2011).

By dealing with behavior, interactive art relates to process and experience. Interactivity happens as a process which consists of a series of actions related to each other, as discussed above. And what happens in this process is what becomes central, in artistic, terms, in interactive art. As said, Myron Krueger, the pioneer of computer vision, stressed the importance of interactivity as an aesthetic concern (Krueger, 1989). His systems were visually very simple: people’s silhouettes and primary shapes. What he was interested in was in how users interacted in real time with the virtual creatures and among themselves, not on creating sophisticated visuals[58]. The focus, thus, was on the process that develops with interactivity. The user does something, and the system responds in a meaningful and, to some degree, predictable way.

There is an aesthetic departure here from the traditional plastic or visual arts’ discourses that impose “an axiomatic subject-object distinction upon the artwork/interactor system” (Penny, 2011). The focus on interactivity is a step away from the object and towards its relation with the interactor. The work is neither just the behaving system nor

---

[58] It is a safe assumption that sophisticated visuals were not really an option for Krueger. It is remarkable enough that he was able to do real-time image analysis over a decade before than anyone else and build a system that worked so efficiently and with virtually no perceived latency for the users.

its behavior, but it encompasses too the situated actions of the visitor, the 'contextualized doing' of whomever interacts with the system, much in accordance to the performative ontology that Andrew Pickering describes as the basis of Cybernetics[59]. As Simon Penny puts it, "the lesson of performativity is that the doing of the action by the subject in the context of the work is what constitutes the experience of the work. It is less the destination, or chain of destinations, and more the temporal process which constitutes the experience" (Penny, 2011). This connects with the ideas of embodied interaction and enaction, and in this respect Penny uses Tim Ingold's notion of skill: a performative, situated and embodied capacity, to propose an aesthetic approach that would find in this notion one of its central ontological concepts (Ingold, 2001; Penny, 2013). Other authors also stress the importance of engagement in discussing these issues (Edmonds, 2010).

Thus, creating artistic interactive systems means designing performative artifacts that will relate in a meaningful way to an interactor. As said above, this means that there are mainly two focus of interest: interface and behavior. Within the system formed by interactor and performative artifact, the first is the point of connection between them. In accordance, its design is crucial in establishing this relation. The artist designing the artifact knows that the interactor relates to his or her environment by doing things on it. And according to this approach the idea is that he or she will acquire the necessary knowledge to relate with the piece through this embodied doing, not through an objective abstract examination of it. This means that, as opposed to the design of a website from the point of view of usability, for instance, the design approach must take the actions of the interactor into account at all levels. The analysis of the artifact's behavior comes through the doing things with it relating to its interfaces (using the buttons, moving in front of it, etc.).

Behavior is the essential element in defining the artifact and the interfaces are the means of establishing the relation to it. All other aspects of the piece are accessory to behavior. This doesn't mean they are not important, but external appearance, for instance, is here one step below in aesthetic relevance than behavior. Petit Mal, the Colloquy of Mobiles, Performative Ecologies or Autopoiesis are pieces that can be analyzed aesthetically from the point of view of appearance, in sculptural terms or in relation to the technology of the time. But according to the approach proposed here it is in terms of performativity, of behavior, that they are most relevant. And since they relate to their visitors in a meaningful way, and indeed are designed to explore such relations and to provoke exploration of behavior in the part of the human interactor, they are aesthetically relevant in terms of interactive behavior.

---

[59] See chapter 4.

7.6 Aesthetics and Emergence

There are some efforts within the ALife Art literature to characterize the aesthetics of emergence. However, none of them offers a very comprehensive approach or a definition of emergence that attempts to cover all the instances of the idea. Melanie Baljko and Nell Tenhaff, in the context of their work on the ALife Art piece LoFi, published one of the most direct attempts to cover this ground in (Baljko and Tenhaff, 2006; 2008). In these papers, they advocate for a need to devise the means to recognize and interpret emergent phenomena in interactive art. They recognize Cariani's contribution and his influence on Mitchel Whitelaw. According to Baljko and Tenhaff, emergence does not converge into a single notion in the literature, but they emphasize one point of consensus, in that "'emergence' is something more than the mere 'emersion' (sic) or 'appearance' of behaviors/properties"[60] (Baljko and Tenhaff, 2006).

In their analysis, the characterization of interactive ALife Art conflates with what above was referred to as poetic interaction and, hence, to interactive art in general: "we feel strongly that interaction with an A-life artwork is simply another kind of human-computer interaction, but also an interesting and unique arena of interaction in that performance-based measures are typically not applied, not relevant. A primary goal for the user is for he or she to engage in interpretive professes in relation to the artwork's behaviors as well as his/her own behavior." Following this, the authors identify emergence as the result of the interaction: "the synthesis of these processes is emergent relative to the experience of the work" (Baljko and Tenhaff, 2008).

According to Baljko and Tenhaff interaction cannot be designed. One can only design interactive media that affords interaction. Hence, interactivity is the result of the participatory actions of the user and the interactive system. User and system co-construct the experience which, according to the authors, is in all cases emergent: "The human user's actions in the context of the interactions are 'participatory' actions, and the artifact's actions are participatory too (even if only in the sense that its states changes are context-dependent). It follows, therefore, that the interaction is 'emergent' to the participatory actions of the engaged actors" (Baljko and Tenhaff, 2008). Although he is not cited in the text, this approach resonates with Andrew Pickering's notions of agency and emergence; the emergent 'dance of agency' that would constitute what Baljko and Tenhaff identify as emergence here.

---

[60] The choice of the word 'emersion' here is slightly odd, since its main literal meaning is actually "the act of emerging" (Encyclopædia Britannica Company, 2014). However, it can also mean the reappearance of a body after being hidden or submerged (Oxford Dictionaries 2014), so it is likely that it is in these sense that they use the word. Since the word is accompanied by 'appearance', we can assume that it means the simple act of appearing of something that was hidden. What the authors want to convey is that emergence is not just the appearing of something out of the blue, nor of something that was there but not yet had been seen.

As they advance in the discussion, the authors move towards a more elaborated definition of the concept of emergence, which can be read as being in a middle ground between Cariani's theory and a subjective understanding of emergence. They refer to Cariani's identification of the 'observational frame' from which emergence can be identified, and they differentiate two types observational frames: "(i) that of the interactant, and (ii) that of the outside observer/designer, for whom the model consists of the A-life artwork and the interactant(s) engaged with it." From this idea, they make explicit what they understand for emergence, and by doing so they point out the subjective implications of the notion: "in terms of emergence in relation to aesthetics, what matters is what is subjectively perceived about the behaviors of the artifact in relation to the behaviors of the interactant(s). In A-life arworks –and more generally, for digital artifacts– synergistic patterns of interactional behaviors are emergent in respect to both observational frames (i) and (ii)" (Baljko and Tenhaff, 2008). The remark 'and more generally, for digital artifacts' recognizes how, despite they discourse is focused on ALife Art pieces, they address in fact the aesthetics of interactive systems in general.

Therefore, with their argumentation, Baljko and Tenhaff identify interactivity and emergence. According to them interactivity, which cannot be designed per se, is necessarily emergent. This argument eliminates de facto the differentiation between emergent and non emergent interactivity. It follows that the notion of emergent interactive behavior is redundant and, therefore, useless in order to analyze the presence of emergence in interactive media.

This identification, however, is problematic. First, despite their references to the observational frame of the interactant and of the external observer, what exactly is emergent and what is not when the behaviors of both system and interactant are observed is left unspecified. This omission is reminiscent of the most ambiguous uses of emergence in some Artificial Life literature, were the concept is both taken for granted and left to the intuition of the reader (precisely the problem that Cariani addressed in the elaboration of emergence-relative-to-a-model)[61]. Second, the move towards subjectivity makes it impossible for this notion of emergence to be scientifically communicated and addressed, as Cariani had too pointed out[62]. Finally, it can be argued that this importance on the subjective perception makes it impossible to determine where interactivity and emergence start. It is essentially the same argument that was made regarding Mark Bedau's importance on the simulation in order to decide whether or not the behavior of a system is emergent[63]: the line between what is emergent and what is not is in Baljko and Tenhaff's argument, too, diffuse at best. Then, if emergence and

---

[61] See sections 5.1.1 and 5.7.1.

[62] As noted in the introduction to chapter 5.

[63] See section 2.7.

interactivity are synonomous, not only the first but both notions fall into this undefined terrain, making it difficult to state whether or not a system is actually interactive.

Another important approach to the aesthetics of emergence was proposed by John McCormack and Alan Dorin in (2001). They too recognize the difficulties of defining emergence within the interactive art discourse. McCormack and Dorin are concerned mainly with generative art: how computers can generate processes which, from a set of simple rules, evolve in time as one can experience how the behaviors (usually as visual patterns) develop. Note that generative art is neither necessarily interactive, nor necessarily related to ALife. In their discussion on emergence, McCormack and Dorin conclude that what is important regarding their aesthetic interests is the impossibility of prediction that this idea implies. An unpredictability, that is, which does not imply inconsistency. It is again a subjective approach. What is initially unpredictable is consistent in retrospect, and can be expected to occur again if the same system is run with similar initial conditions.

In generative art, creating unanticipated forms or behaviors is a very attractive goal, hence the interest in this sort of emergence. In order to understand how this can be done, they resort again to Cariani and emergence-relative-to-a-model. Here the model is analogous to Don Norman's mental models (Norman, 1988): a mental image that a designer has of how a system works and relates to the world, which has a correspondent mental model that is formed by the user of a system[64]. This mental model informs the user of what the system is capable of doing and what the user has to do in order to achieve his or her goals. Thus, for McCormack and Dorin, emergence relative to 'that' model –the mental model of the designer– is the appearance of forms or behaviors that, apparently, were not designed into the system. And this is the basis of what makes these kinds of interactions interesting: "engagement with the computer needs to suggest that the work is 'more' than its design intended it to be" (McCormack and Dorin, 2001). For a detailed account on generative art see Dorin et al. (2012).

Other authors refer to emergence in the context of interactivity. Dan Collins defines emergent behavior as "the name given to the observed function of an entity, and is, generally speaking, unrelated to the underlying form or structure," which exists due to the interactions between the parts and their environment and cannot be reduced through analytic analysis. He calls 'interactive emergence' the interactivity with systems exhibiting such phenomena (Collins, 2002).

---

[64] For Norman, who writes in the context of traditional, object design it is desirable that the mental model of the designer and the user coincide. This idea would become basic for usability: an interface must inform a user of how a system works in an efficient manner, so that this user can form a correct mental model of the system as soon as possible, either with the first phases of usage or even before the first use of the system. Very often, this is precisely the rule that digital art is interested in playing with.

Hendriks-Jansen, in the context of situated robotics, writes 'in praise of interactive emergence' (Hendriks-Jansen, 1994). He refers to Nagel (1961) to differentiate between two types of emergence which, as he describes them, would essentially coincide with the two proposed in this thesis, as one essentially refers to self-organization, while the other to novelty[65]. Hendriks-Jansen then proposes a third type, interactive emergence, which he exemplifies with a wall-following robot that was not programmed specifically to follow walls, but that did have capabilities of obstacle avoidance among others. His argument is that the wall-following behavior is emergent, since the robot does not have any formal definition of a wall, nor is it specifically 'told' to do so. However, that this is a case of emergence can be disputed. Certainly, none of the parts of the robot, nor a hierarchically superior entity is designed to produce the wall-following behavior that is accounted for by the (external) observer. But the wall-following behavior is nothing but the avoidance of a continuous line of obstacles. Therefore, there is nothing new, nothing emergent, in respect to the predefined obstacle-avoiding behavior.

Other approaches basically conflate emergence to openness (Seevinck and Edmonds, 2008; 2009). These authors define the idea as follows: "emergence occurs when a new form or concept appears that was not directly implied by the context from which it arose. The emergent whole is greater than the sum of its parts: it is new and different to what was there before. It is unpredictable or not immediately deducible." As an example, they show the triangle that appears at the intersection of two squares. According to them, the triangle is an emergent form in respect to the two squares that overlap.

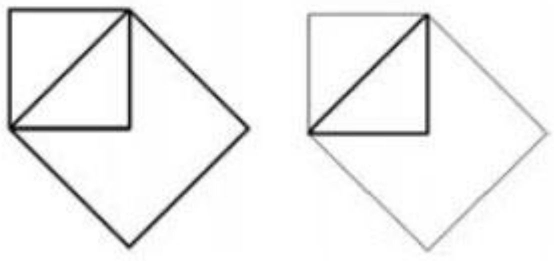


Figure 7.2: Two squares overlapping squares and a triangle (Seevinck and Edmonds, 2009).

In Seevinck and Edmonds' system Plus Minus Now, an interactive installation, users encounter a sand box interface. As they make shapes on it, the computer will project, onto a frontal projection screen, combinations of these shapes with others that were previously drawn by other users. According to the authors, the forms that appear at the intersections are emergent shapes.

---

[65] See section 8.3 for a detailed account on Nagel's theorization of emergence.

This does not fit into either of the forms of emergence that this thesis defends: self-organization and generation of novelty, and indeed it can be argued that these shapes are not emergent in any way that relates to a meaningful formulation of the term. In fact, this is precisely what Baljko and Tenhaff dubbed “the mere ‘emersion’ (sic) or ‘appearance’ of behaviors/properties” (in this case, of shapes). The shapes appear, and they were certainly not there before. But if these shapes are emergent, any shape, color and sound one can experience is emergent too, since none of them is composed of any singular essential element, nor was preexistent in epistemological terms. There’s no behavioral level from which these forms can be originated in terms of self-organization, nor there is any model that the interactor can relate to for the shapes to be emergent in respect to. Therefore, in this case, the conclusion is that the term emergent is misused.

Finally, it is worthy to note that some other approaches to the subject are considerably vague. For instance, Ken Rinaldo (1998) identifies emergence with the appearance of new synergies among researchers and practitioners from different disciplines, basically artistic and scientific, that originated ALife Art. He states that emergence is “the result of the collapse of both scientific and artistic barriers which have contributed to the rise of Artificial Life art.” For him, it is “the new paradigm for a global change encompassing this earth.” On a side note, he offers this definition of the concept: “emergence is used broadly to describe the process by which new artistic and behavioral forms can bubble up to the surface in a culture rich with information and technology. It can also be considered a process of unpredictable development where unexpected elements come into view because of the attendant cross fertilization of ideas that the information makes possible.” It is a definition that does not help to reduce the vagueness with which the concept is used in the text.

In conclusion, the aesthetics of emergence need a clear and concise definition of emergence in order to move away from vague claims about a diffuse notion, as it has been the case on several occasions. This objective can find a stable context within the aesthetics of behaving systems and poetic interaction proposed by Penny in (2011; 2013). In it, emergent interactive behavior would be a subset of the kinds of interactivity that are investigated, both as self-organizing phenomena or as related to the appearance of novelty.

7.7 Interactive Artificial Life Art

ALife Art is part of digital art and, as such, some of it is about utilizing and discussing interactivity and some of it is not. Animations, simulated environments and evolved artwork often do not consist of systems that interact with their visitors, but fall instead into the category of contemplative artwork. This does not exclude the presence of behavior in these works. In some of them, although contemplative in the sense that they are not interactive (in respect to an external user), there is interaction amongst the parts of the system and, thus, there is a system behavior that can be observed.

Among the examples presented in chapter 6, such is the case in Breed, Rule 30, Human Cellular Automata, Sand Lines, Substrate, The Conversation and Evolved Virtual Creatures. The rest of the examples are classified in at least one of the subsections that follow, which examines different modes of interactivity: selection, creation, disruption, agency and cumulated interaction.

	Selection	**Creation**	**Disruption**	**Agentic interaction**	**Cummulative interaction**
Abstract generative systems			·Propagaciones		
Evolutionary and GA	·Genetic images	·A-Volve	·A-Volve ·Autopoiesis	·A-Volve ·Autopoiesis ·Performative Ecologies	·Performative Ecologies
Simulated ecologies		·TechnoSphere ·Iconica	Digital Babylon	·Iconica ·Digital Babylon	·Digital Babylon
Autonomous agents	·El Ball del Fanalet		·Sympathetic Sentience	·Petit Mal ·Tickle Robots ·El Ball del Fanalet ·Sniff	

Table 7.1: ALife Art categories (rows) and types of interaction (columns).

The descriptor 'interactive Artificial Life Art' is used here to delimitate artwork that, while falling into the ALife Art category because of its themes and methods, is conceived as an interactive system in the terms discussed at the beginning of the chapter. These are systems in which the intervention of one or more people (the interactors) is a central aesthetic concern, in the sense that interactivity is an aesthetic concept as described above. In the discussion that follows this poetic interaction is assumed, as well as the proposed definition of interactivity. Although the presence of emergent phenomena may be occasionally mentioned here, detailed discussion on how emergence is present or not in these works is left for the next chapter, where a framework of analysis is presented and exemplified.

7.7.1 Selection

Selection is a simple type of interactivity that is found in pieces such as Sims' Genetic Images, some of the work of William Latham or the evolved buttons of Michael Kass and Andrew Witkin. These pieces operate with genetic algorithms: from an initial pool of possibilities, a series of individual entities are created in what constitutes a generation. But instead of having an algorithm evaluate the fitness of each entity according to the predefined criteria, here the selection is made through human judgment: either the artist himself or the pubic chooses the images to breed. In the first case the judgment is based on personal aesthetic criteria. In the second, it is left open to the decision of the audience. William Latham or Steven Rooke are artists who experimented with the creation of evolved images in this way. They would create a series of images and then select those that were more interesting to them as the basis for subsequent generations. The process would iterate as long as considered necessary until they reached a result they were satisfied with. As said, the alternative is to let the audience be the judge, as did Karl Sims in Genetic Images.

When analyzing the results of such experiments Galanter's critique, reviewed in the previous chapter, comes to mind (Galanter, 2010). There's certainly a sort of 'sameness' in the works of Sims, Latham, Rooke and others who pursued this evolutionary approach to image creation, such as Japanese animator Yoichiro Kawaguchi[66]. In this case what drives the selection is the personal criteria of the artist, who might be focusing on what is the most appealing at any particular time, or be familiarized to the process up to the point that the selection is made thinking on how each selected instance will constitute the basis for further developments, regardless of its particular appeal at the moment it is selected. When the selection is made by the audience, it is even more difficult to state what exactly is driving the choices. According to Galanter, the problem here is that novelty prevails over quality.

In terms of interactivity, while strictly speaking these systems are indeed interactive, they propose nothing more than the equivalent to a selection from a database. Either the

---

[66] This 'sameness' is not exclusive of this group. The same argument could be made about many artists and art movements. However, since usually genetic algorithms are presented as a technique that affords for unexpected choice (and thus for unexpected variety) Galanter's critique is relevant.

artist or the audience chooses, but they choose from a series of closed options, and the act of choosing is a direct mapping of an action to an effect: a pointer click on a screen or standing in front of an image for some time. It is a functional design. There are no poetics, no design for motivation, in the interactive action which constitutes the choice.

An exception to this is El Ball del Fanalet. In it, users do not choose from a series of options. They are actually offered with an all-or-nothing choice: users can only select a proto-creature that matches the color of their paper lantern. A closed option that actually fits quite well with the metaphor of the dance: a potential dance partner shows up, and you either dance with it or you don’t. Thus, the option is even more predetermined that in any of the previous cases. However, the way the choice is presented to the interactor is not perceived as such. And this is so due to the poetics of the piece. The fact that the creature shows up from a hidden world, and that choosing to engage with it is done through the portable illuminating device affords a type of interaction that does not relate to database-driven interactivity. What the user ends up discovering is that, by holding the fanalet steadily on top of a proto-creature of the same color, he or she has selects this creature to interact with (make it grow and teach it dance moves). The fact is that, at that particular moment, no other choice could have been made.

7.7.2 Creation

As discussed in the previous chapter, some ALife Art projects are built around the idea of the ecosystem in a more or less literal way. This metaphor affords the possibility for the users to create some of the creatures that inhabit these ecosystems. Creation is, adapted to the ALife Art context, what was above presented as contributive participation. Iconica, A-Volve and TechnoSphere are examples of systems that offer this possibility to the interactors: individual entities can be created by users, and will then enter the ecosystem to interact with the other beings that are already in it. A different approach is found in Digital Babylon, which is closed to this possibility. In it individual creatures are born and perish, but the visitors’ interaction is not designed to intervene directly in this process or to make choices regarding how particular entities will be created.

In terms of interactivity, the manner in which users interact with TechnoSphere and Iconica in order to generate creatures is through database-like interfaces. In the first, users select from a series of choices: carnivore or herbivore, type of head, body, wheels and eyes, etc. When choosing through the website interface, the users don’t know what is happening inside the system. That is, there is no way that they can have an overall knowledge of what is going on within the ecosystem. Thus, anticipating the possible consequences of one’s choice is out of the question. In this respect, it is a blind choice. And once the choice is made, the digital creature disappears into the system. From this point on will communicate only the updates to the users’ email. Therefore, there is still some communication (one-way messages from the creature to the creator) but no interaction.

Similarly, the choices in Iconica are somewhat blind, too. But in this case it is mostly because of the highly abstracted nature of how the Iconica's ecosystem is presented. Here the visitor to the installation can see, at least up to some degree, what is going on in the system before building his or her own creature. But since Iconica is precisely based on an abstract language and on abstract representations of it, only a potentially very expert user would be doing this creation with a sufficient degree of knowledge in order to anticipate how this might actually affect the system.

In any case the actual creative interaction with the system is limited to this initial database-like choice. There are some elements that the user can combine, and once this combination has been made, the interactivity is over. Indeed, this is only a small part of the experience. Both TechnoSphere and Iconica, especially the first, are built on participation (i.e. on contribution) over interactivity. The interactive part just described constitutes at the same time the way the act of contributing to the piece, of creating content that will be a part of the ecosystem. TechnoSphere is interesting as long as visitors participate in it, which is indeed what happened very successfully in the years that the project was online. Iconica does have another level of interactivity besides creation of elements, which is discussed below in section 7.7.4.

In the case of A-Volve, the weight on interactivity and participation is more balanced. Regarding creation, the visitors to the installation are presented with a much more open interface than that of Iconica or TechnoSpere. In it, they can draw whatever shapes they want on the tactile screen, with the only constrain of the size of the screen and the fact that the system will be drawing the symmetry on its other side. Once the shape is closed, the drawing is interpreted by Sommerer and Mignonneau's algorithms in a manner that is unknown to the user, but this is not different from what happens in the other systems, were the creature's behaviors are also given from what was predefined, without further intervention of the user.

The fundamental difference between A-Volve's creation process and the other two examples is in the openness of the digital interface, which instead of predefined choices offers the user with a white canvas. This points to a classic dichotomy in interface design: the limited menu driven selection vs. the more open multiple choice scenario. In the first the designer controls exactly what choices the user is given and how these are presented, and errors are easy to avoid. And while the user might perceive some of the given choices as limiting, it is very clear what can be chosen and how.

On the other hand, the free canvas approach offers, theoretically, unlimited creative freedom to the user. However, what is designed in the white canvas will always have to be converted into something that the system can operate with (in the case of A-Volve, a creature that will go to the pool). Thus, whilst initially the perceived freedom can be expected to be greater in the white canvas approach, in terms of final experience the frustration can be even greater if the conversion from the apparently unlimited to the predesigned choice (how the 3D is extruded from the 2D, how a drawn line will translate to a behavior, etc.) is not accomplished successfully enough.

In any case, the context in which this is done in A-Volve also differs from TechnoSphere, since in the first it is indeed possible to watch and, in general terms, understand what is happening in the whole ecosystem. The pool that hosts the ecosystem is visible at a single glance, and in fact it appears to the visitor as being significantly smaller scaled environment than the other two. Therefore, a visitor who has taken some time to observe how each of the creatures move around the pool might anticipate what kind of shapes will result in some particular movements and behaviors. Plus, since he or she knows how many entities are in the pool, the decision of creating a particular being will also be more informed. Once this creation is done, like in TechnoSphere and Iconica, other types of interactivity are possible. These are discussed below.

7.7.3 Disruption

Some ALife Art projects are based on systems of agents that work without the intervention of the user, but that at the same time afford interactivity. Users can interact with the agents and affect the processes that are happening at the level of the system. The agents that form these systems can be digitally or physically instantiated (e.g. A-Volve and Digital Babylon as digital intantiations; Sympathetic Sentience, Autopoiesis and Propagaciones as physical or robotic). In all these cases the agents are presented as constitutors of a system, and it is at the systemic level that the interaction becomes relevant. While Sympathetic Sentience, Propagaciones and Digital Babylon are systems that exist and work without any intervention of a potential interactor, the case of A-Volve is a little different. In it, the actual existence of the agents in the system depends on previous action by users. However, once the system is in motion –the ecosystem in the pool is populated– the virtual creatures generate their own dynamics of behavior and equilibrium. A-Volve is considered in this section from this point of view of the populated pool, as a system with agents with which the users can interact with. Autopoiesis is designed to work with the visitors in the installation space, but the robots do perform some autonomous system behaviors regardless of the presence of visitors.

Once these systems are set in motion, they function without the need for the interactors to do anything with them. They all produce their own patterns of behavior and equilibrium. This resonates with the cybernetic notion of homeostasis. When left alone, the dynamics of the system equilibrate each other as the sum of all agentic forces procures a self-organized balancing effect.

Sympathetic Sentience slowly builds up its sound loop, as the robotic units communicate to each other sequentially. Propagaciones comes to a stable systemic state, as does any implementation of Game of Life eventually, were the agents either find a repetitive pattern of reaction to their neighbors, or simply stay in a fixed state. Autopoiesis, despite it being designed to essentially interact with visitors, can generate its own dynamics of movement and sound. In Digital Babylon, the peaks and lows of population happen in consonance with abundance and scarcity of food, as the predators make their way in the ecosystem by feeding themselves on the other species. The piece reaches a rhythmic equilibrium after some functioning time. Finally, A-Volve's creatures live, prey and die, and despite the fact that eventually new creatures will have to be introduced in the system for it not to come to a halt, while this doesn't happen the ecosystem in the pool finds its own equilibrium too.

From this point of view, the intervention of the visitors in these systems is disruptive. That is, the essence of the interactivity is to break the equilibrium or the pattern of behavior, and with this introduce new possible developments into the piece. In Sympathetic Sentience, the visitor can place him or herself, intentionally or not, in a space that interrupts the communication of the robotic units. As explained in the previous chapter, this will break the communication and, thus, introduce silence in the sound loop of the unit that looses reception. In turn, this will end up affecting the whole sound loop of the system, up to the point of complete silence if the disruption of the communication is persistent enough. In Propagaciones, what is actually disrupted is stability. The intervention of an interactor excites the units he or she interacts with, and these units' reactions propagate into the whole system, which then generates a certain amount of activity until, eventually, it reaches a stable state again. In Autopoiesis the visitor will attract the nearby arms, which will move and communicate the visitor's positions to others. Thus, as in Propagaciones, it is this disruption which actually activates what the artist is more interested in in the system.

In A-Volve and Digital Babylon the disruption is somewhat more subtle in terms of systemic effects. In the first, users can trap creatures and either help them or help their predators with it. Similarly, in the latter the visitors attract individuals of the main species (S1), and they can choose to either keep them away from their predators or get them close to being eaten. In both cases, what the visitors do affects the current agents populating the ecosystem. The difference lies in the fact that, while in A-Volve there is only a small number of agents at any time (over 15 or 20 the pool would seem to be much too crowded), in Digital Babylon the agents are counted by several hundreds (typically between 500 and 1200 agents populate the ecosystem at any time). Therefore, in the first users will see very clearly how their action affects the individual agents and the current situation of the system, since one or less agents with such small numbers is very relevant. But the effect that their actions will have in the long term in the ecosystem will be irrelevant, since eventually the system will become empty and the cycle will need to start all over again as new users draw new creatures.

In contrast, in the case of Digital Babylon the effect that one’s action produces on an individual agent is virtually impossible to perceive, as the individuals of S1 appear to move in group and they move too fast to be tracked for more than a few seconds, except if the space is rather empty at the time. What the visitors can read from the installation is the overall behavior of the group. But, contrary to the other examples, each action that a user does in Digital Babylon will have a long-term effect in the interactivity of the piece –which is nonetheless imperceptible to the individual visitor. Since having the friendliest S1 individuals killed or saved affects the overall evolutionary process of the species, what every user is modulating with his or her action is the manner in which future visitors will be able to interact with that particular species, as its overall friendliness will grow or decrease as individual interactions either encourage or discourage it. As discussed in chapter 6, these two projects are based on the metaphor of the ecosystem. They also draw upon evolution as an emergent process. This is discussed in chapter 8.

Some other ALife Art systems work in the same way in terms of generation of systemic behaviors and equilibrium, but they do not afford the possibility of disruption by an interactor. Sand Lines, Substrate, The Conversation, TechnoSphere or even Human Cellular Automata also produce similar cycles or generate similar types of systemic equilibrium. However, there is no possibility for a visitor or user of the system to affect these patterns. They are not interactive in this respect –in fact, among these last examples given, only TechnoSphere is actually interactive.

Disruption, therefore, is a type of interactivity that is found in artwork which is built as a system that works independently of the interactions of the users, but that does afford interactivity in a way that allows the interactor to affect the equilibrium and patterns of behavior of the system that are thus either interrupted, excited or eliminated. When, instead of closed system vs. user, these systems are considered with the inclusion of the ‘perturbing’ user, they can be regarded as second order closed dynamical systems. In this context, Maturana’s autopoiesis (the fact that the system is able to maintain itself) can be applied from an art-experience point of view.

7.7.4 Agentic interaction

In artwork where the interactor relates directly to one or several agentic entities, either within the systems described above or in others where the agents are presented on their own, the type of interactivity that is developed is most directly related to what was described above as poetic interaction. It is not just that the visitor to the installation or website generates a creature to be thrown into the system, or that he or she has the means to interfere in the systemic patterns of behavior. In agentic interaction the visitor is directly addressed by an entity that has its own behaviors in relation to his or her own, so a conversation, a back an forth of actions an responses, can develop as the process of interactivity advances.

Here agency is understood in the terms described by Andrew Pickering: as a basic building block of the influence that an entity can have on its environment. Agency is, literally, the “doing things in the world” (Pickering, 2002). And by doing things in the world is how an entity becomes an agent. This is certainly a very broad definition of the term, and Pickering elaborates it in detail in his work (Pickering, 1995; 2002; 2010). In terms of interactivity, it is sufficient to state that an agent is a non passive entity. That is, one that either responds reactively or with actual interactivity to the actions it is able to perceive from whomever is relating to it.

In terms of art practice, the creation of agents with artistic purposes has been studied by Simon Penny in (2000). An agentic artwork, according to Penny, is not just one that does things in responding to its surroundings. It can be understood as a cultural actor in its own right, and it is under this understanding that the notion of an agent as an artwork –like his own Petit Mal– is interesting. These agents are designed and built within a cultural tradition, and in a particular social context. And it is when they are understood and read as such, and not abstracted from their social and technological context, when they can be better understood.

Agentic interaction, thus, is the type of interactivity that develops when the interactor is in direct dialog with such an agent, be it digital of physically instantiated (i.e. virtual or robotic. This agent may or may not be part of a larger system, but this particular type of interaction is concerned with the local level of interaction with the particular agent, and not with the system as a whole if it exists.

Interactivity with particular agents within a system is found in Iconica, Performative Ecologies, El Ball del Fanalet, Autopoiesis, A-Volve or Digital Babylon. In the case of the last three, the users can affect the overall processes of the system, as discussed above, and they do so by interacting with the individual agents. One at a time in the case of Autopoiesis and A-Volve, and with a group of them in Digital Babylon. In any case, these interactions with the particular agents have to be considered from the point of view of the individuals and, thus, locally, as opposed to the interaction with the global behavior of the system. In all these cases, the interaction is based on the attraction that the visitor exerts towards the agent. In A-Volve, the visitor traps the creature with his or her hand, and can choose to help it get closer to another creature, stay away, or just move it around. In Autopoiesis and Digital Babylon, the agents will approach the users’ position, thus generating group behavior through these individual movements.

In el Ball del Fanalet, the relation between creature and user is more complex, as explained in the previous chapter. Although the agents are found within a system, the interaction is always one-to-one and no overall systemic effects are perceived by the users. In any case, it is a relation that illustrates very well and in a quite literal way the creator-creature scheme of may ALife Art piece: the visitor to El Ball del Fanalet will nurture the proto-object until making it grow, and then will give knowledge to it (teaching it how to dance) until the creature is ready to go on its own, either because it finds another mature creature or because it gets bored and moves one (and away from the creator).

In Performative Ecologies, there's no disruptive interactivity, because the system is always waiting for a user. Only as a substitute for idle moments will perform independent tasks (the teaching of one agent to another) but these will be interrupted (not disrupted) if a visitor enters the installation space. The agents in that installation are designed to interact in a one-to-one basis with their visitors, and there is no coordinated group behavior that originates from local interactions among them. Similarly, in Iconica, besides the creative interactivity mentioned above, the visitor can interact, too, with the individual creatures that populate the abstract ecosystem –although with an intended counter-intuitive language.

Petit Mal is a paradigmatic example of agentic interaction. It is not a robot built to be aesthetically pleasing in terms of how it looks, nor to perform complicated tasks or to engage in complex interactive cycles with its visitors. The strength of the piece, as intended by its author, resides precisely in a very neat representation of what being an agent consists of. Because of the fact that any other artifact is removed, exhibiting agency is what makes Petit Mal what it is. The experience for the visitor with this robotic installation is generated not through reminiscences of what it looks like, but thought how it behaves and engages with its human counterpart. The 'doing in the world' of Petit Mal is all what matters and, as said above, it is a very successful example of what being an interactive agent is[67]. In this respect, Petit Mal is reminiscent of how Grey Walter described the idea cybernetic model[68]: it is simple, transparent and brittle (Walter, 1963).

Sniff is another example of agentic interaction. In relation to Petit Mal, it presents some important differences. First, it is not a robotic entity but a 3D model shown on a screen. Second, both and its looks and its behavior, sniff is clearly a dog. Although it does not strive for realism –it is shown as a white polygonal figure with no textures– it is nonetheless undeniable, even by its name, that it emulates the experience of interacting with a real dog. While in Petit Mal this space is left to the imagination and interpretation of the visitor to the piece, in Sniff it is predefined by the artists. However, the focus is indeed on the behavior and not on the looks. A behavior which is more sophisticated, in terms of programming complexity –not necessarily in terms of aesthetic experience– than that of Petit Mal, since the virtual agent does operate, for instance, with memory, while Penny's robot is designed to exist exclusively in the present tense.

---

[67] While preparing a panel for a conference, Andrew Pickering proposed Petit Mal as an example of what he understands as the 'emergent dance of agency' that generates interactivity in relating to a cybernetic system. Regarding the adequacy of this particular piece as an example, he declared that "it is hard to get away from Petit Mal" (personal communication, 06-24-2014).

[68] See section 4.2.2.

Finally, the Tickle Robots are a special case of this type of entity. These are robots with which the user doesn't really interact with. Rather, it is the robots that act upon him or her. They have no control devices, no buttons or levers for the interactors to operate. Instead, here the human side of the system plays the part of the passive entity. The visitor to the installation lies down in a massage bed, and it is the robot that interacts with his or her body. While in Spear this interaction with the visitor is rather simple, in Tickle and Tickle Salon the complexity escalates, from the tickling game to a quasi-erotic experience.

7.7.5 Cumulative Interaction and the question of system memory

The idea of cumulative interaction was formulated in (Soler-Adillon, 2010a), in order to introduce how genetic algorithms afforded a second layer of interactivity to the Digital Babylon interactive installation in respect to those types of direct interactivity described above. The term is relevant to other interactive artworks that make use of this programming technique, such as Performative Ecologies.

The concern that originated the idea of cumulative interaction appeared while designing Digital Babylon. With this piece, a main concern was to overcome the on/off structure that most interactive art installations are built upon. That is, the pieces have an idle state while they wait for an interactor to activate the system and, once this happens, they start to trigger whatever interactive reactions to him or her they are programmed to exhibit. This interactivity can either be very simple or very complex, but it any case it eventually ends when the interactor moves away and stops the process. At this point, the piece is back to the initial state: idle and awaiting the next interactor as if nothing had happened. In order to overcome this, the question was: how can each of the individual interactions be meaningful to the whole system? How can they accumulate and remain meaningful to how the piece interacts with its visitors?

Some participative environments do address this problem. Each act of participation adds to the piece and, thus, alters it up to some degree. However, the question remains: to what degree this actually affects interactivity? Often what will happen is that, since the content changes, the overall experience with the piece will necessarily chance. But that does not mean that the changes affect how one interacts with the piece.

An illustrative example is Camille Utterback's Untitled 5. In this installation, the visitors move around a designated space in front of a projection screen. By doing so, their path is read by a head mounted camera and the system draws on a screen according to their movements. Each interaction (each drawing) accumulates on top of the others. Thus, one visitor's actions will affect what others have previously done on the screen. However, since the screen is a finite space, eventually one's contribution will be completely erased, and no trace of it will be left in the installation.

Figure 7.3: Camille Utterback. Untitled 5 (source: http://csis.pace.edu/digitalgallery/art_talks/fall05/images/camille_untitled5.jpg).

Overcoming this limitation of space and memory was one of the design objectives of Digital Babylon. The solution to this it came through the use of genetic algorithms. With GA, Digital Babylon accumulates each one of the interactions of the users into the behavior of the agents with which future users will be able to interact with. Therefore, each interaction is meaningful, although certainly in a subtle manner, to future interactors as long as the installation is functioning without interruption. A very similar cumulative process happens in Performative Ecologies. Each individual interactive experience affects the behavior of the individual robots in a cumulative way. Therefore, the actual ways in which interactivity happens for future visitor is a result of the sum of all this accumulations of individual interactions. In both cases, the system's behavior is designed to evolve, with the hope that this will generate particular behaviors that were not predesigned, at least not in the exact manner in which they will appear.

Not all uses of GA in interactive artwork generate cumulative interaction. In the case of Sims' Genetic Images, accumulation of interaction does affect the outcome of the piece, as discussed in the previous chapter. Each user's choice determines future possibilities as long as the piece is not reset. However, regardless of the particular images the interactive experience will always be the same for the visitors to the installation. Therefore, in this respect the concept of cumulative interaction does not apply. It is a participatory environment, but the actual interactivity does not change as time passes. Similarly, in A-Volve the creatures change but the actual way in which the user interacts with them remains the same. In addition, the fact that the ecosystem can eventually come to a halt as all the creatures vanish –generating the need for new users to introduce others– puts a stop to the evolutionary process that is changing the current inhabitants of the pool, and then the strictly predefined behaviors become the only ones relevant again.

This notion of the ‘reset’ is interesting in regards to the idea of the system memory. Many interactive art pieces are designed with the cited cycle of idle state/action/idle state, as if they lived only on short temp memory. After each cycle, the reset represents a sort of amnesia, and the cycle starts all over again. In terms of interactivity, only systems that accumulate experiences (i.e. that learn) can eventually change how they act. This represents a step beyond the interactivity that takes into account the previous actions within the current cycle of interaction. It is a step towards the conversation as a sophisticated form of interaction (Dubberly, Pangaro and Haque, 2009). Once more, it is an idea that can be traced back to Cybernetics, with the work of Gordon Pask on learning devices (Pask, 1971; Pickering, 2010).

Finally, there is another aspect of cumulated interaction that is worth mentioning here: it represents a step towards moving away from the database-driven paradigm that many interactive art pieces are built upon. It is the ‘reversion to rest state’ paradigm that was mentioned above: the piece does something and when the interactor moves away it sits there waiting to start the process all over again. This resonates with Lev Manovich’s characterization of the multimedia object as essentially related to database choices: “From the point of view of the user’s experience, a large proportion of [new media objects] are databases. … They appear as collections of items on which the user can perform various operations –view, navigate, search” (Manovich, 2001: 219). Cumulative interaction is in fact a type of system learning. Thus, it fits into the category of learning systems which, along with adaptive systems and devices, aims at proposing a type of interactivity that is more complex than that of ‘reverse to rest state’ system. That is, interactivity that is one step closer to conversation.

7.8 Concluding Remarks: Issues in Interactive ALife Art

Interactive Artificial Life Art is a descriptor that encompasses art practices inspired by ALife themes and techniques that deal with interactivity. Many of the examples presented in the previous chapter fit into this category. The interactions found within these artworks can be categorized into at least one of the types of interaction proposed: selection, creation, disruption, agentic interaction and cumulative interaction. These categories are to be understood under the more inclusive framework of the definition of interactivity given at the beginning of the chapter.

As a subset of Interactive Art, Interactive ALife Art shares its main issues. As explained above, the aesthetic object is here displaced, in regards to traditional aesthetic discourses, from the appearance to the behavior: from a passively observable quality to an experience that develops in the process of interactivity. The interactive artwork becomes an agent, and as an agent it interacts with the public for which it is designed.

But some issues are more relevant in designing this type of systems than when dealing with interactive art in general. Most importantly, the relevance of the complex processes with which these artworks are built in respect to the interactors. This was discussed in the previous chapter with the idea of over-complication. The issue is whether or not complex processes, such as those based on GA, are relevant to whoever interacts with one of these artifacts. It is a question that needs to be addressed in each example separately, and the proposal here is to do so with the premise of avoiding unnecessary over-complication.

The proposed criterion for this is ‘relevance in terms of interactivity’. Since, as said above, interactivity is here understood as a process that involves both the piece (and thus its behavior and design) and the user, there are two points of view to take into account. From the point of view of the user, this can be explained as follows: if the fact that complex processes are involved in the generation of the behavior of the piece has a decisive influence on how the user interacts with it, then these complex processes are relevant. If, from the point of view of the user experience, nothing would change if the behavior was generated though random processes or fixed parameters, then the use of complex structures is irrelevant. On the other side, the use of these complex processes is a design decision, and it will affect, in one way or another, the outcome of the piece. Therefore, the criterion from the design point of view is the same: the use of these techniques is relevant in terms of generating interactive behavior if a similar outcome could have not been obtained using simpler techniques. Thus, the idea is to apply a sort of Ockham’s razor is applied in both ends: interactor and designer.

Another important aspect that is specific to Interactive ALife Art is the relevance of the metaphor of life onto which these artifacts are built. Biological models, ecology and evolution are central both in the techniques used and in the themes of ALife Art. Here a similar argument than the one above could be applied: if the metaphor is meaningful either in the construction of the piece, to its discourse, or it is relevant in terms of how the user experiences it, then the metaphor is relevant. However, while in terms of interactivity, provided the concept is well defined, the relevance of the use of complex processes can be judged with some degree of objectivity, in terms of the use of life metaphors the discussion necessarily falls into a more subjective realm.

Finally, in some of these system claims of emergent behavior are articulated. These claims are often vague (e.g. with the case of Autopoiesis) and it is not clear what the artists mean by emergence in a particular case, nor how their work is an example of such behavior. The following chapter presents a framework of analysis that aims to overcome this problem.

8. EMERGENCE

This chapter summarizes the proposed approach to the concept of emergence. It starts by addressing some of the concepts that can be understood as oppositional notions to how emergence is here understood. After that, it reviews the proposal of Ernest Nagel, which coincides to that of this dissertation in the sense that it defends that emergence does in fact refer to two different notions. Finally, these two notions (self-organization and generation of novelty) and their implications regarding the idea of emergent interactive behavior are reviewed.

8.1 What Emergence is not

Complex concepts such as emergence are often described or delimitated using negative notions. Stating what emergence is not helps understanding what the limits of the idea are. In the particular case of emergence, the notion of reducibility is perhaps the main oppositional concept, but there are other concepts worth considering in this respect. What follows is a review of some of these oppositional ideas.

As said, what is emergent is usually presented as not being reducible in terms of classical cause-effect links. The basic idea of classic reductionism is that laws and effects at higher levels of complexity are entirely explainable by analyzing more elemental laws and effects. In systemic terms this means that any system is always and completely explainable by analyzing its components and how they are interfaced to each other. In turn, these components can be explained too by breaking them down into smaller subsystems, and so on. Thus, if emergence is not reducible, this means that this scheme does not apply to emergent phenomena: at some point, the black boxes that compose the system cannot be opened. This is precisely the popular notion of the whole being more than the sum of its parts. When emergence appears, it is not enough to break down the system and look into the components and to how they relate to each other. How this is resolved, as it has been shown, varies greatly from author to author, depending on what each one understands emergence exactly. Some of the following sections are dedicated to clarify the proposal of this thesis in this respect.

Secondly, emergence must be differentiated from mere appearance. Semantically, the word emergence does have this meaning too: ‘the emergence of the Chinese economy,’ ‘the emerging field of nanotechnology,’ ‘the fox emerged from beyond the bushes’… all these expressions are correct. But they don’t use the word emergence with the same meaning as the discussions reviewed in this thesis. Neither the Chinese economy, nor the field of nanotechnology or the fox are emergent in that sense. In these cases, the distinction is trivial and the confusion does not go beyond polysemy; here emergence is nothing but a synonym of appearance or ‘rise of’. The problem is when the term is used as ‘the’ concept of emergence, but it is actually linked only to the appearing of things. As seen in the previous chapter, this is the case of (Seevinck and Edmonds, 2009). The differentiation is important, as noted by Baljko and Tenhaff in (2006): “‘emergence’ is something more than the mere ‘emersion’ or ‘appearance’ of behaviors/properties.” This relates, too, to the notion of openness, which is discussed below.

Third, an emergent phenomenon needs to be differentiated from an epiphenomenon. The latter is an effect that appears along with a process but that is not directly related to it in terms of causality. Again, this notion is very similar to that of emergence. Some fine conceptual considerations need to be elaborated in order to distinguish them properly. For Jaegwon Kim, for instance, what differentiates an emergent phenomenon from an epiphenomenon is the absence of downward causation in the latter (Kim, 1999)[69]. According to him, there is always a causal base that supervenes on the phenomena, but only in the case of emergence this phenomena exerts back an influence onto this causal base, while epiphenomena are irrelevant to this base. As discussed in chapter 1, Sensitive Dependence on Initial Conditions (SDIC) is an example of epiphenomena. A popular formulation of SDIC is the butterfly effect: a butterfly flapping its wings in Brazil can set, by a long chain of cause-effect relations, a tornado in Texas. More technically, a change in the initial conditions of a weather forecast model (e.g. using 4 decimals instead of 5) can lead to very different outcome in a particular prediction. The context is that of deterministic nonlinear systems. Here the initial conditions are indeed key to the further development of events: the flap causes the air to move, and this current creates a small air flow that causes something else… and a few weeks later the tornado happens; the numerical decision as the prediction is being prepared causes differences large enough in the calculi of events as to generate widely different predictions. But the tornado or the weather prediction do not have any meaningful effect back on the causal base, on the initial conditions. Unless by another unlikely chain of events, the butterfly will not receive any effect back from the tornado in Texas.

Finally, there is a notion that does not usually appear in the discussion on emergence, but that can be useful to clarify what the concept means: intentionality. The idea is that emergent phenomena are necessarily unintentional, in the sense that the way in which they actually appear is not pre-specified by anyone either inside or outside the system where they happen. This does not exclude the possibility for someone to set up the conditions for a system to exhibit such phenomena (otherwise the whole purpose of this thesis would be rendered useless). One can expect emergence to happen; prepare a system so that certain conditions are met and wait. But the idea is that how the emergent phenomenon actually develops is the result of this irreducible, not epiphenomenal process that is not driven by any specific agent inside or outside the system. In this respect, this absence of intention is a necessary (although not sufficient) condition for self-organization. From the point of view of novelty, similarly the appearance of whatever is novel in the system is only emergent if it is not intentionally produced. Otherwise it would be (simply) a product of design. An emergent novel feature in a designed system appears as the system develops its designed behaviors. It is necessarily unintended; an unexpected –maybe hoped for, but not intended in the particular way it appears– new feature or behavior of the system that has been set in motion. This is precisely were the apparent paradox of de design for emergence is resolved. The emergent phenomenon cannot be designed: only the conditions for it to appear can.

---

[69] See section 2.4.2.

8.2 Emergence, Openness and Game

In relation to the previous subsection, there is yet another concept that needs to be differentiated from emergence: openness. The fact that emergence can be surprising and that it is necessarily linked to some degree of complexity has been the source of some formulations that, arguably, diminish the relevance of the concept by essentially identifying it with openness. Such an example was discussed in the previous chapter, in relation to the work of Seevinck and Edmonds (2008; 2009). In a similar fashion, some authors have used the concept of emergence in the context of game theory to discuss the idea of open vs. closed games (Juul, 2002; Salen and Zimmerman, 2004; Adams and Dormans, 2012).

Jesper Juul differentiates between games of emergence and games of progression. He bases this idea in Harvey Smith's use of the term emergence in (Smith, 2001) to talk about gameplay situations that were unexpected to the game designers. Juul's distinction is used to characterize two different ways of presenting the player of a game with a challenge. The first is that of a game that develops from a relatively simple set of rules. These games are replayable, and lead to a considerable variety of possible outcomes. Card games, board games or strategy games fall into this category. These games, which are in fact almost all traditional games, are mainly based on the game structure that Juul labels 'emergence'. On the other hand, some newer computer-based games are based on the structure of 'progression:' a structure within which the player has to perform a series of predefined actions in order to advance. One of the differences between these two approaches to games is that, while the first foster competition and strategy, those of the second type lead to step-by-step guides that help the user advance: "as a rule of thumb, the simplest way to tell games of emergence from games of progression is to find guides for them on the net. Progression games have walkthroughs: lists of actions to perform to complete the game. Emergence games have strategy guides: rules of thumb, general tricks" (Juul, 2002). According to the author, most pre-electronic games are games of emergence, while videogames combine both structures or favor progression.

The distinction is indeed very useful in order to analyze videogames. The paradigm of progression games is adventure games. The range of examples goes from the text-based adventures from the 1980s like Infocom's Zork, or the graphic adventure's of games like Maniac Mansion, to contemporary videogames like the Grand Theft Auto sage. The advances in such games happen through finding objects, discovering doors, resolving puzzles, etc. Each of these actions, or a group of them, is always a pre-requisite to move forward in the game. Thus, the gameplay is in this sense closed: a series of predefined actions must be made, always in the same manner and in the same order. As Juul notes, replaying these games is not usually appealing for a player, because, once the progression mechanisms of each stage are discovered, the game offers nothing new to the player. In contrast, what Juul refers to as emergence games are games that offer an open structure, in the sense that the gameplay can develop in very different ways, and different strategies can be created in order to try to succeed in the game. Open and closed structures for characterizing games are very well defined within these parameters.

Following Juul, Ernest Adams and Joris Dormans base their analysis of game mechanics on the distinction between mechanics that relate to emergence and mechanics of progression (Adams and Dormans, 2012). Similarly, Katie Salen and Eric Zimmerman present the idea of emergence as a key concept in their understanding of games. For instance, they explain how meaningful play (the fact that an action within a game is relevant in terms of the how it affects the development of the play session) is emergent respect the relation between player action and system outcome (Salen and Zimmerman, 2004). All these authors use Game of Life as a primary example of how simple rules can generate complex and, according to them, emergent results.

These authors use the idea of emergence along the lines of some ALife and ALife Art authors discussed in previous chapters. Juul, for instance, proposes to use emergence to explain how the variation of a game is “a non-obvious consequence of the rules of the game” (Juul, 2002). Whatever is emergent (these game situations and the strategies that players develop in order to win) is “neither anticipated by the game designer, nor is easily derivable from the rules of the game.” As said above, Adams and Dormans (2012) use Juul’s categories as the basis to elaborate their theory of games. Like Salen and Zimmerman (2004) they link emergence to unpredictability. The basis of their argumentation is that games, even when made out of simple rules, generate systems that are complex. And as complex systems they produce unanticipated results that they identify as emergent. The parts –the rules and mechanics of the game– generate a complex whole –the particular gameplay– which is emergent in respect to these parts: “the rules of Pong are relatively simple, but if you imagine all of the ways that a game can play out, from a quick-win match where one player dominates, to an extended, dramatic finish, it is clear that the system of Pong demonstrates emergence” (Salen and Zimmerman, 2004).

This resonates with the most popular –and most ambiguous– characterizations of emergence, and many of the arguments that have been proposed here in relation to those can be made in relation to these game theories. The first important point is that these authors basically link emergence to the impossibility of anticipation, to surprise. Thus, it is an approach that relates to emergence as generation of novelty, very much along the lines of Holland’s importance on the idea of emergence as being linked to the notion of perpetual novelty. However, this relates to the subjective formulations of emergence of Ronald et al. (1999a; 1999b) and Monro (2009) reviewed in chapter 5: it remains unclear in respect to what exactly these gameplay experiences are emergent (i.e. novel). To whom is it surprising that a particular Pong game can be very long? Or (as Juul presents as an example) that the Monopoly game often results in bankruptcy of one of the players despite that no rule states that this will happen?

More so, the fact that these are emergent outcomes of the game experiences conflicts with the fact that, in game design, a basic premise is that games are created through iterative design: that is, games are though of, prototyped and tested, and then they are revised and tested again, and so on until the final product is released. Thus, encountering these particular outcomes is part of the creation process, and while they can surprise the designers on the first appearance, they are incorporated into the space of possibilities afterwards. Therefore, at most, this would be an instance of emergence-relative-to-a-model, where the model is the expected outcome of the designer. Once something unexpected appears, it is incorporated into the model and as a result it ceases to be unexpected.

But this is not how these authors present emergence. Rather, for them it is the fact that the space of possibilities is open what makes this very space of possibilities into a source of emergent phenomena. The rules can create complexity-related unexpected outcomes that are, according to them, always emergent. And they are emergent because they are related to complexity. The problem with that formulation is that, if we accept this, then anything that is not predictable in detail is emergent: a gameplay, but also someone's life path or any conversation. Anything that has an open space of possibilities would be emergent according to this. Essentially, what they do is to identify emergence and openness, and this invalidates the particularities of what is emergent in respect to what is open. If instead of emergent and progressive the terms used were open and progressive (or closed), Juul's theory would still be as valid for the characterization of games. Then, perhaps, emergence could be presented as a third alternative. Instead of talking about 'the open and the closed', as Juul does in his 2002 article, we might be able to discuss 'the open, the closed, and the emergent'.

As said above, at the basis of these characterizations of emergence as a key concept to understand the wide variety of situations that can appear in a particular play session of a game is related to John Holland's theories. Particularly, to his account of emergent phenomena as related to the game of checkers (Holland, 1998). In this monograph on emergence, Holland presents his theories to a non-technical audience. As discussed in previous chapters, Holland understands emergence as both a self-organizing process (agents interacting among them and generating phenomena that can only be explained if this interactions –not the isolated agents solely– are taken into account) and as related to novelty, as a generator of 'perpetual novelty' in the world. It is in order to illustrate the second aspect, novelty, that Holland uses the game of checkers to illustrate emergent novelty.

The references to Holland are used by these game theorists to illustrate how emergence is important to explain how a simple set of game rules can generate divers outcomes that are 'unexpected' regarding these rules (Juul, 2002; Adams and Dormans, 2012) and how this relates to novelty and surprise and to meaningful play (Salen and Zimmerman, 2004). However, Holland does not refer to how emergence generates novelty in the passages they cite, but on how it relates to self-organization. When Holland mentions the game of chess in (1998: 14) or how the interactions among agents have to be taken into account in order to explain emergence in (1998: 122) he is referring to the idea that self-organization is not just about analyzing the attributes of the agents of a complex system in isolation, but on how relate to each other. When he talks about how a particular situation on the chess game does not inform us of the progress of the game, he is referring to that. The system must be observed in motion in order to have a proper understanding of how the game is developing. According to Holland, this is a simple representation of the idea of emergence: local actions developing interactions through time that can be observed as patterns of organization. The same would apply to the understanding of Game of Life as a simple model of self-organization.

Holland does address emergence as novelty in the book, and he does so by using the game of checkers as the main driving example. But here he is not defending, as the game theorists do, that the possible outcomes of a particular play of chess are emergent in this sense. They are open, of course, but when Holland analyses checkers he is in fact studying the emergent properties, not of the game itself, but of a particular AI checkers player developed by Art Samuel in the 1950s (Samuel, 1959; 1967). Holland clearly states that he is interested in the mechanism by which Samuel's player learns, and how this learning is related to emergence as novelty. The basis of his argumentation is that the checkerplayer's play strategies are emergent because they were not explicitly designed by Samuel. The player accumulated the results of different moves in different game settings regarding the final outcome of the game. Samuel used different weights to account for different features on the game, and the player learned how its moves might lead to a gain or loss through this weighing system. From this accumulation, different strategies were observed by the researchers, which had not been explicitly designed into the system: use of subgoals, anticipation of the opponent, and so on. Such strategies are emergent according to Holland. They are, in fact, coherent with the idea of emergene-relative-to-a-model. What is novel here is what the external observer identifies as a pattern of behavior that was not incorporated in the initial model of the system.

In conclusion, the use of emergence in game theory, as presented here, does not really move beyond the basic understanding of emergence as a consequence of openness in a relatively complex system. While this is an interesting concept to account for the wide range of possibilities that derive from something relatively simple such as the rules of a game, the argument here is that using the notion of emergence is unnecessary, and that it leads to two negative effects: first, to confusion by basically equating emergence and openness. Second, it precludes a useful theory of emergence to be developed in the parameters of game theory. Thus, as said above, a more fruitful approach would be to separate closed, open and emergent game structures, instead of conflating the second and third into only one.

8.3 Ernest Nagel: The Two Doctrines of Emergence

Ernest Nagel was an important figure in the Philosophy of Science of the twentieth century. His book *The Structure of Science*, published in 1961, remains a landmark in this discipline and one of the most important works on scientific reductionism (Van Riel and Van Gulick, 2014). He also has been identified as one of the critics of emergence in the sense that the British Emergentist's used the term in discussing evolutionary theory (McLaughlin, 2008; Whitelaw, 2006).

Nagel writes about emergence in the context of theorizing on the reduction of scientific theories, in the sense of theories or sets of experimental laws being explained by others from another domain. Within this discussion, Nagel presents and analyzes what he labels as 'the two doctrines of emergence'. That is, according to him emergence is related to two different theoretical domains. These two domains basically coincide with the two proposed in this thesis. The two ways in which emergence is understood, according to Nagel, are emergence as irreducible hierarchical organization (the equivalent to emergence as self-organization) and 'the evolutionary version of emergence' (emergence as generation of novelty): "the doctrine of emergence is sometimes formulated as a thesis about the hierarchical organization of things and processes, and the consequent occurrence of properties at 'higher' levels of organization which are not predictable from properties found at 'lower' levels. On the other hand, the doctrine is sometimes stated as part of an evolutionary cosmogony, according to which the simpler properties and forms of organization already in existence make contributions to the 'creative advance' of nature by giving birth to more complex and 'irreducibly novel' traits and structures. … This evolutionary version of the emergence doctrine is not entailed by the conception of emergence as irreducible hierarchical organization, and the two forms of the doctrine must be distinguished" (Nagel, 1961: 366-367).

Thus, Nagel proposes to distinguish between the emergence of the 'whole being more than the sum of the parts' from the emergence as a concept related to novelty. He relates the first to the hierarchical organization of nature, as said above, and also to the idea of the 'nonpredictability of certain characteristics of things'. Nagel describes this type of emergence as follows: first, there is a series of elements standing in interaction that are the necessary and sufficient condition for an object (a whole) that is characterized by a series of properties. Second, complete knowledge of the elements in isolation (and of other wholes they might form and the properties that these might have) is assumed. If, despite this knowledge, there is at least one property of this whole that is impossible to predict, then the property is emergent and, consequently, the whole is emergent too.

Nagel, who is writing in the context of the analysis of theories, situates these emergent wholes and properties just mentioned in the realm of epistemology. This means that it is in the scientific discourse, in the formulation of theories, that the idea of emergence is relevant, not in reality itself in ontological terms: "it is not properties but statements (or propositions) which can be deduced. Moreover, statements about properties of complex wholes can be deduced from statements about their constituents only if the premises contain a suitable theory concerning these constituents-one which makes it possible to analyze the behavior of such wholes as ‘resultants’ of the assumed behaviors of the constituents. Accordingly, all descriptive expressions occurring in a statement that is allegedly deducible from the theory must also occur among the expressions used to formulate the theory or the assumptions adjoined to the theory when it is applied to specialized circumstances, Thus a statement like ‘Water is translucent’ cannot indeed be deduced from any set of statements about hydrogen and oxygen which do not contain the expressions ‘water’ and ‘translucent;’ but this impossibility derives entirely from purely formal considerations and is relative to the special set of statements adopted as premises in the case under consideration. … The doctrine of emergence (in the sense now under discussion) must be understood as stating certain logical facts about formal relations between statements rather than any experimental or even ‘metaphysical’ facts about some allegedly ‘inherent’ traits of properties of objects” (Nagel, 1961: 368-369).

Thus, says Nagel, nothing is emergent in absolute terms: no given property is inherently or absolutely an emergent trait. But this does not mean that emergence is simply a result of ignorance. The assertion that this type of emergence is only the result of a lack of knwoledge, that if we were to really fully know all the properties of the elements we would be able to do any prediction, is mistaken: “while it is an error to claim that a given property is ‘inherently’ or ‘absolutely’ an emergent trait, it is equally an error to maintain that in characterizing a trait as an emergent we are only baptizing our ignorance” (Nagel, 1961: 371).

The key to understanding this type of emergence is, according to him, the notion of ‘prediction:’ “the doctrine [of emergence] employs the phrase ‘to predict’ in the sense of ‘to deduce with strict logical rigor’” (Nagel, 1961: 371). This distinction is crucial, and it can help understand, if the argument is accepted, one of the most controversial aspects in the discussions on emergence. The idea is that emergence as self-organization (or as hierarchical organization) is not about prediction in the sense of being able to anticipate what is going to happen, but about deduction in respect to the set of statements under consideration. This is reminiscent of Cariani’s emergence-relative-to-a-model, in the sense that emergence is always emergence in relation to something previously (and clearly) stated[70]. Furthermore, this is perfectly consistent with the fact that prediction is indeed possible when other statements (or theories) are added to be taken into consideration. That is, when the model incorporates new aspects.

This leads to another crucial aspect of Nagel’s theory: according to this, the fact that the result has been already observed does not invalidate it as a case of emergence. This is

---

[70] Mitchell Whitelaw has noted that Nagel is the first to introduce emergence as an epistemological problem, and that Cariani’s work connects with this idea (Whitelaw, 2006).

perhaps the biggest problem with Cariani's ERTM. If emergence is always about novelty relative to something, whatever is novel ceases to be emergent once observed and incorporated into the model that serves as the observing framework. The problem with this is that it invalidates as emergent basically all of the examples that relate to the idea of self-organization which, although in different particular configurations, tend to repeat observable patterns. With Nagel's differentiation between predictability and logical deduction, emergence can still be valid as a construct in this type of phenomena, That is, the actual observation does not invalidate the fact that is wasn't logically deducible before the observation. This is what allows self-organization emergent phenomena maintain their status even once they have been observed.

Despite the differences, the 'relative-to' is as crucial in Nagel as it is in Cariani. However, while Cariani formulates his theory with the comparison of observations and abstract models in mind, Nagel bases his argumentation in the logical arena: "[the issue is] whether a given statement is deducible from a 'given' set of statements, and not whether the statement is deducible from some 'other' set of statements. As we have already seen, when we are said to improve or enlarge our knowledge concerning the 'nature of H and O,' what we are doing in effect is replacing one theory about H and O with another theory; and the fact that H and O combine to form water can be deduced from the second theory does not contradict the fact that the statement cannot be deduced from the initial set of premises" (Nagel, 1961: 371). All of this, for Nagel, constitutes the admittance of "the essential correctness of the doctrine of emergence when construed as a thesis concerning the 'logical relation' between certain statements" (Nagel, 1961: 372). Hence Whitelaw's claim that Nagel presents emergence as "a kind of logical truism" (Whitelaw, 2006: 210).

In any case, Nagel insists in that this type of emergence is, in fact, even more powerful than the proponents of the doctrine claim. The usual assumption, he says, is that emergence often happens when higher level properties are not accounted for by looking at lower level elements (separation of chemistry and physics, for instance). The typical counterexample of emergence is the behavior of a clock, which is supposedly predictable from the knowledge of the properties and organization of its constituent cogs and springs. But, while this is not untrue, mechanical theory only accounts for some behaviors of the clock. It does not explain the changes in temperature of the clock or the magnetic forces that the relative motions of the parts of the clock might generate: "It appears that nothing but arbitrary custom stands in the way of calling these 'nonmechanical' features of the clock's behavior 'emergent properties' relative to mechanics. On the other hand, such nonmechanical features are certainly explicable with the help of theories of heat and magnetism, so that, relative to a wider class of theoretical assumptions, the clock may display no emergent traits" (Nagel, 1961: 373).

Therefore, even in this first formulation, emergence is always so relative to a theory or a series of statements. Nothing is inherently or absolutely emergent. It is always emergent from a certain point of view and always in an epistemological sense. In other words, what is emergent is emergent relative to a model.

The second 'doctrine of emergence' that Nagel identifies is that of emergence as an evolutionary process or 'as a temporal, cosmogonic process' (about how things come into being). Emergence as novelty is here linked basically to emergent evolution: "the doctrine of emergence as an evolutionary cosmogony, whose primary stress is upon the alleged 'novelty' of emergent qualities. The doctrine of emergent evolution thus maintains that the variety of individuals and their properties that existed in the past or occur in the present is not complete, and that qualities, structures, and modes of behavior come into existence from time to time the like of which has never been previously manifested anywhere in the universe" (Nagel, 1961: 374).

This second type of emergence is to be understood as separated from the previous one, since none of the two conceptions, according to Nagel, entails the other. This means that emergence is not necessarily linked to novelty according to the first definition: "it may very well be the case that a property is an emergent relative to a given theory but is not novel in a 'temporal' sense" (Nagel, 1961: 375). While the first one is mostly concerned with the logical relations of statements, the second can only be resolved by empirical inquiry. Therefore, emergence as novelty is "a straightforward empirical problem."

However, the problem is, here: how do we know that it is actually novel? And, more so, what kind of changes can afford novelty? When is a pattern, a property or a process novel? In relation to this, Nagel states: "in general, whether or not an attribute is to be regarded as an 'essential' one depends on the context of the question and on the problem under consideration. But if this is so, then in view of that stipulation the distinction between an emergent trait and a nonemergent one would shift with changes in interest and with the purposes of an inquiry. These difficulties are not cited as being fatal to the doctrine of emergence. They do indicate, however, that, unless the doctrine is formulated with greater care than is customary, it can easily be construed as simply a truism" (Nagel, 1961: 377). To some degree, emergence-relative-to-a-model is a response to these problems. What is novel is so relative to a particular model, which is elaborated after observation, so newly observed behaviors and patterns can be accounted for in relation to it.

Finally, Nagel ends with some considerations on the type of knowledge that can lead us to judge something as new or not new that connect with Rosen's idea of emergence as the (almost necessary) deviation of the observed reality in regards to the model through which one is observing it: "it might of course be said that such novel types of dependence are not 'really novel' but are only the realizations of 'potentialities' that have always been present in 'the natures of things'; and it might also be said that, with 'sufficient knowledge' of these 'natures,' anyone having the requisite mathematical skills could predict the novelties in advance of their realization. We have already commented sufficiently on the latter part of this rejoinder, and can therefore discount it without further ado as both invalid and irrelevant. As for the first part of the objection, it must be admitted that it is irrefutable; but it will also be clear that what the objection asserts has no factual content, and that its irrefutability is that of a definitional truism" (Nagel, 1961: 380). The second part of the objection is what ERTM inherits from Rosen (and from Cybernetics): Any model is a necessary abstraction that takes into account only a limited set of elements, while others are left out. And what is left out can elaborate affect the functioning of what is being modeled.

In conclusion, Nagel differentiates two types of emergence that are equivalent to self-organization emergence and emergence as generation of novelty. Both formulations are relevant as epistemological constructs, and can be understood to be resolved with the approach of ERTM. The main difference is that while Cariani's theory takes observation (i.e. empirical history) into account, Nagel's resides in the logical realm.

8.4 Emergence as Self-Organization and Emergence as Novelty

The proposal to understand emergence as both self-organization and as generation of novelty finds its roots in Nagel[71], and relies strongly on Cariani's emergence-relative-to-a-model. The aim of this approach is to accommodate both versions of emergence in a way that can be used as a tool for analysis and, eventually, as the basis for the future guidelines of a creative strategy in interactive communication experiences.

A basic premise for understanding emergence as proposed here is that the approach is epistemological and pragmatic. Coincidently with Nagel and Cariani, emergence is understood as an epistemological construct. I.e. the focus is not on whether or not properties and behaviors identified as emergent are actually emergent in the ontological sense, but on how they are perceived, understood and communicated. It is also pragmatic, because the final aim of the proposed framework is to be used not only as an analytical tool but as the basis for the creation of devices that are capable of producing emergent phenomena and behavior as it is here understood. The following are some remarks on the basic ideas of the proposed framework.

8.4.1 Self-Organization: Homeostasis and Cybernetic Emergence.

Self-organization is the process by which a group of agents (particles, molecules, animals, robot parts…) generates, through local interactions, an observable pattern of organized behavior at the group level. This process is not directed or coordinated by any particular agent, neither by a specific sub-group (Soler-Adillon and Penny, 2014). 'Self-organization emergence' refers to the classic idea of emergence: phenomena that result from multiple interactions among agents within a system (the parts), in the form of patterns of behavior of the system as a whole. These patterns are emergent in the sense that they could not be understood, nor anticipated, through the analysis of the elements and their behaviors in isolation. In Nagel's words, they could not have been 'deduced with strict logical rigor' in advance.

Self-organization emergence is not necessarily related to novelty. The Cybernetics' emergence (usually referred to as self-organization) was concerned with how local interactions within systems might lead to systemic results that were not 'deduced with strict logical rigor' in advance, but only discovered through experimentation. The performative idiom of Cybernetics is all about these type of processes. The ultrastability of Ashby's Homeostat, the behaviors of Walters' tortoises, or the performances of Pask's Musicolour and Colloquy of Mobiles were all examples of this approach to knowledge.

---

[71] On a personal note, and with the understanding that this is irrelevant in academic terms, I found out about Nagel's distinction long after elaborating the idea of the distinction.

Embedded in this practice is the idea of homeostasis; the tendency of the system to reach and maintain equilibrium. This idea is central in Cybernetics. The parts of the system communicate with feedback loops that result in a systemic behavior. This behavior is based on all the parts maintaining their role in the equilibrium that allows the systemic behavior to persist –an an idea that connects to that of autopoiesis, when applied to living systems. Up to some degree of complexity, the analysis of the parts permits the informed investigator to foresee the system's behavior, but if the complexity increments this anticipation might cease to be possible (the system becomes exceedingly complex, to use Beer's words). In this respect, homeostasis is emergent if it can only be experimented, but not anticipated or 'deduced with strict logical rigor'.

Like other ALife related concepts, the interest in self-organization appeared in the 1960s within the context of both Artificial Intelligence and Cybernetics and reappeared in the 1980s and 1990s in the context of Complexity Sciences (Umpleby, 2009). Von Foerster (1960) and Ashby (1962) were amongst the first to theorize about the idea of self-organizing systems and their relation to the environment. The general idea was that, while these systems were informationally closed (the rules of interaction among agents do not change), they are open to the interactions with their environment (they are dynamic systems).

Self-organization, however, is not unanimously defined, and defining the proposed term 'self-organization emergence' poses the problem that many definitions of self-organization assume and use the term emergence to define it. For Kaufmann, for instance, self-organization is "the spontaneous emergence of order" (Kaufmann, 1993). Furthermore, authors such as Gershenson and Heylighen (2003) consider that an actually complete definition should first clarify "the philosophical problem of defining 'self,' the cybernetic problem of defining 'system,' and the universal problem of defining 'organization'."

In any case, the definitions usually stress the importance of the absence of a central control, of local interaction, and of the order that it is perceived (as a pattern) by an external observer. Self-organization is "the appearance of structure or pattern without an external agent imposing it" or "the spontaneous emergence of global coherence out of local interactions" (Heylighen, 2003). It is the global 'cooperation' of the elements of a dynamic system that spontaneously reaches an attractor state (i.e. it organizes around a certain control parameter) (Rocha, 1998). The importance of the observer connects the issue to epistemology. In this respect, Gershenson and Heylighen (2003) propose that "self-organizing systems, rather than a 'type' of systems, are a 'perspective' for studying, understanding, designing, controlling, and building systems. … The crucial factor is the observer, who has to describe the process at an appropriate level(s) and aspects, and to define the purpose of the system. All these 'make' the system to be self-organizing. In that sense, self-organization can be everywhere: it just needs to be observed."

Arguably, from the epistemological point of view self-organization and emergence are inseparable, and this is the position defended in this thesis with the notion of self-organization emergence. As said, an equivalent idea was developed by Ernest Nagel, and many authors theorizing on both ideas used them in this sense. In accordance to this, some explanations of self-organization explicitly conflate the idea to that of emergence: "[it is] the appearance of special or temporal structures with a significantly superior size or duration than that of its constituent parts. Self-organization is, thus, the dynamic 'explanation' of emergent properties"[72] (Solé, 2012: 49-50).

Authors such as Langton, Holland, Wolfram, Bedau or Kelso, reviewed in preceding chapters, elaborate on emergence as related to self-organization processes. Cellular automata (CA) are abstract models of such systems, in which the simplicity of the interacting parts allows the researchers to elaborate on what emergence is exactly. CA are used in this context by some of these authors as an example of how the local interactions produce what they consider to be emergent results. These abstract models are useful to illustrate this in the sense that it is very clear what happens at the local level, and thus how it relates to the possibility of anticipating what is observed as a pattern of behavior. CA are in this respect at the opposite end of the scope of the opaqueness of Ross Ashby's Homeostat, which consisted of several units the interior workings of which were largely unknown even to Ashby himself. The problem, in either case, is on defining what exactly is constitutive of emergence and what is not.

The main problem with this type of emergence, as noted by Nagel, is the notion of prediction. If emergence is linked to the impossibility of prediction, then the question is what happens once an emergent effect has been observed. Take as an example the flocking behavior of the boids or a glider in Game of Life. Are they emergent only the first time they are observed by anyone at all? Or are they emergent when observed by someone who had never observed them before? Does the observed effect cease to be emergent once observed? The answer is no: these authors do not consider emergence to be solely a matter of surprise. Bedau, for instance, considers that is emergent whatever cannot be anticipated and has to be found out through simulation (Bedau, 2008). The problem with this, as discussed in chapter 2, is that the line between when the need for simulation starts and ends cannot be clearly defined, unless this is done relative to a particular observer and stated in advance. Similarly, Holland considers that the emergent effects are only deducible if the local interactions among the agents that constitute the system are taken into account (Holland, 1998).

---

[72] The original text, in Spanish, is: "la aparición de estructuras espaciales o temporales cuyo tamaño o duración es muy superior al de los elementos constituyentes. La autoorganización es por lo tanto la 'explicación' dinámica de las propiedades emergentes."

A common factor in all these approaches it that emergent phenomena appear in complex systems, and that these phenomena have to be differentiated from other complex system-related phenomena, such as those related to indeterminacy, or the results of Sensitive Dependence on Initial Conditions. Emergent phenomena are not merely stochastic. However, there is always some degree of indeterminacy involved. When Kelso (1995) writes about emergent dynamic patterns that are driven by control parameters (attractors) he is describing this. There is no way to know how exactly the control parameter will affect the system and anticipate what motions in particular will appear (e.g. in which direction the molecules of boiling oil will roll). However, once observed, if the same conditions are presented again (what Jaegwon Kim calls the causal base) the emergent effect can be expected to reappear.

Debates on concepts such as downward causation and supervenience find their context here. While supervenience refers to this dependence of the causal base just described, downward causation is an idea that elaborates on how these emergent phenomena affect back the constituent parts at the causal base. In this respect the question of the predictability of a self-organizing phenomenon ceases to be problematic. If certain conditions are given, the phenomenon can be expected to present itself, with subtleties that will be impossible to anticipate. The boids will always flock if the simulation meets the conditions, although no exact path of the boids can be traced in advance. On the other hand, if predictability is about knowing whether or not the flocking will happen, or a structure like the glider, or the R-Pentomino can appear in Game of Life in the sense that the effect is novel, we must frame it within the parameters of emergence in relation to novelty, and question in respect to what exactly it is novel.

Downward causation is usually rooted in the ontological debate on emergence, and Bedau’s distinctions between strong and weak downward causation –that parallels strong/weak emergence– find its meaning in regards to this. A sort of ontologically discharged notion of downward causation was the idea of ‘channeling’ of Campbell and Bickhard (2001) discussed in chapter 2. It is a notion that explains how the attractor or control parameter driving the overall self-organized behavior influences the particular behavior of each of the agents in the system, while not the relations among them.

The definition proposed here aims at accommodating all the important aspects of the approaches to emergence as self-organization: superveience, downward causation (thought the concept of channeling), the absence of a central control (i.e. distributed decision-making), the importance of the observer and the idea of agent as a behaving component in a system:

Self-organization emergence is the appearance of a(n observed) pattern at the level of a system that is produced by the local interactions of the system’s agents. This pattern has an effect on the agents, by channeling their behavior.

As said, crucially these emergent patterns of behavior cannot be directed by any agent or group of agents within the system, nor from an outside entity. I.e. there can be no intentionality in the actual unfolding of the behavior. As elaborated in the following section, in order to identify this type of emergence it is necessary to define the system under analysis in terms of systemic level and parts (components) level, and the interactions among the components. Only then one can clearly state what is a system-level observed pattern and how it relates to the local interactions of the agents.

In the example of the boids, the systemic level is the group of simulated birds. Each of the birds is an agent (a behaving component) in the system, and the local interactions are the rules that each agent uses to decide its moves: proximity to others, separation and steering according to neighbors. The self-organization emergent pattern that is observed is the flocking or schooling of the group of agents, which moves strikingly like a real group of birds of fish. This observed pattern influences the whole group and, thus, in a way it influences back each of the agents that generated it. The issue of whether or not this is ontologically valid as a case of emergence, if this is strong or weak downward causation, or even if real birds and fish use a similar local navigation system in order to move around in their environments is beyond what the proposed approach aims at covering.

8.4.2 Novelty: Creativity and the Coming Into Being of Things

The second type of emergence is what Nagel identifies as the 'evolutionary version of emergence,' or emergence as generation of novelty. This is the type of emergence that relates to the emergent evolution of the the British emergentists or Bergson's becoming. It is also Cariani's emergence-relative-to-a-model or what Holland refers to as the 'perpetual novelty' that emergence is related to. However, while this first group of authors is concerned with the appearance of novelty in natural systems, Holland and Cariani are concerned with how novelty can appear in artificial computer-related (or cybernetic) systems.

This connects with the question of whether or not artificial systems can be creative. However, while emergence understood in this sense can serve as the basis to elaborate a theory on how novelty appears, creativity is a wider concept that encloses novelty among other things, depending on the approach. Margaret Boden, for instance, has described novelty as involving ideas and artifacts that are that are not only new but also surprising and valuable (Boden, 2010: cited in McCormack, 2012). According to Schmidhuber, a creative agent is "one that never stops generating non-trivial, novel, and surprising behaviours and data" and that is capable of learing (Schmidhuber, 2012). In general, creativity is associated with the human individual, but systems, nature or society can be regarded as inventive too (Perkins, 1996; McCormack, 2012). Some authors are more skeptical towards the very concept of creativity, and consider it to be a discursive artifact related to particular activities, but not an idea that can be abstracted (Nake, 2012). Oliver Brown distinguishes between generative and adaptive creativity, in a fashion that resonates with Cariani’s creative and combinatoric emergence, although he explains it in terms of the immediate perceived value of the creative effort (adaptive creativity) o the absence of such immediately evident value (generative creativity) (Brown, 2012). Dorin and Korb define computer creativity as “the introduction and use of a framework that has a relatively high probability of producing representations of patterns that can arise only with a smaller probability in previously existing frameworks” (Dorin and Korb, 2012).

The proposal here is to understand how emergence can be the basis to understand the appearance of novelty as produced by artificial systems in the context of interactive art. As said, the approach is strongly based in Cariani’s emergence-relative-to-a-model, as it consists in defining a framework in order to analyze artistic interactive systems and, then, observe if novelty appears. Thus, the main steps consist in identifying the system under analysis and specifying the primitives with which it operates and the behaviors that are observed. These particular models are the elements relative to which novel behaviors will either appear or not.

8.5 Emergent Interactive Behavior

Once the framework has been presented and applied to a few examples, it is useful to recall one of the main driving idea of this thesis and see how the presented analysis model can help in building the way towards its achievement: the idea of emergent interactive behavior.

The goal of explicitly combining the ideas of emergence and interactive art originated in parallel with the appearance of Complexity Sciences in the late 1980s, and, in particular, of Artificial Life. Emergent Interactive Behavior was proposed as a paradigm that aims to create systems that can behave beyond what is explicitly specified in their design: "Emergent interactive behavior would not be derived from a set or pre-determined alternatives. Rather, behaviors might arise through a contingent and unconnected chain of triggers" (Penny, 1996). The basic idea is that this type of systems should exhibit behaviors that surprise their own designer. This element of surprise is not desirable in most cases when dealing with interactive devices. Functional interactivity is what one expects to find when using a word processor or a database interface. But in an artistic context interactivity becomes an idea to experiment with, and notions such as emergent interactive behavior or evolved behaviors became central for some of the ALife Art practitioners.

First, it is important to distinguish between the idea of interactive behavior that is emergent and the idea of combining emergent behavior and interactivity. There are cases of systems and devices that are interactive and that exhibit emergent behaviors, but what is emergent does not affect the way in which they interact with their users. Arguably, such an example is Sympathetic Sentience. The overall sonic result of the installation is, as shown above, emergent in the sense of self-organization. But this does not change the way in which the installation relates to its users in terms of interactivity. That is, whilst the work does exhibit interactive behavior, the particular behavior related to interactivity is not emergent.

Strictly speaking, emergent interactive behavior can be understood as a more strict term if it refers exclusively to behaviors that are interactive (i.e. that constitute how the piece relates to the user) and that are emergent. In this respect, the closest examples reviewed up to this point are the cases or combinatoric emergence of Performative Ecologies and Digital Babylon, which though the use of genetic algorithms modulate the actions that the systems do in relation to the user (being more or less prone to approaching him or her in the latter; performing one type of dances or another in the former). In this respect, the notion of Cumulative Interaction, which as stated above relates to learning and adaptive systems, represents an approach to creating this type of systems.

Finally, it is interesting to remind here the recent work of Peter Cariani and the creation of new concepts in minds that was reviewed in section 5.6. In a way, this approach implies a displacement of the point of view of analysis of emergence-relative-to-a-model from the cybernetic device to the human brain, but it is still relevant in terms of interactivity because the idea focuses on how computers, which are bounded and closed systems, can facilitate the creation of new concepts in human minds, and at the same time humans can expand the state-sets of computers by introducing new primitives into the digital systems. This combination of human and machine, according to Cariani, is in fact an open-ended system capable of generating new primitives and, thus, creative emergent novelty: “enhancing human creativity using flexible human-machine interfaces and other ‘tools for creativity’ is a much more efficient route at present for generating open-ended novelty than attempting to build autonomous self-organizing systems that are creative in their own right” (Cariani, 2009b). Since the brain is essentially a combinatorically creative system, this introduction of newness is one of the possible generators of creative emergence. As said in chapter 5, the details of how this works, which are beyond the scope of this thesis, can be found in Cariani (2009b; 2012).

8.6 Concluding Remarks: Emergence and the Metaphors of Life

Understanding emergence as related to both self-organization and novelty offers a large enough scope to accommodate for most of the work done in ALife Art. Although there are remarkable exceptions, many ALife artists and researchers conceive their work around biological, ecological and evolutionary metaphors. Most of these approaches relate in one way or another to one of the forms of emergence presented here. The relation between evolution and generation of novelty emergence is clear. In fact, Nagel labeled this type of emergence ‘the evolutionary version’ of the term. From Bergson to Ray to Cariani, evolution and novelty are strongly related, and even synonymous when it comes to emergence. In addition, some of these biological and ecological metaphors are also linked to novelty regardless of the evolutionary paradigm, especially if they incorporate learning or adaptive mechanisms.

Regarding self-organization emergence, there are mainly two ways in which it is pursued in ALife projects. First, through the idea of the bottom-up approach to generating processes: Sympathetic Sentience and The Conversation are examples of this approach. The first puts together a series of sound emitting robots that end up generating a group piece, while the second is a modern implementation of the homeostat, a piece were all the parts strive to do its part in achieving and maintaining the overall equilibrium. The second approach is more directly biologically inspired. The Colloquy of Mobiles, with its males and females, or Autopoiesis, with its awkwardly looking robotic arms than nonetheless aim at being reminiscent of living beings, are examples of this. They approach the idea of self-organization with a more direct biological metaphor. Finally, all the simulated ecologies, for obvious reasons, reside in this paradigm, too. Depending on the case, these projects can contain behaviors that are the result of either case of emergence.

9. THE SOE/GNE FRAMEWORK

Along the preceding chapters, a series of points of view on emergence have been discussed and analyzed. Building on this review, this chapter proposes a framework of analysis that is based on the notion that emergence refers to the two distinct ideas discussed in the previous chapter: self-organization and the appearance of novelty. After explaining this conceptual base and the framework itself, a series of interactive ALife Art pieces will be analyzed with it in order to test it.

9.1 Proposed Analytical Framework: The SOE/GNE Framework

The SOE/GNE Framework is an analytical framework conceived to be used as a tool for analysis of the presence of emergent phenomena in interactive artwork. It is based on the understanding that emergence is related to both self-organization and generation of novelty. The terms self-organization emergence (SOE) and generation of novelty emergence (GNE) refer to this idea. The distinction, as said above, finds it theoretical roots in Ernets Nagel's *The Structure of Science*. While SOE is an idea that connects to a wide variety of authors, from Cybernetics to ALife Art, GNE is mostly based on Peter Cariani's emergence-relative-to-a-model and the distinction between creative and combinatoric emergence. Research based on the distinction between emergence as self-organization and as generation of novelty has been presented in previous works, such as Soler-Adillon et al. (2014) and Soler-Adillon and Penny (2014). In addition to this, the understanding of the basic parts of a system is based on Herbert Simon's distinction among system, components and interfaces (Simon, 1996), and the particular model of each analyzed system will strive for the three modeling principles presented in (Walter, 1963): models shall be simple, transparent, and semantically brittle[73].

Methodologically, the way in which the framework is presented and how the analysis is implemented is partially based on the systems analysis proposals in the field of information theory (e.g. Kendall and Kendall, 2006) and particularly on the approach based in defining three basic aspects: concepts, indicators and analysis procedures (Codina, 2000). This approach has been used in the 'interactive decoupage,' a tool for reading and analyzing audiovisual interactive media (Freixa, 2009; Freixa et al., 2014a; 2014b; 2014c).

According to this, the framework is presented here in three steps: concepts, procedures and phases. The first is the presentation of the conceptual apparatus onto which the framework is built in. The second introduces the tools that are to be used in the analysis; it is the manner in which it is actually implemented. The third makes explicit how these procedures should be applied in order to achieve the desired analytical results.

---

[73] See section 4.2.2.

9.2 SOE/GNE Framework: Concepts

In defining the conceptual apparatus of the framework, it is necessary to clarify what is understood by each of the fundamental concepts that are in use in it. This section presents these conceptual explanations, along with a generic diagram that will be the basis of the modeling procedure, as reviewed in the section to follow.

9.2.1 Fundamental Concepts

- System (and environment)

  The system here is the interactive artwork. In its description, it is important to address how it relates to its environment through sensors and effectors. This is what Cariani calls the semantic axis. From this point of view, in the case of interactive systems the user or interactor is a part of the environment. Thus, the system is the device or group of devices that affords interaction. The system operates with primitives in order to perform calculations and make decisions on how to act on the environment. This calculation and decision making is Cariani's syntactic axis. Finally, the pragmatic axis is the evaluation of these actions in terms of achievement of goals. The result of these evaluations may result in changes in the calculations and decision making (combinatoric emergence) and even on the sensors and effectors themselves in the case of creative emergence. However, pragmatics is arguably only relevant in adaptive and learning systems.

  An important addition here in respect to Cariani's theorization (which is made with robotic cybernetic devices in mind) is that the interactive systems that are of interest here can have an important part of elements that are virtual. Here 'virtual' is understood in the sense of non-physically instantiated, or simply put: digital. Screen-based works, for instance, have a virtual environment that has to be taken into account in parallel with the physical environment of the work. The virtual environment is part of the system, but it is relevant here to address the fact that it will often present virtual agents or components, with their corresponding digital sensors and effectors (through which they interact among each other or with either the virtual or external environment).

  The importance of this point is that, if virtual sensors, effectors and environment are accepted, then creative emergence can happen within the virtual (digital) part of the system: a digital agent algorithmically developing a new way of reading some feature of another digital agent would be a case of creative emergence. Although this is not something that will easily appear, it is arguably more likely to occur than the actual creation of a new physical sensor of effector by an artificial agent. Combinatoric emergence, too, can appear within the digital domain, for instance by the hand of genetic programming.

Another important consideration here is that the system can include groups of agents that perform some or all of the operations regarding one of the components, and this is precisely one of the bases of self-organization emergence. E.g. system's sensors are distributed among a series of agents (Sympathetic Sentience) or the calculations and decisions might only be performed by individual agents, and not by a central brain (Game of Life; Digital Babylon).

- Sensors and effectors

  Sensors and effectors are the interfaces between system and environment. In the case of interactive systems they include the points of contact between system and interactor. Occasionally interfaces among the different components of the system, whether they are digital or physically instantiated, might be relevant for the analysis.

- Primitives

  The primitives are the basic informational building blocks with witch the system operates: the units it reads with the sensors and that are combined in the calculations. They are at the same time the basis of the decision making process. The proximity of a user, the fact that he or she pressed a button, the state of a neighboring cell… all of these can be primitives in different systems. In emergence-relative-to-a-model the recombination of these primitives can lead to combinatoric emergence, while the introduction of new primitives is what affords creative emergence.

- Computation and decision making

  This is the algorithmic part of the system. The primitives are combined in order to determine the state of the system, and decisions on how to act are made according to this.

- Evaluation of performed actions

  This is relevant in learning and adaptive systems: performed actions are evaluated vis-à-vis system's goals. Actions that were successful in getting the system closer to its goals will be fostered, while those that result in the contrary will not. Eventually, this evaluation might lead to a reconfiguration of how the calculations and decisions are made, and even on how the system perceives and acts on its environment (or how the virtual agent perceives and acts on its virtual environment).

- Behaviors

  As said, behavior is here understood epistemologically. It is an observed pattern that refers to a consistent series of actions and responses that the system exhibits when interacting with its environment. The behavior can be a system behavior or an agent behavior. In fact, it is the very notion of behavior what distinguishes and agent from a passive component in a system. When something within the system makes decisions and actions on its own (i.e. it behaves), then it is an agent. When something rests passively and simply passes along information or makes predefined calculations, it is a non-behaving component. Thus, agents are the behaving components in a system.

- Self-Organization Emergence (SOE)

  SOE is a (logical) 'disconnect' between what is described in Phase 1 and what is observed in Phase 2: a behavior produced through the accumulation of agent behaviors that was not explicitly programmed. In turn, this effect will exert some sort of influence on the agents, by channeling their behavior.

  The conditions for this type of emergence are thus: (1) the system has agents – behavioral components–, (2) these agents generate a behavior that is not directed by either one of them nor by an external entity (i.e. there is no central brain controlling the behavior) and (3) this behavior, once it has appeared, channels the individual actions of the agents.

- Generation of Novelty Emergence (GNE)

  GNE is the appearance of a new behavior or feature of the system after the system and its behaviors have been described. Phases 1 and 2 configure the model relative to which emergence will be evaluated. This emergent novelty is combinatoric when it is the result of new combinations of the primitives. When new primitives enter into the system (through the evolution of new sensors) then it is a case of creative emergence. As discussed in chapter 5, creative emergence in artificial systems is very rare.

9.2.2 The System Diagram

The description of the system will be accompanied by a system diagram. This diagram is the basis of the particular model that is built for each system, and when applied it represents the first of the procedures. What is presented here is a generic system with all the possible components incorporated. When analyzing each of the particular interactive artworks the diagram will be adapted ad-hoc at each time, eliminating the irrelevant parts, if any, and specifying which parts correspond to agents and system respectively whenever this is relevant. The generic diagram is an adaptation of Cariani's model as presented in (Cariani, 2012)[74], with the inclusion of the virtual parts (sensors, effectors and environment) for those systems who have them.

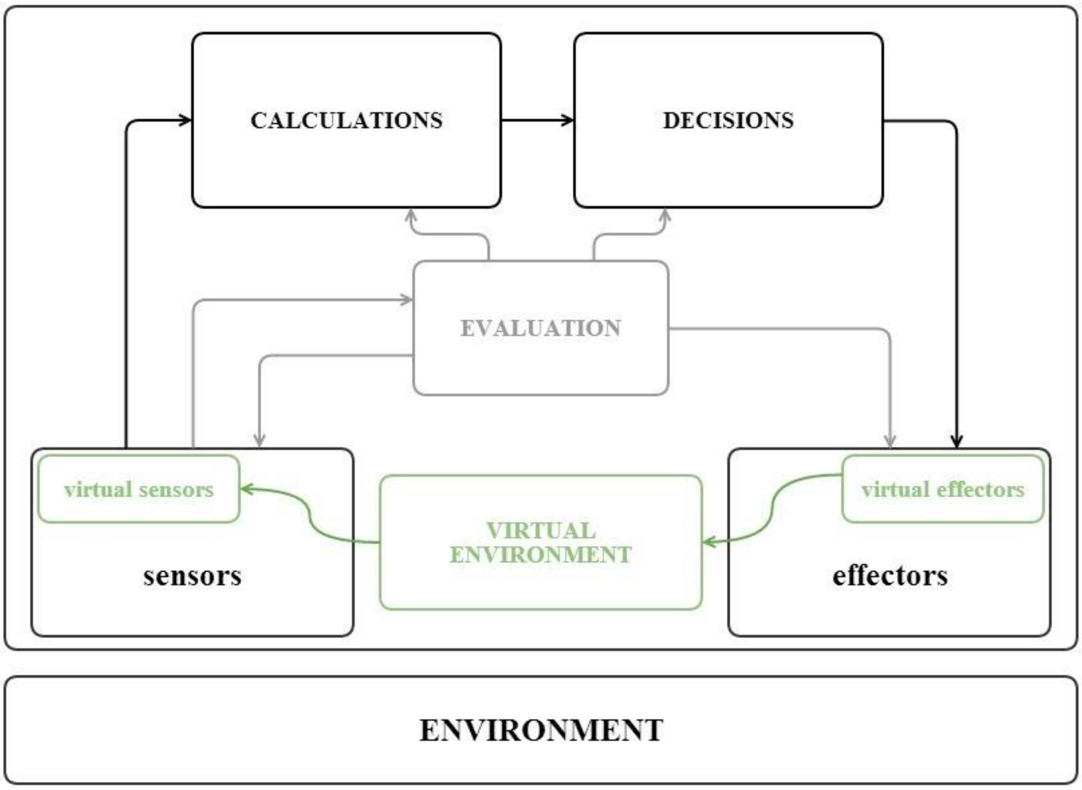


Figure 9.1: Generic system diagram.

---

[74] See figure 6.2.

9.3 The SOE/GNE Framework: Procedures

Once the basic concepts have been defined, the procedures are the actual methods through which the analysis is implemented. In this proposal, they consist in two steps: generation of the particular system diagram, by incorporating or leaving out the corresponding components of the generic diagram, and the filling of the tables, which is how all the relevant parts and behaviors of the system are made explicit.

9.3.1 Creating the Specific System Diagram

As said, a specific diagram for each of the analyzed works will be created as a first analytical step. This diagram constitutes the graphical representation and is the model respect to which emergence-relative-to-a-model will be evaluated. Thus, it is an explicit implementation of Cariani's idea, which does include also the possibility of serving as the basis to assess for emergence as self-organization phenomena, along with the emergence as generation of novelty with which Cariani's proposal is based on.

The creation of each of the particular diagrams starts from the generic diagram (figure 9.1) and proceeds by eliminating the irrelevant parts for the particular work under analysis. For instance, a robotic piece is likely not to have virtual components, and thus the virtual sensors, effectors and environment will be removed from the diagram. The idea is to indicate only the relevant components and the connections among them, so the particular way of working of the system under analysis is better understood.

The diagrams of three of the works analyzed in chapter 6 are shown below (figures 9.5, 9.8 and 9.9), along with the rest of the analytical procedures. As an example, the following figures show the diagrams of two more works. Figure 9.2 shows the diagram for Performative Ecologies. This piece does not have any virtual parts, since all the interactor sees is the robotic pieces; there are no screen-based elements in the work. Thus, all the virtual components are removed from the drawing. The diagram also reflects the pragmatic axis (decision-making), which does exist in this piece, as the robots use their sensors in order to evaluate how much of a user's gaze each of their dance moves is able to attract. This affects how they will run the calculations involved in creating new moves and the decisions of what moves will be repeated from one interactor to the next. As shown in the image, this evaluation does not directly affect how the sensors or the effectors work.

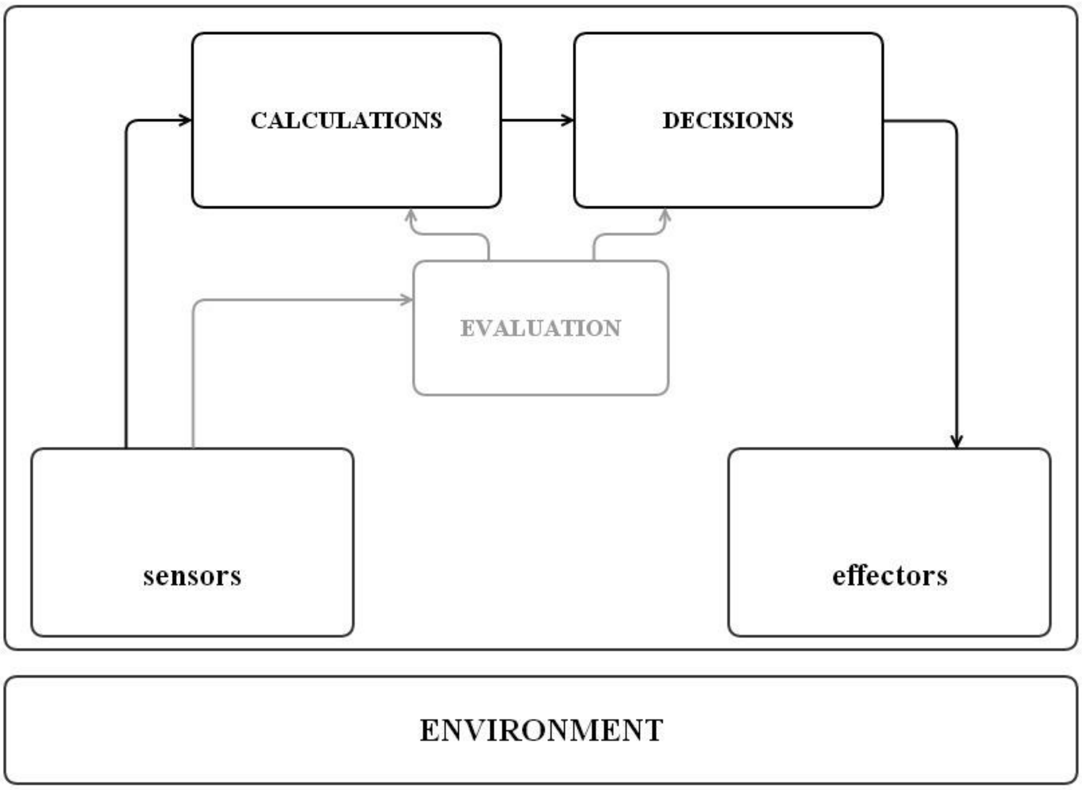


Figure 9.2: Performative Ecologies' diagram.

Figure 9.3 shows the diagram of Sniff. Differently from the previous example, this piece does have a virtual environment; the screen where the 3D dog is shown. However, it does not have virtual sensors and effectors represented in the diagram, because there are no local interactions with any other virtual agent. Sniff does also have the evaluative axis activated. According to the descriptions of the piece, the puppy changes its behavior depending on the state of engagement of the users. This means that, while sensing what the user is doing at each time, it evaluates this user's actions in terms of engagement with it, and this affects the decisions on the next moves to be made.

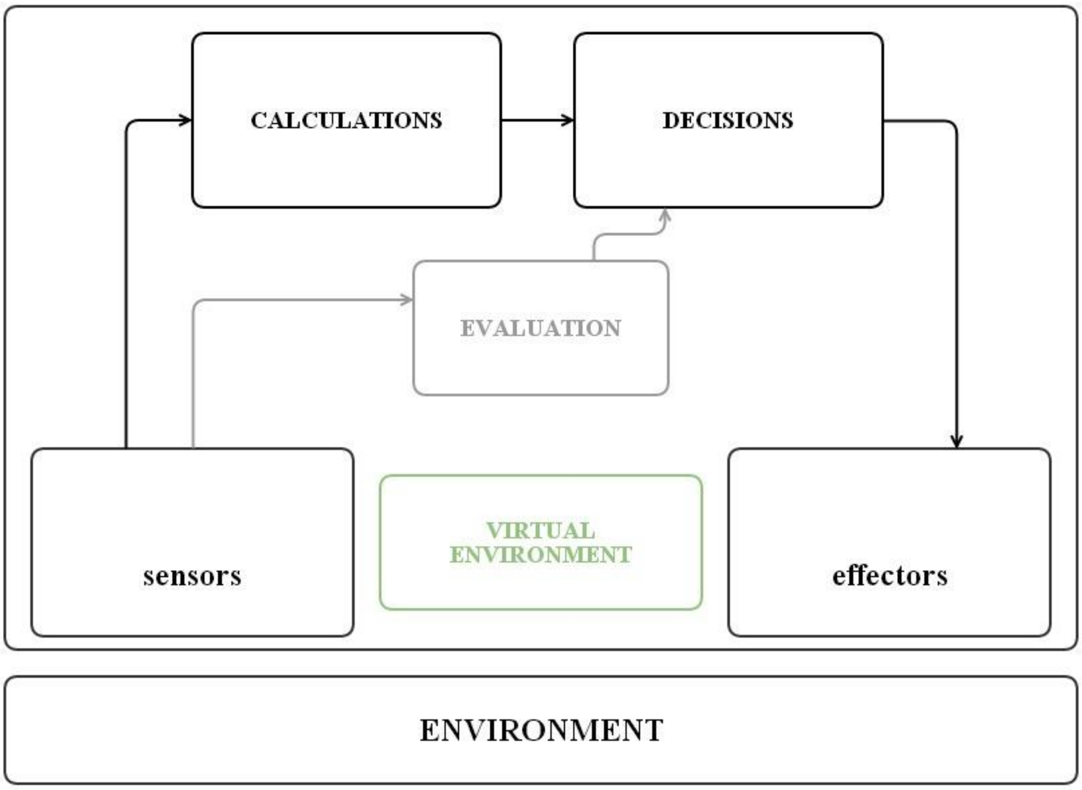


Figure 9.3: Diagram for Sniff.

As said above, figures 9.5, 9.8 and 9.9 below show the diagrams of three more pieces: the Boids, Sympathetic Sentience and Digital Babylon. Each of the diagrams serves as a tool to show the particularities of the corresponding piece. In the case of the Boids, which is not an interactive piece, the bottom part of the environment is removed, and so are the physical sensors and effectors. Only the virtual environment, sensors and effectors remain. On an opposite side, Sympathetic Sentience, which is a robotic piece, does not have any virtual part. Neither of these two pieces performs any evaluation of how the environment is affected by their actions or vice-versa, so this part is removed in both diagrams. Finally, Digital Babylon represents a middle-ground example, with both physical and virtual environment, sensors and effectors and an evaluation of how the agents perform in the environment. Details on the analysis of these three pieces are found below.

9.3.2 Analysis Tables

The second analytical step is to specify each of the relevant aspects of the system in terms of accounting for emergent behaviors in it. The proposal is to do so by using the following tables, which are an adaptation of the 'interactive decoupage' presented in (Freixa et al., 2014a). This first version of the decoupage was aimed at the analysis of online interactive news features and documentaries (webdocs). As it was presented, the decoupage has two steps: general description and in-depth analysis. Here this analytical framework is adapted to the needs of the proposal. Thus, the only original tables actually used are those corresponding to the general description (tables 0, a and b), since they provide useful general information. Some of the parameters have been removed or modified, especially those specific to the analysis of online work. In addition, many of the tables and parameters from the second part of the original decoupage proposal (in-depth analysis of media) could also be used, and this would provide a very detailed account on the types and subtypes of media that each work uses. However, in order to focus on the aspects that are relevant to the SOE/GNE Framework, these tables are omitted and four new specific tables are proposed.

The following are the adapted and newly proposed tables:

Authorship of analysis and relevant aspects for reception (module 0):

0 – Authorship of analysis			
Code	Indicator	Value	Description and procedure
001	Date of analysis	[mm/dd/yyyy]	Indicate the date the analysis was conducted
002	Author(s) of the analysis	[text]	Name of the author(s) and email(s)
003	Reception	[text]	Conditions of reception (seen at exhibition, through video and/or written documentation, etc.)

Table 9.1: Module 0: authorship of analysis and reception.

Identification data (module a):

a – Identification data			
Code	Indicator	Value	Description and procedure
a01	Title	[text]	Title of the work.
a02	Authorship	[text]	Author(s) and institution(s) supporting the product if any.
a03	Basic description by authors	[text]	Descriptive paragraph or abstract as provided by authors.
a04	Category	[text]	Category or genre to which the authors assign the work, if any.
a05	URL of project	[text]	URL of the project.
a06	Technology used	[text]	If available, information on the technological implementation of the work: sensors and effectors used, programming languages, etc.
a07	Launch date	[yyyy]	Indicate the launch date/first public presentation of the piece.
a08	Other versions	[text]	Indicate, if any, relevant information regarding other versions of the work.

Table 9.2: Module a: identification data.

Description and global assessment by the analyst (module b):

b – General description			
Code	Indicator	Value	Description and procedure
b01	Synopsis	[text]	Descriptive paragraph(s) where the analyst provides a general description of the application. It is a subjective description, wherein the different aspects of the piece are intertwined.
b02	Contextual description	[text]	The context of the work is described: whether it is part of some a series of work, collection or event; whether there are support websites, etc.
b03	Public	[text]	Identify the types of potential users.
a04	Category according to the classification in chapter 6	[text]	Category(ies) or genre(s) to which the work fits, if any, according to the proposal elaborated in chapter 6.
b05	Dominant aspects	[text]	Indicate those aspects which, from a subjective assessment point of view, are considered to identify and distinguish the application.

Table 9.3: Module b: general description and global assessment by author.

Analytical description of interactive system (module h[75]):

h – Phase 1: Description of Interactive system			
Code	Indicator	Value	Description and procedure
h01	System	[text]	General description of the system.
h02	Sensors	[text]	Description of system's sensors. Indicate whether they belong to the system as a whole, to an agent, or both.
h03	Effectors	[text]	Description of system's effectors. Indicate whether they belong to the system as a whole, to an agent, or both.
h03	Virtual Environment	[text]	Description of the virtual environment, if any. What it is, how it is presented to the interactor, etc.
h05	Primitives	[text]	Description of system's primitives. Indicate whether they are relevant to the computation's of the system as a whole, to an agent, or both.
h06	Computations	[text]	Description of system's computations: what operations does it perform with the primitives? Indicate whether they are relevant to the computation's of the system as a whole, to an agent, or both.
h07	Decisions	[text]	Description of system's mappings between computations and performed actions. Indicate whether they are relevant to the system as a whole, to an agent, or both.
h08	Evaluations	[text]	If relevant: how does the system evaluate its actions vis-à-vis system's goals.

Table 9.4: Module h: interactive system.

Observation of system's behaviors (module i):

i – Phase 2: Interactive system's behaviors			
Code	Indicator	Value	Description and procedure
i01.n	System's behavior N	[text]	Observed behavior of the system.
i02.n	Agents's behavior N	[text]	Observed behavior of a system's agent.

Table 9.5: Module i: System's behaviors.

---

[75] The lettering of the module moves here to 'h' in order to favor the consistency with the rest of the tables of the decopage, which end at the letter 'g'.

Evaluation of SOE behaviors (module j):

j – Phase 3: Evaluation of SOE			
Code	Indicator	Value	Description and procedure
j01	Presence of SOE	[yes/no]	State whether or not SOE is present in the system.
j02.n	SOE emergent behavior N	[text]	Observed emergent behavior.

Table 9.6: Module j: Evaluation of SOE.

Evaluation of GNE (module k):

k – Phase 4: Evaluation of GNE			
Code	Indicator	Value	Description and procedure
k01	Conditions of new observation	[text]	Describe the conditions of the observation: elapsed time respect the first one, etc.
k02	Presence of GNE	[yes/no]	State whether or not GNE is present in the system.
k03	Combinatoric or creative emergence?	None.	State whether the GNE found corresponds to combinatoric emergence, creative emergence or both.
k04.n	GNE emergent behavior N	[text]	Observed emergent behavior.

Table 9.7: Module k: Evaluation of GNO.

9.4 The SOE/GNE Framework: Phases

Finally, once the concepts and the methods are clear, the analysis is implemented in two phases, each of which doubles into two subphases. The first of them is descriptive, and it incorporates the modeling of the system and the description of the system and system's agents' behaviors through the diagram and the tables. The second phase is evaluative. It consists in the assessment of the presence or absence of emergent behavior (either as SOE or GNE), according to the model created and to further observation of the system when necessary.

	Subphases	**Procedure**	**Based on**
Phase 1: Description	(a) Identification and modeling of the system	Diagram and tables 0, a, b and h	System / artwork documentation
	(b) Description of the system's behaviors	Table i	Observation of the system in motion
Phase 2: Evaluation	(a) Evaluate the presence of SOE	Table j	Phase 1
	(b) Evaluate the presence of GNE	Table k	Phase 1, Phase 2a and further observation of the system in motion.

Table 9.8: The phases of the SOE/GNE Framework.

Thus, the proposed modeling framework is based on four subphases: first, a study of the system documentation in order to describe how the system is designed. The identification of the system is based on what Cariani identifies as the three axes: semantic (reading and affecting the environment), syntactic (computation and decision making) and pragmatic (evaluation of performed actions). This first step will include, when necessary, a virtual environment, and virtual sensors and effectors. While Cariani does not consider this aspect in his modeling of cybernetic devices, it is a useful addition in order to analyze interactive artworks which often consist partially in components that are not physically instantiated. The notion of environment, which in Cariani is left abstracted, includes here the potential user of the interactive artwork.

While this first analytical subphase is based on the system's documentation, the second subphase is based on observation: it consists in identifying the behaviors of the system and of its agents if they exist. Behavior is here understood as an epistemological term. It is an observed pattern that refers to a consistent series of actions and responses that the system exhibits when interacting with its environment and with the system's users.

This first phase is based both on system documentation and on observation of the system in motion. All the analytical steps corresponding to the first phase shall be made according to the system designer's descriptions (either through formal documentation or through other means). The behaviors shall be described through the observation of the system in motion (i.e. through simulation).

The third subphase consists in evaluating whether or not what was documented in Phase 1 and what was observed in Phase 2 constitutes a case of self-organization emergence. Finally, the last of the subphases consists in further observation of the modeled system in order to assess whether or not new structures, functions or behaviors appear, in what would be a case of generation of novelty emergence.

According to the proposed framework, only once the description of the behaviors has been made, a statement can be made regarding the presence of self-organization emergence and emergence as generation of novelty. In computationally mediated experiences such as those analyzed here, the notion of pre-programmed behaviors is important. While still within a degree of ambiguity on what was exactly preprogrammed and what was not, the general idea here is to determine whether or not a particular behavior was directly encoded by the programmer. If it is not, it is might be just a serendipitous appearance, or it could be a case of emergence if it follows the criteria that was made explicit.

9.5 Applications of the SOE/GNE Framework

What follows is the application of the framework to three of the pieces presented in chapter 6. The second and third are interactive, while the first is not. The reason to choose Boids as an example for the analysis, despite the fact that the framework is designed for interactive systems, is to show how it can easily accommodate other types of projects and also to test it against a prototypical case of emergence. What follows assumes a certain degree of familiarity of the reader with the analyzed works. Since these works were discussed earlier in the thesis, the general presentation is not replicated here. Only a small reminder introduces the tables. In another context this introductory presentation would precede the presentation of the tables and diagrams.

9.5.1 Analysis of Boids

Boids refers to the computer simulation model developed by Craig Reynolds and presented in SIGGRAPH in 1987. This model has been applied to a wide range of experiences (many of them are listed on Reynold's website). The analysis here looks at the original proposal, which was the recurring example within the ALife community in the 1990s. Since Boids is not an interactive piece, it does not have any sensors and effectors to interact with its environment. However, we can consider it to have virtual sensors and effectors regarding the individual agents. Calculations and decisions also exist, and in this case are made by the individual agents, too. This is shown in the diagram that follows.

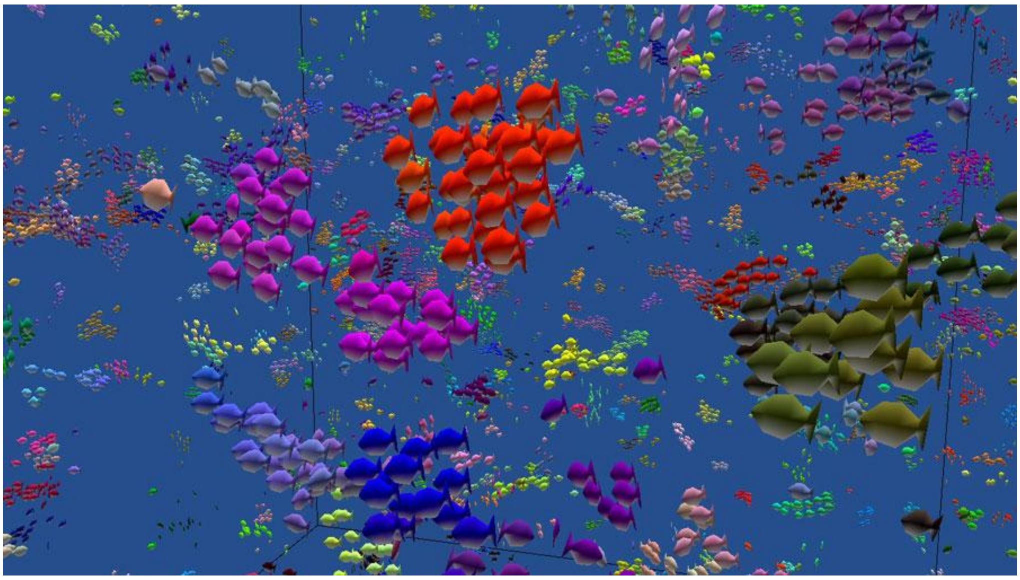


Figure 9.4: A 2006 implementation of the boids in PSCrowd, a library for the PlayStation 3 system (source: http://www.red3d.com/cwr/temp/PSCrowd/Fig_1_10000_fish.jpg).

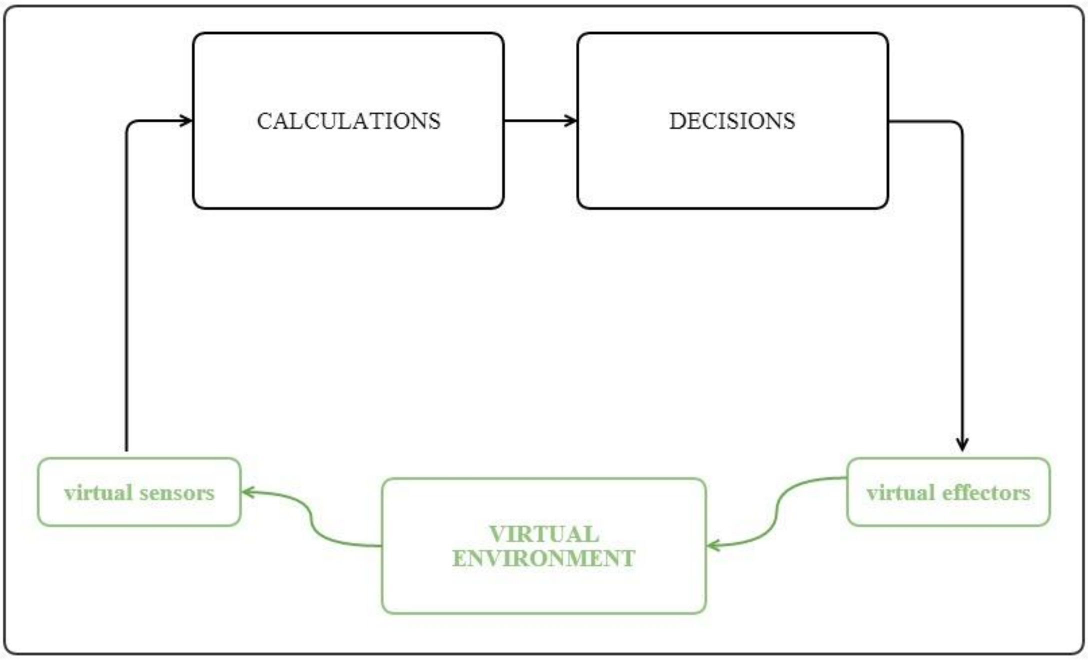


Figure 9.5: general diagram of Boids.

What follows are the tables for the analysis of Boids:

0 – Authorship of analysis		
Code	Indicator	Value
001	Date of analysis	01/29/2015
002	Author(s) of the analysis	Joan Soler-Adillon
003	Reception	Web documentation.

Table 9.9: Authorship of analysis of Boids.

a – Identification data		
Code	Indicator	Value
a01	Title	Boids
a02	Authorship	Craig Reynolds
a03	Basic description by authors	In 1986 I made a computer model of coordinated animal motion such as bird flocks and fish schools. It was based on three dimensional computational geometry of the sort normally used in computer animation or computer aided design. I called the generic simulated flocking creatures boids. The basic flocking model consists of three simple steering behaviors which describe how an individual boid maneuvers based on the positions and velocities its nearby flockmates: Separation: steer to avoid crowding local flockmates; Alignment: steer towards the average heading of local flockmates; Cohesion: steer to move toward the average position of local flockmates. Each boid has direct access to the whole scene's geometric description, but flocking requires that it reacts only to flockmates within a certain small neighborhood around itself. The neighborhood is characterized by adistance (measured from the center of the boid) and an angle, measured from the boid's direction of flight. Flockmates outside this local neighborhood are ignored. The neighborhood could be considered a model of limited perception (as by fish in murky water) but it is probably more correct to think of it as defining the region in which flockmates influence a boids steering. Source: http://www.red3d.com/cwr/boids/
a04	Category	Computer simulation technique.
a05	URL of project	http://www.red3d.com/cwr/boids/
a06	Technology used	Computer algorithm. It can be implemented in any programming language and environment.
a07	Launch date	1987
a08	Other versions	Many implementations of the Boids have been developed over the years. An example is this Play Station 3 experiment conducted by Reynolds himself in 2006: http://www.research.scea.com/pscrowd/

Table 9.10: Identification data for Boids.

b – General description		
Code	Indicator	Value
b01	Synopsis	The term boids is a contraction of bird-oid, as in ‘like a bird’. The boids are a computer model for simulating the movement of a flock of birds, school of fish, or a herd. It works through local rules: each agent calculates how to move according only to its neighboring agents. There is no central control deciding how the group shall move.
b02	Contextual description	The boids were created as a computer simulation technique and has been used by Reynolds himself and other in computer animation in movie effects.
b03	Public	Technical developers and, ultimately, general movie audiences.
a04	Category according to the classification in chapter 6	Autonomous agents.
b05	Dominant aspects	The boids are a classic example of emergence: three simple rules that apply only to the local agents are the driven force for a group behavior (flocking, schooling or herding).

Table 9.11: General description for Boids.

h – Phase 1: Description of Interactive system		
Code	Indicator	Value
h01	System	The group of agents. The flock, school or herd of autonomous agents.
h02	Sensors	None.
h03	Effectors	None.
h02	Virtual sensors	The agents access the position and heading of the agents located within a certain degree of proximity.
h03	Virtual Effectors	The agents are able to effect their own steering and movement within a certain degree of maximum velocity.
h04	Virtual Environment	The space were the agents move is a computer simulated environment, which can or cannot have obstacles in it. In the most simple of the implementations, only the agents are present in the environment.
h05	Primitives	Position of nearby agents; heading of nearby agents.
h06	Computations	Calculate medium heading of nearby agents; calculate medium position of nearby agents; calculate distance to each nearby agent.
h07	Decisions	Steer towards the medium heading of neighbors; stay close to the center point of the group; avoid collisions with other agents.
h08	Evaluations	None.

Table 9.12: System description for Boids.

i – Phase 2: Interactive system's behaviors		
Code	Indicator	Value
i01.1	System's behavior 1	Flock: Move around in groups that resemble schools of fish or flocks of birds; avoid obstacles and remain together.
i02.1	Agents' behavior 1	Move with group: head and move towards group direction.
i02.2	Agents' behavior 2	Stay within group: only abandon group to join another group; avoid collisions with mates and with obstacles.

Table 9.13: Description of behaviors for Boids.

j – Phase 3: Evaluation of SOE		
Code	Indicator	Value
j01	Presence of SOE	Yes.
j02.1	SOE emergent behavior 1	The group behavior (flocking) is exclusively the result of the local interactions. There is no central or external entity controlling group behavior. The way the group actually moves is generated from the local interactions, yet the movement of each agent is channeled through this group behavior.

Table 9.14: Evaluation of SOE for Boids.

k – Phase 4: Evaluation of GNE		
Code	Indicator	Value
k01	Conditions of new observation	Observation of documentation of computer simulations; observation of computer simulation after several minutes; academic literature.
k02	Presence of GNE	No.
k03	Combinatoric or creative emergence?	None.
k04.n	GNE emergent behavior N	None.

Table 9.15: Evaluation of GNE for Boids.

Thus, according to this the Boids are indeed an example of self-organization emergence. The flocking behavior is created through the accumulation of the interactions of the agents. There is nothing directing the flock. In turn, the flocking clearly channels the movements of the individual agents, in a feedback loop of group-agent influence. The effect can be certainly surprising to someone who sees it for the first time, even with full knowledge of the local rules. But once seen, it will be expected if the simulation is repeated. This is where the definitions of emergence that base its argumentation on the ideas of surprise and non-anticipation become problematic. Understanding emergence as self-organization eliminates these problems, since being emergent or not does not reside on surprise or lack of anticipation, but on a judgment that is made a posteriori. This framework is an attempt to clarify this.

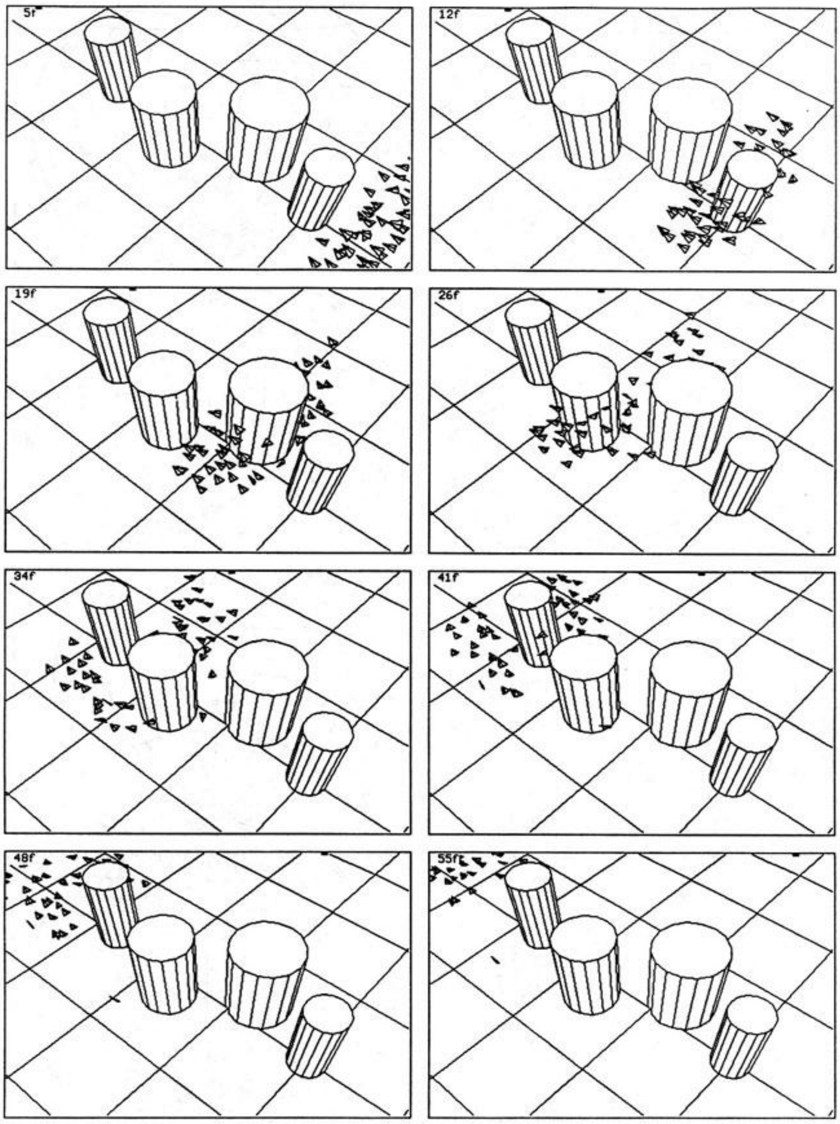


Figure 9.6: Boids motion test from 1986. A group moves around a series of cylinders (source: http://www.red3d.com/cwr/boids/images/flock_around_cylinders.jpg).

9.5.2 Analysis of Sympathetic Sentience

Sympathetic Sentience is an interactive installation created by Simon Penny in 1994. It is a good example to analyze because it explicitly attempts to generate complex behavior from a simple set of low-tech robotic units. The conclusion of the analysis is that it does generate self-organizing emergent behavior.

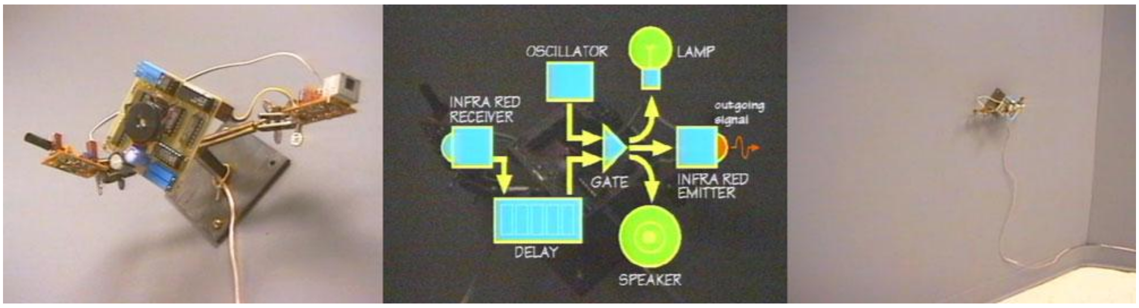


Figure 9.7: Sympathetic Sentience (video frames). Unit detail, diagram and unit in installation space (Penny, 2008).

In this case, there are no virtual elements in the system that are relevant to the analysis. Hence their removal from the diagram:

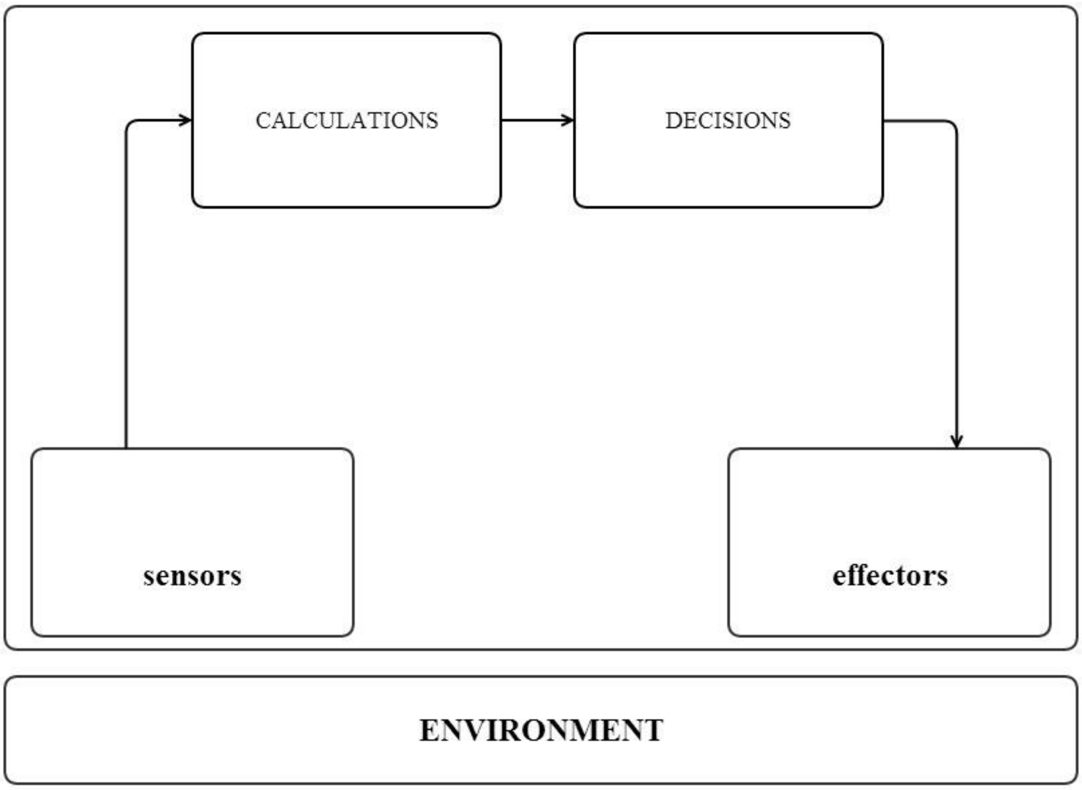


Figure 9.8: general diagram of Sympathetic Sentience.

0 – Authorship of analysis		
Code	Indicator	Value
001	Date of analysis	02/05/2015
002	Author(s) of the analysis	Joan Soler-Adillon
003	Reception	Web and video documentation.

Table 9.16: Authorship of analysis of Sympathetic Sentience.

a – Identification data		
Code	Indicator	Value
a01	Title	Sympathetic Sentience II
a02	Authorship	Simon Penny
a03	Basic description by authors	Sympathetic Sentience is an interactive sound installation which generates complex patterns of rythmic sound through to phenomenon of emergent complexity. Each of the eight identical units is alone capable of the most extremely simple on/off rythm. Rythmic complexity develops through the concatenation of the units. In the installation each unit passes its rythm to the next. Each unit then combines its own rythm with the rythm it recieves, and passes the resulting new rythm along. Thus the rythms slowly cycle around the group, increasing in complexity. A visitor can interupt this chain of communication by moving through the space. This results in a supression of communication activity and hence reduction of complexity. Along interuption results in complete silencing of the whole group. When the interupion is removed, slowly a new rythm will build up. The build-up of a new rhythm cycle can take several minutes. The rhythm cycles are never constant but continually in development. To gain experience of the full complexity of the piece, it is recommended to spend several minutes with the piece in an uninterupted state. Source: http://simonpenny.net/works/sympatheticII.html
a04	Category	Interactive Sound Installation.
a05	URL of project	http://simonpenny.net/works/sympatheticII.html
a06	Technology used	PIC chips, sound sensors and emitters, LEDs, infrared transmitters and receptors and other electronic parts.
a07	Launch date	1994
a08	Other versions	Simpathetic Sentience II (1997) and Sympathetic Sentience III (1999) extended the capabilities of the original work.

Table 9.17: Identification data for Sympathetic Sentience.

b – General description		
Code	Indicator	Value
b01	Synopsis	Sympathetic Sentience is an interactive installation formed by a group of robots that connect with each other trough infrared light beams. Each unit produces a pattern of rhythm, emits it, and passes it through infrared signals to the next unit, which will add a small modification to it and pass it along. The operation goes on in a loop that creates a systemic rhythmic pattern. This pattern can be altered if the communication among units is interrupted by someone who stands in the way of the infrared beams. This is the type of interaction that Sympathetic Sentience affords. The interruptions cause the receiving units to interpret the lack of communication as silence, and therefore to modify their own rhythmic pattern accordingly, to the extreme that if the interruption is persistent it can result in complete silence.
b02	Contextual description	Sympathetic Sentience is an interactive installation. The artist's interest resides in creating complex experiences from simple electronics and low-tech technology.
b03	Public	Art gallery public, interested in interactive and robotic art.
a04	Category according to the classification in chapter 6	Autonomous agents.
b05	Dominant aspects	Sympathetic Sentience creates a sort of awkwardly realistic sound environment. Through very simple single-pitched sounds the coupling of the units generates a group sound that resembles the sound of a group of birds singing within a small space. The piece does not claim to model or simulate any of this, but the result can be easily perceived as similar to being in the middle of a bird area on a zoo (for the gallery setting resembles that more than a natural environment experience). Also, the fact that moving around and standing close to the units can lead to silence might also be reminiscent of the birds quieting down as someone or something approaches them.

Table 9.18: General description for Sympathetic Sentience.

h – Phase 1: Description of Interactive system		
Code	Indicator	Value
h01	System	The group of units.
h02	Sensors	Infrared sensors that allow each unit to receive messages from the previous unit.
h03	Effectors	Infrared emitter that each unit uses to communicate its rhythmic pattern to the next.
h02	Virtual sensors	None.
h03	Virtual Effectors	None.
h04	Virtual Environment	None.
h05	Primitives	Rhythmic patterns. The on and off states of sound emitting during a short period of time.
h06	Computations	Read infrared light beam as a transmission for a rhythmic pattern.
h07	Decisions	Slightly modify the received pattern to create own.
h08	Evaluations	None.

Table 9.19: System description for Sympathetic Sentience.

i – Phase 2: Interactive system's behaviors		
Code	Indicator	Value
i01.1	System' behaviors 1	System sound: a complex systemic sound that builds up and evolves over time.
i02.1	Agents' behavior 1	Generate sound: each unit generates its own rhythmic pattern.
i02.2	Agents' behavior 2	Pattern communication: units receive a sound pattern from the previous unit in the group and pass along a pattern to the next.

Table 9.20: Description of behaviors for Sympathetic Sentience.

j – Phase 3: Evaluation of SOE		
Code	Indicator	Value
j01	Presence of SOE	Yes.
j02.n	SOE emergent behavior N	The group's 'symphony' is the result of the local interactions, with no conductor and no pre-scripting. Thus, the way the group actually sounds at every given moment is exclusively a result of local interactions among units. In turn, the way in which group sound is evolving affects back how each unit's sound is being affected, thus channeling the changes in each one.

Table 9.21: Evaluation of SOE for Sympathetic Sentience.

k – Phase 4: Evaluation of GNE		
Code	Indicator	Value
k01	Conditions of new observation	This assessment is done base on the online documentation and literature on the piece.
k02	Presence of GNE	No.
k03	Combinatoric or creative emergence?	None.
k04.n	GNE emergent behavior N	None.

Table 9.22: Evaluation of GNE for Sympathetic Sentience.

Like the Boids, Sympathetic Sentience is also an emergent system according to this framework. Here instead of moving the robotic units generate a complex auditive experience. In this case the metaphor of the symphony is useful to understand why this piece's outcome is emergent: there is no conductor directing how the sound unfolds. It is the accumulation of each of the interactions what creates it. But it is not just noise. As the sounds are being created, each unit is influenced by the overall generated symphony, because the sounds it receives change as the overall sound change. Thus, the feedback loop of agent-system influencing is here present again, and it is the basis of the behavior identified as emergent.

9.5.3 Analysis of Digital Babylon

Digital Babylon is an interactive installation of my own from 2005. It is relevant to apply the framework to it because of full availability of documentation and also of a degree of complexity in the piece that allows to show how the tables can stretch to accommodate for such a case. The general diagram and the tables show the degree of complexity that was already designed into the installation.

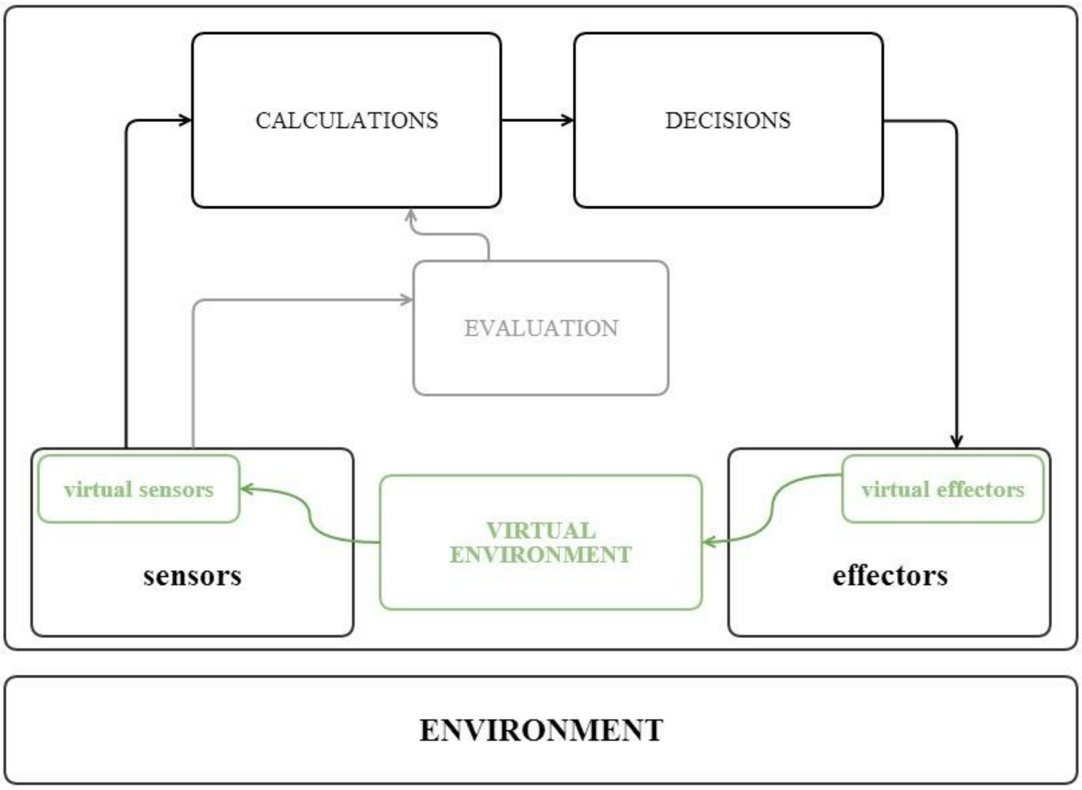


Figure 9.9: general diagram of Digital Babylon.

Tables for Digital Babylon:

0 – Authorship of analysis		
Code	Indicator	Value
001	Date of analysis	02/05/2015
002	Author(s) of the analysis	Joan Soler-Adillon
003	Reception	Web and video documentation.

Table 9.23: Authorship of analysis of Digital Babylon.

a – Identification data		
Code	Indicator	Value
a01	Title	Digital Babylon.
a02	Authorship	Joan Soler-Adillon
a03	Basic description by authors	Digital Babylon is an Artificial-Life interactive installation in which the visitor's presence alters the work's computer generated ecosystem This piece gives two choices to the visitors: contemplation and interaction. It is a digital ecosystem that works on its own without the need of human intervention. There are two creatures (the triangle beings and the round-headed tailed predator) and the plants. All of these interact with each other. The creatures eat, kill, die, mate and reproduce. By doing the latter, the two species evolve. If the visitor decides to intervene, the action will be done with his/her body, moving around a designated space. When doing so, the visitor's body's position will be mapped on the virtual world, and the triangle beings will be more or less attracted to the visitor, depending on how much friendly each one of them is. At this point, the participant decides either to help or to harm the triangle beings. Keeping them away from the predator will be helping them, and dragging them to the predator (to death) will kill some of them and harm the species. Depending on these actions, the next generations (where friendliness is one of many factors that the virtual parents pass on to their children) will be affected, becoming the species as a whole more or less friendly with visitors and, therefore, more or less inclined to interact with the human beings visiting the space. Source: http://joan.cat/en/dbn/
a04	Category	Interactive Installation.
a05	URL of project	http://joan.cat/en/dbn/
a06	Technology used	Computer, firewire camera, projector; programmed in Processing and C++.
a07	Launch date	2005
a08	Other versions	None.

Table 9.24: Identification data for Digital Babylon.

b – General description		
Code	Indicator	Value
b01	Synopsis	Digital Babylon is an installation that presents the user with a virtual 2D real-time generated ecosystem on a projected screen. Striving for simplicity in terms of representation, the ecosystem consists of a black background populated by a species represented by triangles and another one by blue tailed ellipses. Green plants appear in the form of a series of dots. The system works autonomously, without the intervention of the interactor. The two animal species relate as predator and prey. Both reproduce using genetic algorithms, and therefore are evolutive. On starting the system, there are a few randomly generated individuals and some plants. From here on, the ecosystem is left to find its equilibrium, which might eventually be disrupted by the human visitors.
b02	Contextual description	Digital Babylon is an interactive installation. It was presented as a Master's Thesis in New York University's Interactive Telecommunications Program, and publicly displayed at the NYU's facilities and presented at Universitat de Barcelona in 2006.
b03	Public	Art gallery public, interested in interactive art and ALife art.
a04	Category according to the classification in chapter 6	Simulated ecologies.
b05	Dominant aspects	Digital Babylon has a double layer of interaction. First, there is the direct interaction with a user entering the space: some of the elements in the piece will follow him or her. Second, there's the accumulated interaction, which builds up over time and affects how the species that interacts with the users does so.

Table 9.25: General description for Digital Babylon.

h – Phase 1: Description of Interactive system		
Code	Indicator	Value
h01	System	The ecosystem formed by the main species (S1), the predator (S2) and the plants.
h02	Sensors	A ceiling-mounted camera that tracks the position of the visitor in the space.
h03	Effectors	Projected screen showing the ecosystem and the mapping of the visitor's position.
h02	Virtual sensors	S1 have the following virtual sensors: -They can sense the presence of plants -They can sense the presence of the visitor -Within a range of proximity, they can sense the presence of predators -When ready to mate, they will know if there is another S1 individual ready and they will sense its position S2 have the following sensors: -Within a range of proximity, they can sense the presence of plants -Within a range of proximity, they can sense the presence of individuals of S1 -Within a range of proximity, they can sense the presence of other S2
h03	Virtual Effectors	Both S1 and S2 have virtual effectors that cause them to move around in the virtual ecosystem.
h04	Virtual Environment	The ecosystem, formed by individuals of S1, S2 and plants. It is a 2D environment of 800x600 pixels and the limits of the screen act as borders, so no agent can go through them.
h05	Primitives	-Position of each S1 agent -Heading of each S1 agent -Level of energy of each S1 agent -Position of each S2 agent -Level of energy of each S2 agent -Position of each plant -Position of each plant's points (units of food) -Position of visitor
h06	Computations	S1 does the following computations: -Check for predators within proximity -Check for the presence of the user -If going to a plant: check if close enough to check for food availability -If chasing a mate, check for mate's position S2 does the following computations: -Check for other S2 within proximity -Check for prey (S1) within proximity -In no prey is available, check for plants within proximity

h07	Decisions	S1 makes the following decisions: -If looking for food and food available: move to food -If trying to mate and mate is known: move towards mate -If predator close by, steer away -If visitor within installation space, weight all the previous decisions with the percentage of ‘affinity’ to interactor. E.g. if 100%, disregard all previous decisions and move to visitor; if 0%, ignore completely; if inbetween, move accordingly. S2 makes the following decisions: -Flock within group -If individual S1 close enough, ignore flocking and attack it -If no S1 to attack, steer towards the plants
h08	Evaluations	There’s no explicit evaluation done by the agents. However, since they reproduce using genetic algorithms, the information of successful individuals in terms of survival is passed along from generation to generation, thus generating and indirect evaluation method regarding the ability of survival of the agents.

Table 9.26: System description for Digital Babylon.

i – Phase 2: Interactive system’s behaviors		
Code	Indicator	Value
i01.1	System’s behavior 1	Rhythmic pattern: after some time, the ecosystem enters a rhythmic pattern of abundance/scarcity of food, which lead to high and low peaks of population. When the food is abundant, S1 agents feed easily and this leads to mating and, thus, a population growth which in turn consumes more food. Eventually, the food will all disappear, so the individuals will wonder around until several of those that are weakest die, and then more food starts appearing where they did. S2 is equally affected by these waves of scarcity and abundance, as it becomes more easy or more difficult for them to feed and, thus, to survive depending on each moment.
i01.2	System’s behavior 2	Plant generation: Wherever an agent of S1 or S2 dies, a new plant appears.
i02.1	Agents’ behavior 1	S1 wander around: if the agent is in need for food but there’s no food around, it just wanders around waiting for something to happen.
i02.2	Agents’ behavior 2	S1 eat: if the agent is in need for food and it knows there are plants available, it will go to the closest one and then see if, when it arrives there, any food unit in the plant is free. If it is, it will eat it. If not, it will move to the closest plant and try again.
i02.3	Agents’ behavior 3	S1 mate: if the agent has the sufficient amount of energy, it will try to mate. At this point, if no other S1 agent is in mating mood, it will wander around. Else, it will chase the one agent that is in mating mood and ignore all those that change into this mood afterwards. If it reaches the chased agent they will perform a mating dance. If this dance finishes before the sufficient energy for mating wears off, a new agent is born.
i02.4	Agents’ behavior 4	S1 avoid predator: at any time, if the agent perceives a nearby predator, it will try to avoid it and move away. The inertia of its movement, however, can cause it go get close enough to be attacked and, eventually, killed.
i02.5	Agents’ behavior 5	S1 get close to visitor: if a visitor enters the installation space, his or her position will be mapped onto the virtual ecosystem. Some agents will get close to this position, while others will ignore it completely.
i02.6	Agents’ behavior 6	S2 flock: each S2 agent will move towards an overall direction accordingly to its group, in a slow passed flocking behavior.
i02.7	Agents’ behavior 7	S2 haunt: if an S1 agent gets close by, attack it and try to eat it.
i02.8	Agents’ behavior 8	S2 move towards plants: while staying within the group, move towards the plants, as this is where the S1s might be found.

i02.9	Agents' behavior 9	S2 nest routine: after some time hunting, the group will go back to the nest (in the lower-right part of the screen). There some agents will stay an create an egg and some others will go out, not interact with anyone, and die. After that, a new egg pops us and the appearing group starts hunting.
i02.10	Agents' behavior 10	S2 kill plants: if an S2 touches a plant, the plant becomes brown and slowly fades away, as it ceases to serve as food for the S1s.

Table 9.27: Description of behaviors for Digital Babylon.

j – Phase 3: Evaluation of SOE		
Code	Indicator	Value
j01	Presence of SOE	Yes.
j02.1	SOE emergent behavior 1	System’s rhythmic pattern: the patterns of abundance and scarcity of resources is not pre-programmed. It is the result of the multiple interactions among S1, S2 agents and plants. In turn, the cycles channel the interactions among agents and plants and affect their chances of survival.
j02.2	SOE emergent behavior 2	S2’s flocking behavior: the group behavior is the result of the local interactions. There is no central or external entity controlling group behavior. The way the group actually moves is exclusively a result of local interactions, yet the movement of each agent is channeled through this group behavior.
j02.3	SOE emergent behavior 3	S1’s plague behavior: in relation to the eating behavior (i02.2), the S1 agents eat the plants in a way that strongly resembles a plague, especially when there are only a few plants and a lot of S1 agents, and in an area with no interferences. The agents rush to the nearest plant and, at the moment it fills out, they suddenly move away to the next closest plant, and so on, until they have devoured everything. This effect was accentuated in an early version consisting of only S1s and plants, with no mating behavior.

Table 9.28: Evaluation of SOE for Digital Babylon.

k – Phase 4: Evaluation of GNE		
Code	Indicator	Value
k01	Conditions of new observation	Observation of a simulation after a 13 hour run.
k02	Presence of GNE	Yes.
k03	Combinatoric or creative emergence?	Combinatoric.
k04.1	GNE emergent behavior 1	Group behavior of S1s: during scarcity phases, instead of wandering around the agents remained close to each other as if they were acting in group. Their movements had exaggerated the turning around, in a quasi-3D effect. This effect can be seen at https://vimeo.com/9192967

Table 9.29: Evaluation of GNE for Digital Babylon.

In Digital Babylon there are some effects that relate to self-organization phenomena: the cyclic processes of the system, the flocking behavior of the predators, and the plague behavior of the prey. The same arguments given before in favor of SOE for Boids and Sympathetic Sentience apply here.

In addition, according to the analysis this piece presents a case of combinatoric emergence. After a long simulation, behaviors that were not present in other presentations and simulations of the piece appeared on screen, as explained in the table field k04.1. Figures 9.7 and 9.8 show the distribution of the S1 individuals when wandering around and waiting for food to appear. While in the first image they distribute along the screen space in a rather uniform manner, according to the initial conditions of the system, the second image (taken after a 12 hour simulation) shows the S1 individuals gathering around the center of the screen, as if moving in group.

Figure 9.10: Digital Babylon. S1 individuals, during the first hour of installation running time, wandering around as they wait for food to be found.

Figure 9.11: Digital Babylon. S1 individuals in the same environmental condition as those in figure 9.10, but 12 hours within system running time.

The video documentation[76] shows the actual behaviors of the digital creatures and how they interact with the predator and the environment. In the evolved behaviors, individuals not only wander close to each other, but they hardly ever separate from their group. This turns out to be a good strategy for species survival, since even when several individuals die from starvation, since they all stay close to one another, the food appears were there are already S1 mates so this food is found and used in the benefit of the species almost immediately, before any predator can get close to the area. This observed behavior was not predefined –i.e. it was not preprogrammed– in the installation piece; it appeared after a long simulation as a result of the iterative application of genetic algorithms. It is according to this that it is argued here that it is a case of combinatoric emergence.

In fact, as it has been said in preceding chapters, the use of genetic algorithms is arguably a facilitator of this type of emergence, since they consist precisely in exploiting the combinatoric possibilities of what is defined as the digital DNA, which in general terms coincides with Cariani's primitives.

---

[76] See https://vimeo.com/9192657 (minute 1:05) for the behavior equivalent to figure 9.7 and https://vimeo.com/9192967 (minute 0:55) for the behavior portrayed in figure 9.8).

9.6 Concluding Remarks

The aim of the SOE/GNE Framework is to contribute to the clarification of how to assess the presence of emergent phenomena in interactive artwork. As discussed, this is a complex concept with ramifications in several disciplines. Therefore, the first step in such an endeavor is to clearly define and delimitate what the concept refers to. This is done here as explained in chapter 8: emergence is related to both self-organization and to generation of novelty. In the case of interactive art, this is done by artificial systems that relate to a human interactor.

The proposed framework is grounded on this differentiation, and it presents a methodology that allows for a detailed description of the systems under analysis –a description that can be pursued even with further detail if the rest of the tables from the original ‘interactive decoupage’ are used. This description forms a base from which the system and the system’s behaviors can be understood and described; it is the model from which emergence-relative-to-a-model is looked at, and it is also the basis for the description of self-organized behaviors.

In this chapter, the framework has been applied to three ALife Art related examples. However, this methodology is not exclusive of ALife Art, nor even of interactive art. It can be applied to any media-related experience in which it is interesting to determine whether or not emergence is present. Social media environments or videogames are arguably solid candidates to futures applications of the framework, provided the pertinent modifications are made in order to accommodate it to the particularities of each genre, as much as it does accommodate for each of the particular examples under analysis as shown above.

10. CONCLUSIONS

This concluding chapter addresses the research goals and research questions presented in the introduction, and evaluates whether or not they have been achieved and how they have been answered in the presented dissertation. Since the conclusions of each of the thesis chapters are already included at the end of each one, this final pages focus only on the mentioned items and on offering some pointers to future lines of research that might spread out from the work presented here.

10.1 Achievement of Research Goals

Research goal #1 (RG1): Review the state of the art, covering a body of literature that accounts for the main approaches to the concept of emergence that relate to Artificial Life and, thus, to Artificial Life Art.

Emergence appears as a subject in very different discourses and contexts. Some of them relate directly to the subject of study of this dissertation, while others relate only indirectly or, arguably, do not do so in any meaningful way. As it has been shown, emergence appears in scientific and philosophical contexts as either associated with the appearance of novelty and with self-organizing processes in complexity-based processes. In occasions, the association is with both formulations simultaneously. From the point of view of interactive art, both approaches to the concept are relevant, as emergence has been and can be used as an idea to both generate autonomous behavior in agent-based interactive systems and to provoke the appearance of novel functions and behaviors in systems that relate meaningfully to the interactor.

The seven chapters that constitute the body of the thesis are all part of this proposed literature survey. With a stronger weight on the chapters of the first part of the thesis, the survey is pervasive, as authors and approaches are presented and discussed throughout the whole work. The line that separates philosophical and scientific approaches to emergence is often diffuse, and even more so is the differentiation among Artificial Life, Cybernetics and other scientific approaches. However, the distinction is useful in order to discuss diverse approaches and influences.

In the first part of the thesis, the chapter on philosophy covers the origins of the term emergence have been covered, along with some of the most influential authors both directly and indirectly related to the field of Artificial Life (e.g. Jaegwon Kim and Mark Bedau). Other authors and texts, which are less influential in general terms, are also reviewed due to the relevance of the discussed ideas regarding the rest of the discussion. In the science chapter, approaches related to Chaos Theory and Dynamical Systems Theory have been reviewed, and also the fundamental contributions to Artificial Life of authors such as Christopher Langton and Steven Wolfram. The Cybernetics chapter covers the work of Ross Ashby, Grey Walter and Gordon Pask that constitutes a body of work that is, as it has been discussed, a direct antecessor of Artificial Life. Finally, Chapter 5 reviews Peter Cariani's theory of emergence-relative-to-a-model and its influences. This theory is fundamental for the proposal elaborated in the presented dissertation.

The second part of the thesis, while moving away from the strict state of the art scheme, continues to amplify the theoretical ground on which the thesis is based. Along with a list of examples of ALife Art, chapter six discusses the techniques of ALife that are relevant to this movement: abstract systems such as Cellular Automata and fractals, Genetic Algorithms and evolutionary programming, generative art and situated robots. It also contains a discussion of eighteenth century automata as an antecessor of ALife devices. Chapter seven reviews the literature on the concept of interactivity in order to relate it to emergence, and chapter 8 covers the game design use of the idea of emergence and Ernest Nagel's proposal, which is coincidental, in general terms, to the differentiation between emergence as self-organization and as generation of novelty.

Arguably, the overall body of work formed by the reviewed literature constitutes a sound and sufficient basis in order to account for the main ideas and approaches related to the concept of emergence, from the point of view of its relation to Artificial Life Art.

Research goal #2 (RG2): Make a proposal to understand emergence in relation to interactive art that is inclusive and that strives for simplicity.

One of the problems with multi-discursive concepts such as emergence (and for that matter also with interactivity, self-organization or creativity) is that they are understood differently in distinct contexts and even within them by different authors. Furthermore, it is not rare that the premises from which an author writes about the concept are left unspecified, leaving to the reader's intuition what is meant by these ideas exactly. Arguably, this is based on the assumption that the reader knows what these ideas are, and indeed every informed reader will have a notion in mind when reading about them. But because of this multi-discursive nature, the mental model that the reader has can be too far apart from that of the author.

Thus, in order to articulate a valid academic discourse, it is necessary to state with precision to what one is referring to when discussing these concepts. This can be done either by proposing new definitions or by stating which authors one is basing his or her argumentation. It is not about trying to find the one right formulation –that arguably does not exist– but about stating with clarity what one is writing about.

At the same time, it is important that the proposed approaches to these ideas are inclusive. That is, that the proposed definitions don't narrow down the scope of the concepts up to the point of making them inconsistent with most of the literature. The difficulty of the task is on offering clarity and specificity while, the same time, accommodating for the broadness of the concepts under discussion.

According to this, this dissertation proposes to understand emergence as related to two concepts: self-organization and novelty. As sated, this idea finds its roots in Nagel and it incorporates Cariani's emergence-relative-to-a-model and the literature on emergence understood as self-organization. The proposal defines self-organization emergence as 'the appearance of a(n observed) pattern at the level of a system that is produced by the local interactions of the system's agents. This pattern has an effect on the agents, by channeling their behavior'. Emergence as generation of novelty is understood through Cariani's method, which includes the differentiation between combinatoric and creative emergence.

This scheme is useful in order to discern when emergence is used in relation to novelty and when it is not. Arguably, it is also valid to identify the miss-uses of the term emergence, such as in those cases when it is essentially conflated to the mere appearance of something or to the open properties of systems that offer a vast number of possible outcomes.

Research goal #3 (RG3): Propose an analytical framework for interactive art pieces that deal with the concept of emergence.

Once the terms under discussion are clear, it is necessary to present a method for assessing when they are present in the works under analysis and when they are not. In the case of emergence and interactive artwork, the goal is to elucidate a way with which the presence of emergent behavior can be accounted for. The intention in elaborating it is that the presented method can be applied to all kinds of interactive pieces in order to discuss the presence or absence of emergent features and behaviors in them.

In this respect, chapter 8 presents and exemplifies the SO/GNO Framework. An analytical framework aimed at identifying and discussing the cases were emergence is claimed to be present in interactive artwork. This framework is based, methodologically, on a previous project that analyses interactive web-based documentaries. Conceptually, it is based on the work presented in the dissertation (and referred to in RG2): the differentiation between emergence as self-organization and as generation of novelty, and the Cariani-inspired importance of making explicit the model of each evaluated work in order to account for them.

10.2 Answering the Research Questions

As stated in the introduction, the research questions of the presented investigation deal with its two core concepts: emergence and interactivity, and with the idea of emergent interactive behavior, which is its driving force. First, and with the presented dissertation in mind, it is necessary to state whether or not the idea of emergence can be formulated into an idea that, according to RG2, is both clear and inclusive in respect to the literature on the concept, and that is a useful idea in discussing interactive art. If this is answered positively, a clear statement on what is understood as interactivity must also be made in order to present a soundly based idea of what is understood by emergent interactive behavior.

Research question #1 (RQ1): Is emergence a useful concept to discuss interactive art?

In academic terms, a useful concept is one that can be defined with clarity, stating its ambiguities if it is the case, so they can be discussed and analyzed in detail. Without this, the discussions lack in occasions of a soundly based background that facilitates the dialog among researchers. As stated before, emergence is sometimes used in a context that allows for a kind of ambiguity that makes it difficult to discern what authors are referring to exactly and when something begins and ends to be related to emergence. Only an explicit clarification of what is meant by it eliminates this type of problems.

In accordance to this, the answer to this question is: yes, provided what is meant by emergence is clearly stated. As it has been shown, the concept is used diversely across different disciplines, and different approaches provide distinct accounts of the idea. Whilst it is arguably unrealistic to provide a definition that accommodates all these approaches, in order for the concept to be academically useful in the discussions on interactive art –and for that matter on any other discipline– a clearly stated definition of the term must be provided. This is what will allow for the ‘scientific communication’ claimed by authors like Peter Cariani. The SO/GNO Framework is an attempt towards solving this problem.

Research question #2 (RQ2): How can we formulate the concept in an unambiguous manner to use it in this context?

According to the answer to the answer of RQ1, a clear statement of what is understood by emergence must be made in order to articulate a coherent discourse around the concept. In the context of interactive art, it is fundamental that the formulation clearly delimitates what is emergent and what is not, and that goes along with a similar delimitation of the concept of interactivity (see RQ4).

As stated in the evaluation of RG2, emergence is here defined as being related to both self-organization and to the appearance of novelty. The terms 'self-organization emergence' and 'generation of novelty emergence' are an attempt to clarify this. For the former, the definition given above is presented: 'self-organization emergence is the appearance of a(n observed) pattern at the level of a system that is produced by the local interactions of the system's agents. This pattern has an effect on the agents, by channeling their behavior.' The latter is based on Peter Cariani's emergence-relative-to-a-model: emergence as novelty is only possible to be accounted for after a model of the system under observation has been presented. Then, when the primitives with which this system works recombine to generate new functions, effects or concepts, combinatoric emergence occurs. When new primitives enter into the calculations and decision making processes of the system, and generate the new functions, effects or concepts in this way, then it is a case of creative emergence.

Research question #3 (RQ3): Is the idea of designing emergent behavior coherent?

Emergent interactive behavior was an idea formulated in the 1990s in the context of Artificial Life Art. The idea aimed at creating interactive systems that exhibit or generate behaviors that are emergent. Arguably, this idea can be viewed as one of the ultimate goals of Artificial Life: create systems that generate their own behaviors, and that through these emergently generated behaviors they relate to an interactor in novel and autonomous ways, thus freeing the device from the constraints of the designer's pre-specifications.

The answer to the question is: yes, in accordance to the answer of RQ1. Emergent interactive behavior, as discussed in secition 8.7, can either be understood as the appearance of emergent behavior in an interactive system or as the appearance of emergence in the interactive behavior of the system (i.e. the way in which the system interacts with its users changes according to the parameters defined as emergence). In either case, provided that both the idea of emergence and of interactivity are clearly stated, emergent interactive behavior can become not only a valid idea but a very powerful driving force for the creation of interactive art and digital media in general.

Research question #4 (RQ4): How can we define interactivity coherently with the idea of emergent interactive behavior?

Within the paradigm of poetic interaction, as discussed in chapter 7, interactivity can be described as 'a series of related actions between two agents where (1) at least one of them is an artificial system that (2) processes its responses according to a behavior specified by design and (3) takes into account some of the previous actions executed by both agents.' This definition is inclusive enough to accommodate for the practices that relate to ALife Art among many others, and succinct enough to be a useful approach to define such a complex idea.

There is no claim here that this is the ultimate definition of interactivity. However, it is a characterization of the term that aims at making it operable in order to discuss the idea in the contexts of art and communication. As stated in the chapter dedicated to this, it is important to understand that interactivity is here defined from the point of view of system design, that it understands interactivity as a process of communication, and that it is to be distinguished from participation in a technologically mediated environment.

The first means that the idea is viewed from the design perspective. That is, what is most important is how the system is built and programmed in order to respond to an interactor. The second means that, according to this first point, interactivity is not just an affordance of the system (i.e. that it has buttons and knobs for things to be activated and modulated), nor is it solely concerned with the psychological process of perception of interaction by the user of the system. Finally, the differentiation between interactivity and participation is important to stress that here the focus in on how the system performs its actions in order to relate to the interactor, not on how it mediates the interaction among users or on how it allows the users to create content (e.g. text in a web forum or post a photograph on a social network). Therefore, the focus of interest here is on interactivity as it was understood before the social media era, and even before the web. As stated in the following section on future work, this ‘classic’ idea of interactivity is arguably about to regain importance in the light of to the current developments of personal communication technology.

10.3 Future Work

The most direct line of work that follows the presented dissertation is a thorough analysis of ALife Art and interactive art pieces that claim to exhibit emergent behavior, or that are related to this idea in one way or another. This study would consist in the application of the analytical model to these pieces, in order to describe how they are composed, how they behave, and whether or not they present emergent behaviors. This could allow for the clarification of a landscape of ALife Art and its relation either to self-organization emergence or to emergence as generation of novelty.

A second line of research directly related to this would be to elaborate the guidelines for the design of systems that afford the appearance of emergent behavior in either of its two forms. As it has been defined here, emergence cannot be directly designed. Only the conditions for it to appear can. But understanding what is emergence exactly and how can it be accounted for is the basis for elaborating such guidelines. With them, designers of interactive systems can undertake the creation of systems that aim at producing either novel behaviors in respect to what was pre-specified on them, or self-organized structures that drive how the agents in the system behave.

As said in the introduction, this dissertation is a multidisciplinary investigation. In accordance to this, the possible application of this study spread beyond interactive art, into the field of communication in general. Social media is arguably a ground were these ideas can be applied. The interconnectedness of large numbers of nodes and the vast amount of messages that circulate and relate to one another is, theoretically, a fertile ground for emergence to appear. In this context, non directed self-organized patterns and new types of communication might be captured through the lens of emergence.

Furthermore, as we move from the social media era to the era of the universal connection of devices (the Internet of things), as it is arguably happening right now in 2015, these considerations will only grow in importance. The more that there is connected and sending information back and forth, in a cybernetic universal machine run though the Internet, the more novelty and self-organization will need to be accounted for.

In addition, this new era of connected devices will bring back to the forefront of research and popular interest the idea of interactivity as it was understood in the 1990s; the relation between the human and the machine (both as the big machine and the particular device). If during the social media era the interest has shifted towards participation, arguably as more and more devices acquire digital interfaces –and interface with the digital– the question of how we actually interact with these devices and with the whole that is connected to it will gain weight, shifting again the balance to a more equilibrated point of interest both towards participation and interactivity as it was presented here. Within this context, the importance of notions such as emergent interactive behavior is very likely to re-emerge.

11. REFERENCES

Adams, E., & Dormans, J. (2012). *Game Mechanics. Advanced Game Design*. Berkeley: New Riders.

Ali, S. M., Zimmer, R. M., & Elstob, C. M. (1998). The Question Concerning Emergence: Implications for Artificiality. In D. M. Dubois (Ed.), *Computing Anticipatory Systems: CASYS - First International Conference. AIP Conference Proceedings 437* (pp. 138–156).

Almirón, N. (2001). *De Vannevar Bush a la WWW. Una genealogia de la humanització de les tecnologies de la informació: els pares de la interficie humana*. València: 3i4 Edicions.

Alsina, P. (2009). *Arte y vida: sistemas emergentes en las prácticas artísticas vinculadas a las tecnociencias* (Doctoral dissertation). Universitat de Barcelona.

Anderson, P. W. (1972). More is Different. *Science*, *177*(4047), 393–396.

Aristotle. (1994). *Metaphysics*. Cambridge, Massachussetts: MIT Classics Online. Retrieved from http://classics.mit.edu/Aristotle/metaphysics.8.viii.html

Ascott, R. (2001). Behaviorist Art and the Cybernetic Vision. In R. Packer & K. Jordan (Eds.), *Multimedia. From Wagner to Virtual Reality* (pp. 95–103). New York: WW Norton & Company.

Ascott, R. (2003). The Construction of Change. In N. Wardrip-Fruin & N. Montfort (Eds.), *The New Media Reader* (pp. 127–132). Cambridge, Massachussetts: MIT Press.

Ashby, W. R. (1957). *An Introduction to Cybernetics*. New York: Wiley.

Ashby, W. R. (1962). Principles of the self-organizing system. In H. Von Foerster & G. W. Zopf (Eds.), *Principles of Self-Organization: Transactions of the University of Illinois Symposium* (pp. 255–278). London: Pergamon Press.

Ashby, W. R. (1976). *Design for a Brain*. New York: John Wiley and Sons.

Ashby, W. R. (1981). *Mechanisms of Intelligence: Ross Ashby's Writings on Cybernetics*. (R. Conant, Ed.). Salinas, CA: Intersystems Publications.

Assad, A., & Packard, N. H. (2008). Emergence. In M. A. Bedau & P. Humphreys (Eds.), *Emergence. Contemporary Readings in Philosophy and Science2* (p. 231,234). Cambridge, Massachussetts: MIT Press.

Baecker, R. (2009). The conversation. Retrieved September 01, 2014, from http://www.rlfbckr.org/work/the_conversation

Baecker, R. M., & Buxton, W. A. S. (1988). *Readings in Human-Computer Interaction: A Multidisciplinary Approach*. San Mateo, CA: Kaufmann.

Baljko, M., & Tenhaaf, N. (2006). Different experiences, different types of emergence: A-life sculpture designer, interactant, observer. In *Proc. AAAI Fall Symposium Series, Interaction and Emergent Phenomena in Societies of Agents*. Arlington, VA.

Baljko, M., & Tenhaaf, N. (2008). The aesthetics of emergence. *ACM Transactions on Computer-Human Interaction*, *15*(3), 1–27. doi:10.1145/1453152.1453154

Barnes, J. (1979). *The Presocratic Philosophers: Empedocles to Democritus, Volume 2*. London: Routledge and Paul.

Beckermann, A., Flohr, H., & Kim, J. (Eds.). (1992). *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism*. Berlin-New York: Walter de Gruyter.

Beckers, R., Holland, O., & Deneubourg, J.-L. (2000). Stigmergy, self-organization, and sorting in collective robotics. In H. Cruse, J. Dean, & H. Ritter (Eds.), *Prerational Intelligence: Adaptive Behavior and Intelligent Systems Without Symbols and Logic* (pp. 1008–1022). Dordrecht, Holland: Springer Netherlands.

Bedau, M. A. (1997). Weak Emergence. In J. Tomberlin (Ed.), *Philosophical Perspectives: Mind, Causation, and World* (Vol. 11, pp. 375–399). Oxford: Blackwell Publishers.

Bedau, M. A. (2008). Downward Causation and Autonomy in Weak Emergence. In M. A. Bedau & P. Humphreys (Eds.), *Emergence. Contemporary Readings in Philosophy and Science* (pp. 155–188). Cambridge, Massachussetts: MIT Press.

Bedau, M. A., & Humphreys (Eds.). (2008). *Emergence. Contemporary readings in philosophy and science*. Cambridge, Massachussetts: MIT Press.

Bergson, H. (1911). *Creative Evolution* (1st ed.). New York: H. Holt and company. Retrieved from http://catalog.hathitrust.org/Record/001384767

Bernabé, A. (1988). *De Tales a Demócrito. Fragmentos presocráticos.* Madrid: Alianza Editorial.

Bernstein, D. S. (2002). Feedback control: an invisible thread in the history of technology. *IEEE Control Systems*, *22*(2), 53–68.

Bertalanffy, L. von. (1969). *General System Theory. Foundations, Development, Applications*. New York: George Braziller.

Bird, J., Layzell, P., Webster, A., & Husbands, P. (2003). Towards Epistemically Autonomous Robots: Exploiting the Potential of Physical Systems. *Leonardo*, *36*(2), 109–114.

Birringer, J. (2005). Interactivity: “user testing” for participatory artworks. *International Journal of Performance Arts and Digital Media*, *1*(2), 147–173. doi:10.1386/padm.1.2.147/1

Boden, M. A. (2009). *Creativity and artificial evolution*. Retrieved from http://www.informatics.sussex.ac.uk/courses/creative-systems/papers/maggie/crtvy-evol.pdf …

Boden, M. A. (2010). *Creativity and art. Three roads to success*. London: Oxford University Press.

Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, *47*(1-3), 139 – 159.

Brooks, R. A. (1999). *Cambrian intelligence. An early history of the new AI*. Cambridge, Massachussetts: MIT Press.

Brown, O. (2012). Generative and Adaptive Creativity: A Unified Approach to Creativity in Nature, Humans and Machines. In *Computers and Creativity* (pp. 361–382). New York: Springer.

Brown, P. (2001). Sand lines, kinetic painting. Retrieved September 01, 2014, from http://www.fundacion.telefonica.com/es/at/vida/vida10/paginas/v3/ebrown.html

Bucy, E. P. (2004). Interactivity in Society: Locating an Elusive Concept. *The Information Society*, *20*(5), 373–383. doi:10.1080/01972240490508063

Bucy, E. P., & Tao, C.-C. (2007). The Mediated Moderation Model of Interactivity. *Media Psychology*, *9*(3), 647–672. doi:10.1080/15213260701283269

Burks, A. W. (1970). Von Neumann’s Self-Reproducing Automata. In A. W. Burks (Ed.), *Essays on Cellular Automata* (pp. 3–64). Urbana, IL: University of Illinois Press.

Burnham, J. (1968). *Beyond Modern Sculpture*. New York: George Braziller.

Bush, V. (1945). As We May Think. *The Atlantic Monthly*, *176*(1), 101–108.

Campbell, J. (1999). Interactive Hallucination. Retrieved November 21, 2014, from http://www.jimcampbell.tv/portfolio/installations/interactive_hallucination/

Campbell, J. (2000). Delusions of Dialogue: Control and Choice in Interactive art. *Leonardo*, *33*, 133–136.

Campbell, R., & Bickhard, M. H. (2001). *Physicalism, emergence and downward causation*. *Online: http://www. lehigh. edu/~ mhb0/ …* (pp. 1–33). doi:http://hdl.handle.net/1885/41068

Cariani, P. (1989a). Adaptivity, emergence, and machine-environment dependencies. In *Proceedings of the 33rd Ann. Meeting, Int. Society for System Sciences (ISSS)* (p. Vol. III 31–37). Edimburg: Ledington PWJ.

Cariani, P. (1989b). *On the design of devices with emergent semantic functions* Doctoral dissertation). State University of New York at Binghamton (Binghamton University). Retrieved from http://www.cariani.com/CarianiNewWebsite/Cybernetics_files/CarianiPhDIntegrated1989.pdf

Cariani, P. (1990a). Adaptive coupling to the world through self-organizing sensors and effectors. In A. Meystel, J. Herath, & S. Gray (Eds.), *Proceedings of the Fifth IEEE Conference on Intelligent Control* (pp. 73–78). Philadelphia.

Cariani, P. (1990b). Implications from structural evolution: semantic adaptation. In C. M (Ed.), *Proceedings, international joint conference on neural networks* (pp. 47–51). Washington.

Cariani, P. (1991). Some epistemological implications of devices which construct their own sensors and effectors. In F. J. Varela & P. Bourgine (Eds.), *Toward a Practice of Autonomous Systems. Proceedings of the First European Conference on Artificial Life* (pp. 484–493). Cambridge, Massachussetts: MIT Press.

Cariani, P. (1992a). Adaptivity and emergence in organisms and devices. *World Futures*, *32*(2), 111–132.

Cariani, P. (1992b). Emergence and artificial life. In C. G. Langton, C. Taylor, J. D. Farmer, & S. Rasmussen (Eds.), *Artificial Life II* (pp. 775–797). Reading, Massachussetts: Addison-Wesley Publishing Company.

Cariani, P. (1993). To evolve an ear: epistemological implications of Gordon Pask's electrochemical devices. *Systems Research*, *10*(3), 19–33.

Cariani, P. (1997). Emergence of new signal-primitives in neural systems. *Intellectica*, *25*(2), 95–143. Retrieved from http://www.intellectica.org/SiteArchives/archives/n25/25_06_Cariani.pdf

Cariani, P. (1998a). Epistemic autonomy through adaptive sensing. In *Proceedings of the 1998 IEEE ISIC/CIRA/ISAS Joint Conference* (pp. 718–723). Gaithersburg, Maryland.

Cariani, P. (1998b). Towards an evolutionary semiotics: the emergence of new sign-functions in organisms and devices. In G. van de Vijver, S. N. Salthe, & M. Delpos (Eds.), *Evolutionary systems* (pp. 1–16). Dordrecht, Holland: Kluwer Academic. Retrieved from http://link.springer.com/chapter/10.1007/978-94-017-1510-2_26

Cariani, P. (2003). Cybernetic systems and the semiotics of translation. In S. Petrilli (Ed.), *Translation translation* (Vol. 200, pp. 349–368). Amsterdam, Netherlands: Rodopi.

Cariani, P. (2008). Design Strategies for Open-Ended Evolution. In S. Bulloch, J. Noble, R. Watson, & M. A. Bedau (Eds.), *Artificial Life XI. Proceedings of the Eleventh International Conference on The Simulation and Synthesis of Living Systems* (pp. 94–101). Cambridge, Massachussetts: MIT Press.

Cariani, P. (2009a). Emergence and Creativity. In *Emocão Art.ficial 4.0: Emergência* (pp. 20–42). Sao Paolo, Brazil: Itaú Cultural.

Cariani, P. (2009b). Strategies for Creating New Informational Primitives in Minds and Machines. In *Proceedings of the Dagstuhl Seminar on Computational Creativity* (pp. 1–13). Wadern. Retrieved from http://drops.dagstuhl.de/opus/volltexte/2009/2192/pdf/09291.CarianiPeter.Paper.2192.pdf

Cariani, P. (2009c). The homeostat as embodiment of adaptive control. *International Journal of General Systems, 38*(2), 139–154. doi:10.1080/03081070802633593

Cariani, P. (2010). On the Importance of Being Emergent. *Constructivist Foundations*, *5*(2), 86–91.

Cariani, P. (2011). The Semiotics of Percept-Action Systems. *International Journal of Signs and Semiotic Systems, 1*(1), 1–17.

Cariani, P. (2012). Creating New Informational Primitives in Minds and Machines. In J. McCormack & M. D'Inverno (Eds.), *Computers and Creativity* (pp. 395–430). New York: Springer.

Caston, V. (1997). Epiphenomenalisms, Ancient and Modern. *The Philosophical Review*, *106*(3), 309–363.

Codina, L. (2000). Evaluación de recursos digitales en línea: conceptos, indicadores y métodos. *Revista Española de Documentación Científica, 23*(1), 9–44.

Collins, D. (2002). Breeding the Evolutionary: Interactive Emergence in Art and Education. In *Annual Digital Arts Symposioum: Neural Network*. Tucson, AZ: University of Arizona.

Corning, P. A. (2002). The Re-emergence of "Emergence": A Venerable Concept in Search of a Theory. *Complexity*, *7*, 18–30. doi:10.1002/cplx.10043

Crampton, J. W. (2002). Interactivity Types in Geographic Visualization. *Cartography and Geographic Information Science*, *29*(2), 85–98. doi:10.1559/152304002782053314

Crawford, C. (2003). *Chris Crawford on Game Design*. San Francisco: New Riders.

Crutchfield, J. P. (2008). 15 Is Anything Ever New? Considering Emergence. In M. A. Bedau & P. Humphreys (Eds.), *Emergence. Contemporary Readings in Philosophy and Science*. Cambridge, Massachussetts: MIT Press.

Crutchfield, J. P., Farmer, J. D., Packard, N. H., & Shaw, R. S. (1986). Chaos. *Scientific American*, *254*(12), 46–57.

Daniel Langlois Foundation. (2014). Robert Rauschenberg. Open Score (performance). Retrieved February 02, 2015, from http://www.fondation-langlois.org/html/e/page.php?NumPage=642

Dawkins, R. (2006). *The Blind Watchmaker*. London: Penguin Books.

Deneubourg, J. L., Goss, S., Franks, N., Sendova Franks, A., Detrain, C., & Chrétien, L. (1990). The dynamics of collective sorting robot-like ants and ant-like robots. In *Proceedings of the first international conference on simulation of adaptive behavior on From animals to animats* (pp. 356–363). Cambridge, Massachussetts: MIT Press.

Dix, A., Finlay, J., Abowd, G. D., & Beale, R. (2004). *Human-Computer Interaction*. Upple Saddle River, New Jersey: Pearson Education.

Dorin, A., & Knorb, K. B. (2012). Creativity Refined: Bypassing the Gatekeepers of Appropriateness and Value. In *Computers and Creativity* (pp. 339–360). New York: Springer.

Dorin, A., McCabe, J., McCormack, J., Monro, G., & Whitelaw, M. (2012). A framework for understanding generative art. *Digital Creativity*, *23*(3-4), 1–21.

Downes, E. J., & McMILLAN, S. J. (2000). Defining Interactivity: A Qualitative Identification of Key Dimensions. *New Media & Society*, *2*(2), 157–179. doi:10.1177/14614440022225751

Driessens, E., & Verstappen, M. (2006). Tickle robots. Retrieved September 01, 2014, from http://notnot.home.xs4all.nl/ticklerobots/ticklerobots.html

Driessens, E., & Verstappen, M. (2007). Breed. Retrieved September 02, 2014, from http://notnot.home.xs4all.nl/breed/Breed.html

Dubberly, H., & Pangaro, P. (2009). What is conversation? How can we design for effective conversation? *Interactions Magazine (ACM)*, *XVI*(4).

Dubberly, H., Pangaro, P., & Haque, U. (2009). What is interaction? Are there different types? *Interactions Magazine (ACM)*, *XVI*(1).

Durlak, J. T. (1987). A Typology for Interactive Media. In M. L. McLaughlin (Ed.), *Communication Yearboook* (Vol. 10, pp. 743–757). Newbury Park, CA: Sage.

Edmonds, E. (2010). The art of interaction. *Digital Creativity*, *21*(4), 257–264.

Encyclopædia Britannica Company. (2014). Emersion. Retrieved January 01, 2014, from http://www.merriam-webster.com/dictionary/emersion

Engelbart, D. (2001). Augmenting Human Intellect: A Conceptual Framework. In R. Packer & K. Jordan (Eds.), *Multimedia. From Wagner to Virtual Reality* (pp. 64–90). New York: Norton & Company.

Fererman, L. (1995). *Tierra, Evolution in Another Universe*. Retrieved from http://life.ou.edu/tierra/video.html

Fernández, M. (2008). Gordon Pask: Cybernetic Polymath. *Leonardo*, *41*(2), 162–168.

Flake, G. W. (1998). *The Computational Beauty of Nature*. Cambridge, Massachussetts: MIT Press.

Frasca, G. (2001). *Videogames Of The Oppressed: Videogames As A Means For Critical Thinking And Debate*. Georgia Institute of Technology. Retrieved from http://www.ludology.org/articles/thesis/FrascaThesisVideogames.pdf

Freixa i Font, P. (2009). Modelo para el découpage de audiovisuales interactivos. Retrieved from http://repositori.upf.edu/handle/10230/16197

Freixa i Font, P., Soler-Adillon, J., Sora, C., & Ribas, J. I. (2014a). Contributions of the interactive decoupage to reading and analyzing interactive audiovisual works in cybermedia. *Hipertext.net*, *12*. Retrieved from http://raco.cat/index.php/Hipertext/article/view/274410/364576

Freixa i Font, P., Soler-Adillon, J., Sora, C., & Ribas, J. I. (2014b). Snow Fall y A Short History of the Highrise: dos modelos de comunicación audiovisual interactiva del New York Time. In *XX Congreso Internacional de la Sociedad Española de Periodística*.

Freixa i Font, P., Sora, C., Soler-Adillon, J., & Ribas, J. I. (2014). El découpage interactivo: una propuesta metodológica para el estudio y análisis de aplicaciones audiovisuales interactivas. In *IV Congreso Internacional de la Asociación Española de Investigación de la Comunicación. AE-IC 2014 Bilbao "Espacios de comunicación."*

Fuller, M. (2000). The human cellular automata. Retrieved September 09, 2014, from http://www.spc.org/fuller/projects/the-human-cellular-automata/

Fundacion Telefónica. (2000). Ken Rinaldo Autopoiesis. Retrieved September 25, 2014, from http://www.fundacion.telefonica.com/es/at/vida/vida10/paginas/v3/rinaldo.html

Fundacion Telefónica. (2014). Rule 30. Retrieved September 02, 2014, from http://vida.fundaciontelefonica.com/en/project/rule-30/

Galanter, P. (2008a). Complexism and the role of evolutionary art. In J. Romero & P. Machado (Eds.), *Facing the future: evolutionary possibilites for Human-Computer Interaction* (pp. 311–332). Berlin: Springer.

Galanter, P. (2008b). What is complexism? Generative art and the cultures of science and the humanities. In *International conference on Generative Art. Generative art design lab*. Milano, Italy.

Galanter, P. (2010). The problem with evolutionary art is ... In *EvoCOMNET'10 Proceedings of the 2010 international conference on Applications of Evolutionary Computation* (pp. 321–330). Berlin: Springer-Verlag.

Galanter, P. (2012). Computational Aesthetic Evaluation: Past and Future. In J. McCormack & M. D'Inverno (Eds.), *Computers and Creativity* (pp. 255–293). New York: Springer.

Gao, Q., Rau, P.-L. P., & Salvendy, G. (2010). Measuring perceived interactivity of mobile advertisements. *Behaviour & Information Technology*, *29*(1), 35–44. doi:10.1080/01449290802666770

Gere, C. (2008). *Digital Culture*. London: Reaktion Books.

Gershenson, C., & Heylighen, F. (2003). When Can We Call a System Self-Organizing? In W. Banzhaf, T. Cristaller, P. Dittrich, J. T. Kim, & J. Ziegles (Eds.), *Advances in Artificial Life, 7th European Conference, ECAL 2003* (pp. 606–614). Springer.

Glynn, R. (2008). Conversational Environments Revisted. In *19th European Meeting of Cybernetics & Systems Research*. Viena.

Glynn, R. (2009). Performative ecologies. Retrieved September 03, 2014, from http://www.ruairiglynn.co.uk/portfolio/performative-ecologies/

Gordon, D. M. (2003). The Organization of Work in Social Insect Colonies. *Complexity*, *8*(1), 43–46. Retrieved from http://eclectic.ss.uci.edu/~drwhite/Complexity/

Gordon, D. M. (2010). *Ant Encounters: Interaction Networks and Colony Behavior*. Princenton, New Jersey: Princenton Univerisity Press.

Green, J. (2010). Interactivity and Agency in Real Time Systems. In *Soft Borders Conference & Festival Proceedings* (pp. 1–5). São Paulo, Brazil: Centro Universitário Belas Artes de São Paulo.

Greene, N. (1989). Voxel space automata: Modeling with stochastic growth processes in voxel space. *Computer Graphics*, *23*(4), 175–184.

Haugeland, J. (1998). Mind Embodied and Embeded. In *Having Thought* (pp. 207–237). Cambridge, Massachussetts: Harvard University Press.

Hayles, K. (2004). N. Katherine Hayles responds. *electronic book review*. Retrieved January 01, 2015, from http://www.electronicbookreview.com/thread/firstperson/intereactive

Heeter, C. (2000). Interactivity in the Context of Designed Experiences. *Journal of Interactive Advertising*, *1*(1), 3–14. doi:10.1080/15252019.2000.10722040

Helmrich, S. (2000). *Silicon Second Nature*. Berkeley: University of California Press.

Helmrich, S. (2001). After Culture: Reflections on the Apparition of Anthropology in Artificial Life, a Science of Simulation. *Cultural Antrhropology*, *16*(4), 613–628.

Hendriks-Jansen, H. (1994). In Praise of Interactive Emergence: Or Why Explanations Don''t Have to Wait for Implementations. In *Proceedings of the Fourth International Workshop on the SYnthesis and Simulation of Living Systems* (pp. 70–79).

Hertz, H. (1899). *The Principles of Mechanics Presented in a New Form*. London: Macmillan.

Heylighen, F. (2003). The Science of Self-organization and Adaptivity. In *Knowledge Management, Organizational Intelligence and Learning, and Complexity, in: The Encyclopedia of Life Support Systems*. EOLSS Publishers Co. Ltd.

Heylighen, F., & Joslyn, C. (2001). Cybernetics and Second-Order Cybernetics. In R. A. Meyers (Ed.), *Encyclopedia of Physical Science & Technology. Volume 4.* (pp. 155–170). New York: Academic Press.

Hodgson, G. (2011). Downward Causation - Some Second Thoughts. Retrieved from http://www.geoffrey-hodgson.info/downward-causation.htm

Holland, J. (1992). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. Cambridge, Massachussetts: MIT Press.

Holland, J. (1995). *Hidden Order. How Adaptation Builds Complexity*. Reading, Massachussetts: Addison-Wesley Publishing Company.

Holland, J. (1998). *Emergence. From Chaos to Order*. New York: Basic Books.

Holland, O. (1997). Grey Walter: The Pioneer of Real Artificial Life. In C. G. Langton & K. Shimohara (Eds.), *Artificial Life V: Proceedings of the Fifth International Workshop on the Synthesis and Simulation of Living Systems* (pp. 34–41). Cambridge, Massachussetts: MIT Press.

Holland, O. (2003). The first biologically inspired robots. *Robotica*, *21*, 351–363.

Ingold, T. (2001). Beyond art and technology: the antropology of skill. In M. B. Schiffer (Ed.), *Anthropological Perspectives on Technology* (pp. 17–32). Albuquerque: University of New Mexico Press.

Jee, J., & Lee, W.-N. (2002). Antecedents and Consequences of Perceived Interactivity. *Journal of Interactive Advertising*, *3*(1), 34–45. doi:10.1080/15252019.2002.10722066

Jensen, J. F. (1998). " Interactivity " Tracking a New Concept in Media and Communication Studies. *Media and Communication Studies*, *Nordicon R*(1), 185–204.

Jensen, J. F. (1999). The Concept of "Interactivity" in Interactive Television. In J. F. Jensen & C. Toscan (Eds.), *Interactive Television: TV of the Future or the Future of TV?* (pp. 25–66). Aalborg: Aalborg Universitetsforlag.

Jensen, J. F. (2008). The Concept of Interactivity – revisited Four new typologies for a new media landscape. In *Proceedings of the 1st international conference on Designing interactive user experiences for TV and video* (pp. 129–132).

Jerison, D., & Stroock, D. (1995). Norbert Wiener. *Notices of the American Mathematical Society*, *42*(4), 430–438.

Johnston, J. (2008). *The Allure of Machinic Life. Cybernetics, Artificial Life, and the New AI*. Cambridge, Massachussetts: MIT Press.

Juul, J. (2002). The Open and the Closed: Games of Emergence and Games of Progression. In *Computer Games and Digital Cultures Conference Proceedings* (pp. 323–329). Tampere.

Karamanos, N., & Gibbs, P. (2012). A model for student adoption of online interactivity. *Research in Post-Compulsory Education*, *17*(3), 321–334. doi:10.1080/13596748.2012.700108

Kauffman, S. A. (1984). Emergent properties in random complex automata. *Physica D: Nonlinear Phenomena*, *10*(1-2), 145–156.

Kauffman, S. A. (1993). *The Origins of Order. Self-Organization and Selection in Evolution*. New York: Oxford University Press.

Keller, E. F. (2003). *Making Sense of Life. Explaining Biological Development with Models, Metaphors, and Machines.* Cambridge, Massachusetts: Harvard University Press.

Keller, E. F. (2007). Slime Mold. In *Evocative Objects. Things We Think With* (pp. 297–306). Cambridge, Massachusetts: MIT Press.

Keller, E. F., & Segel, L. (1970). Initiation of slime mold aggregation viewed as an instability. *Journal of Theoretical Biology*, *26*(3), 399–415. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5462335

Kelso, J. A. S. (1995). *Dynamic Patterns. The Self-Organization of Brain and Behavior*. Cambridge, Massachussetts: MIT Press.

Kendall, K. E., & Kendall, J. E. (2006). *Systems Analysis and Design*. Boston: Pearson.

Kim, J. (1992). "Downward Causation" in Emergentism and Nonreductive Physicalism. In A. Beckermann, H. Flohr, & J. Kim (Eds.), *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism* (pp. 119–138). Berlin-New York: Walter de Gruyter.

Kim, J. (1999). Making sense of emergence. *Philosophical Studies*, 3–36. Retrieved from http://www.springerlink.com/index/r86gq867212xt8v2.pdf

Kim, J. (2006). Emergence: Core ideas and issues. *Synthese*, *151*(3), 547–559. doi:10.1007/s11229-006-9025-0

Kiousis, S. (2002). Interactivity: A concept explication. *New Media & Society*, *4*(3), 355–383.

Klir, G. J. (1968). *An Approach to General Systems Theory*. New York: Van Nostrand Reinhold.

Kluegl, F., Puppe, F., Raub, U., & Tautz, J. (1998). Simulating Multiple Emergent Phenomena - Exemplified in an Ant Colony. In *Artificial Life VI*.

Klütsch, C. (2005). The summer 1968 in London and Zagreb: starting or end point for computer art? In *Proceedings of the 5th conference on Creativity & cognition* (pp. 109–117).

Kowaliw, T., McCormack, J., & Dorin, A. (2011). An interactive electronic art system based on artificial ecosystemics. In *Artificial Life (ALIFE), 2011 IEEE Symposium on* (pp. 162 – 169). IEEE.

Krueger, M. (1991). *Artificial Reality II*. New York: Addison-Wesley Professional.

Kwa, C. (2002). Romantic and Baroque Conceptions of Complex Wholes in the Sciences. In J. Law & A. Mol (Eds.), *Complexities. Social Studies of Knowledge Practices.* (pp. 23–52). Durhan, North Carolina: Duke University Press.

Langton, C. G. (1984). Self-Reproduction in Cellular Automata. *Physica D: Nonlinear Phenomena*, *10*(1-2), 135–144.

Langton, C. G. (1986). Studying Artificial Life With Cellular Automata. *Physica D: Nonlinear Phenomena*, *22*(1-3), 120–149.

Langton, C. G. (1988). Artificial life. In C. G. Langton (Ed.), *Artificial Life: Proceedings of an Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems* (pp. 1–47). Boston: Addison-Wesley Publishing Company. Retrieved from http://library.iscpif.fr/files/CSSS2010-20100805-Beslon-lecture4.pdf

Langton, C. G. (1990). Computation at the edge of chaos: Phase transitions and emergent computation. *Physica D: Nonlinear Phenomena*, *42*(1-3), 12–37.

Langton, C. G. (1995). Editor's Introduction. In *Artificial life. An Overview* (pp. ix–xi). Cambridge, Massachussetts: MIT Press.

Latour, B. (1993). *We Have Never Been Modern*. Cambridge, Massachussetts: Harvard University Press.

LeBouthillier, A. E. (1999). W. Grey Walter and his Turtle Robots. *The Robot Builder*, *11*(5), 1–3.

Lee, K. M., Park, N., & Jin, S.-A. (2006). Narrative and interactivity in computer games. In P. Vorderer & J. Bryant (Eds.), *Playing video games: Motives, responses, and consequences* (pp. 259–274). Nahwah, NJ: Erlbaum.

Leiner, D. J., & Quiring, O. (2008). What Interactivity Means to the User Essential Insights into and a Scale for Perceived Interactivity. *Journal of Computer-Mediated Communication*, *14*(1), 127–155.

Levy, S. (1992). *Artificial Life. The quest for a new creation*. London: Jonathan Cape.

Licklider, J. C. R. (1960). Man-Computer Symbiosis. *IRE Transactions on Human Factors in Electronics*, *HFE-1*, 4–11.

Liu, Y., & Shrum, L. J. (2002). What Is Interactivity and Is It Always Such a Good Thing? Implications of Definition, Person, and Situation for the Influence of Interactivity on Advertising Effectiveness. *Journal of Advertising*, *31*(4), 53–64.

Lorenz, E. N. (1972). Texas?, Predictability: Does the Flap of a Butterfly's Wings in Brazil Set Off a Tornado in. Retrieved from http://eaps4.mit.edu/research/Lorenz/Butterfly_1972.pdf

MacGregor, B. (2002). Cybernetic serendipity revisited. In *Proceedings of the 4th conference on Creativity & cognition* (pp. 11–13).

Machotka, P. (1980). Daniel Berlyne's contributions to empirical aesthetics. *Motivation and Emotion*, *4*(2), 113–121.

Mandelbrot, B. (1967). How long is the coast of britain? Statistical self-similarity and fractional dimension. *Science*, *156*(3775), 636–638. doi:10.1126/science.156.3775.636

Mandelbrot, B. (1977). *The Fractal Geometry of Nature*. New York: WH Freeman and Company.

Manovich, L. (2001). *The Language of New Media*. Cambridge, Massachussetts: MIT Press.

Martínez, G. J., Seck-Tuoh-Mora, J. C., & Zenil, H. (2013). Computation and Universality: Class IV versus Class III Cellular Automata. *Journal of Cellular Automata*, *7*(5-6), 393–430.

Maturana, H. (1987). Everything is Said by an Oberver. In W. Thompson (Ed.), *Gaia: A Way of Knowing* (pp. 65–82). Hudson, NY: Lindisfarne Press.

Mayer, J., Khairy, K., & Howard, J. (2010). Drawing an elephant with four complex parameters. *American Journal of Physics*, *78*(6), 648–649.

McCormack, J. (2005). Open Problems in Evolutionary Music and Art. In F. Rothlauf, J. Branke, S. Cagnoni, D. W. Corne, R. Drechsler, Y. Jin, … G. Squillero (Eds.), *Lecture Notes in Computer Science [Proceedings of Applications of Evolutionary Computing, (EvoMUSART 2005)* (Vol. 3449, pp. 428–436). New York: Springer-Verlag.

McCormack, J. (2008). Facing the future: evolutionary possibilities for Human-Computer Interaction. In J. Romero & P. Machado (Eds.), *The art of artificial evolution: a handbook on evolutionary art and music* (pp. 417–451). Berlin: Springer.

McCormack, J. (2012). Creative Ecosystems. In *Computers and Creativity* (pp. 39–60). New York: Springer.

McCormack, J., & Dorin, A. (2001). Art, Emergence, and the Computational Sublime. In A. Dorin (Ed.), *Proceedings of Second Iteration: A Conference on Generative Systems in the Electronic Arts* (pp. 67–81). Melbourne, Australia.

McLaughlin, B. P. (2008). The Rise and Fall of British Emergentism. In M. A. Bedau & P. Humphreys (Eds.), *Emergence. Contemporary Readings in Philosophy and Science* (pp. 19–59). Cambridge, Massachusetts: MIT Press.

McNeil, I. (1990). *An Encyclopaedia of the History of Technology*. London: Routledge.

Meadows, D. H. (2009). *Thinking in Systems. A Primer*. (D. Wright, Ed.). London: Earthscan.

Mechant, P. (2012). An illustrated framework for the analysis of Web2.0 interactivity. *Contemporary Social Science*, *7*(3), 263–281. doi:10.1080/21582041.2012.716524

Medien Kunst Netz. (1994). Sommerer/Mignonneau «A-Volve». Retrieved September 03, 2014, from http://www.medienkunstnetz.de/works/a-volve/

Medien Kunst Netz. (2000). Ken Rinaldo «Autopoiesis». Retrieved September 25, 2014, from http://www.medienkunstnetz.de/works/autopoiesis/

Mill, J. S. (1882). *A System of Logic*. New York: Harper & Brothers. Retrieved from https://archive.org/details/systemofratiocin00milluoft

Mindell, D. A. (2002). *Between Human and Machine. Feedback, Control and Computing before Cybernetics.* Baltimore, Maryland: The Johns Hopkins University Press.

Moggridge, B. (2007). *Designing Interactions*. Cambridge, Massachussetts: MIT Press.

Monro, G. (2009). Emergence and Generative Art. *Leonardo, 42*(9), 446–447.

Morgan, C. L. (1927). *Emergent Evolution*. London: Williams and Norgate. Retrieved from https://archive.org/details/emergentevolutio032372mbp

Morris, C. (1946). *Signs, Language and Behavior*. New York: George Braziller.

Murphie, A., & Potts, J. (2003). *Culture & Technology*. New York: Palgrave.

Myskja, K. (2008). Rule 30. Retrieved September 02, 2014, from http://kristoffermyskja.com/work.php

Nagel, E. (1961). *The Structure of Science: Problems in the Logic of Scientific Explanation*. New York: Harcourt, Brace & World, Inc.

Nake, F. (1971). There Should Be No Computer Art. *Bulletin of the Computer Arts Society*, 18–19.

Nake, F. (2012). Construction and Intuition: Creativity in Early Computer Art. In *Computers and Creativity* (pp. 61–94). New York: Springer.

Newhagen, J. E. (2004). Interactivity, Dynamic Symbol Processing, and the Emergence of Content in Human Communication. *The Information Society*, *20*(5), 395–400. doi:10.1080/01972240490508108

Noble, J. (2009). *Programming Interactivity. A Designer's Guide to Processing, Arduino, and OpenFramework.* Sebastopol, CA: O'Reilly Media.

Norman, D. (1988). *The Design of Everyday Things*. New York: Doubleday.

O'Brien, F. (1967). *The Third Policeman*. Dublin, Ireland: Dalkey Archive Press.

O'Sullivan, T., Hartley, J., Saunders, D., Montgomery, M., & Fiske, J. (1994). *Key Concepts in Communication and Cultural Studies*. London: Routledge.

Oxford Dictionaries. (2014). Emersion. Retrieved January 01, 2014, from http://www.oxforddictionaries.com/es/definicion/ingles_americano/emersion

Pangaro, P. (2009). How Can I Put That? Applying Cybernetics to "Conversational Media." In *Proceedings of the Annual Meeting of the American Society for Cybernetics* (pp. 1–16).

Parés, N., & Pares, R. (2001). Interaction-Driven Virtual Reality Application Design A Particular Case: El Ball del Fanalet or Lightpools. *Presence*, *10*(2), 236–245.

Pask, G. (1959). Physical Analogues to the Growth of a Concept. In A. Uttley (Ed.), *Mechanisation of Thought Processes* (pp. 877–922). London: HMSO.

Pask, G. (1960). The natural history of networks. In M. C. Yovits & S. Cameron (Eds.), *Self Organising Systems* (pp. 232–261). London: Pergamon Press.

Pask, G. (1971). A Comment, a Case History and a Plan. In J. Reichardt (Ed.), *Cybernetics, Art and Ideas1* (pp. 76–99). London: Studio Vista.

Pask, G. (1975). *Conversation, Cognition and Learning. A Cybernetic Theory and Methodology*. New York: Elsevier.

Pask, G. (1976). *Conversation Theory: Applications in Education and Epistemology*. New York: Elsevier.

Pask, G. (1996). Heinz von Foerster's Self-Organisation, the Progenitor of Conversation and Interaction Theories. *Systems Research*, *13*(3), 349–362.

Pattee, H. H. (1982). Cell Psychology: An Evolutionary Approach to the Symbol-Matter Problem. *Cognition and Brain Theory*, *5*(4), 325–341.

Pattee, H. H. (1985). Universal principles of measurement and language functions in evolving systems. In J. L. Casti & A. Karlqvist (Eds.), *Complexity, Language, and Life: Mathematical Approaches* (pp. 268–281). Berlin: Springer-Verlag.

Pattee, H. H. (1987). Instabilities and Information in Biological Organization. In F. E. Yates (Ed.), *Self-Organizing Systems: The Emergence of Order* (pp. 325–388). New York: Plenum Press.

Pattee, H. H. (1989). Simulations , Realizations , and Theories of Life. In C. G. Langton (Ed.), *Artificial Life* (Vol. 1989, pp. 63–77). Redwood City, CA: Addison-Wesley Publishing Company.

Paul, C. (2003). *Digital Art*. New York: Thames & Hudson.

Penny, S. (1995). Sympathetic Sentience II. Retrieved November 20, 2013, from http://simonpenny.net/works/sympatheticII.html

Penny, S. (1996). The Darwin Machine: Artificial Life and Interactive Art. *New Formations UK*, *29*, 59–68.

Penny, S. (2000). Agents as Artworks: and Agent Design as Artistic Practice. In K. Dauthenhahn (Ed.), *Human Cognition and Social Agent Technology* (pp. 395–414). Amsterdam, Netherlands: John Benjamins Publishing Company.

Penny, S. (2008). Robotic art and interactive installation 1988-1997. DVD documentation.

Penny, S. (2009). Art and Artificial Life -A primer. In *Proceedings of the Digital Arts and Culture Conference*. University of California, Irvine. Retrieved from http://escholarship.org/uc/item/1z07j77x

Penny, S. (2010). Twenty Years of Artificial Life. *Digital Creativity*, *21*(3), 197–204.

Penny, S. (2011). Towards a Performative Aesthetics of Interactivity. *Fiberculture*, *19*, 72–108.

Penny, S. (2013a). *The aesthetics of behaving systems*. Unpublished.

Penny, S. (2013b). Trying to be calm. Ubiquity, cognitivism, and embodyment. In U. Ekman (Ed.), *Throughout. Art and culture emerging with ubiquitous computing* (pp. 263–278). Cambridge, Massachussetts: MIT Press.

Perkins, D. N. (1996). Creativity: beyond the Darwinian paradigm. In M. A. Boden (Ed.), *Dimensions of creativity* (pp. 119–142). Cambridge, Massachussetts: MIT Press.

Pickering, A. (1995). *The Mangle of Practice: Time, Agency, and Science*. Chicago: The University of Chicago Press.

Pickering, A. (2002). Cybernetics and the Mangle: Ashby, Beer and Pask. *Social Studies of Science*, *32*(3), 413–437.

Pickering, A. (2008). Ontological Theatre. Gordon Pask, Cybernetics, and the Arts. *Cybernetis And Human Knowing*, *14*(4), 43–57.

Pickering, A. (2010). *The Cybernetic Brain. Sketches of Another Future*. Chicago: The University of Chicago Press.

Prigogine, I., & Stengers, I. (1984). *Order out of Chaos*. New York: Bantam Books.

Prophet, J. (1996). Sublime Ecologies and Artistic Endeavors: Artificial Life and Interactivity in the Online Project “TechnoSphere.” *Leonardo*, *29*(5), 339–344.

Prophet, J. (2001). “TechnoSphere”: “Real” Time, “Artificial” Life. *Leonardo*, *34*(4), 309–312.

Prophet, J., & Shelley, G. (1995). Techosphere. Retrieved September 01, 2014, from http://www.janeprophet.com/old-website/technoweb.html

Protevi, J. (2006). Deleuze, Guattari and Emergence. *Paragraph*, *29*(2), 19–39. doi:10.1353/prg.2006.0018

Prudence, P. (2010). Breed – Driessens & Verstappen: Evolutional Diffusion Lattices. Retrieved September 02, 2104, from http://www.dataisnature.com/?p=565

Prusinkiewicz, P., & Lindenmayer, A. (2004). *The algorithmic beauty of plants*. New York: Springer-Verlag.

Quiring, O. (2009). What do users associate with “interactivity”?: A qualitative study on user schemata. *New Media & Society*, *11*(6), 899–920. doi:10.1177/1461444809336511

Rafaeli, S. (1988). Interactivity: From new media to communication. *Sage Annual Review of Communication Research: Advancing Communication Science*, *16*, 110–134.

Rafaeli, S., & Sundweeks, F. (1997). Networked Interactivity. *Journal of Computer-Mediated Communication*, *2*(4), 0–0.

Ramirez, Jr, A., & Burgoon, J. K. (2004). The effect of interactivity on initial interactions: the influence of information valence and modality and information richness on computer-mediated interaction. *Communication Monographs*, *71*(4), 422–447. doi:10.1080/0363452042000307461

Reichardt, J. (1968). *Cybernetic Serendipity. The computer and the arts*. London: International, Studio.

Reichardt, J. (1971). Cybernetics, art and ideas. In *Cybernetics, Art and Ideas* (pp. 11–17). New York: New York Graphic Society.

Resnick, M. (1995). Learning About Life. In C. G. Langton (Ed.), *Artificial life. An Overview* (pp. 229–242). Cambridge, Massachussetts: MIT Press.

Reynolds, C. (1995). Boids (Flocks, Herds, and Schools: a Distributed Behavioral Model). Retrieved August 18, 2013, from http://www.red3d.com/cwr/boids/index.html

Reynolds, C. (2011). Interactive Evolution of Camouflage. *Artificial Life*, *17*(2), 123–136.

Ribas, J. I. (2001). Difusió cultural i comunicacio audiovisual interactiva. *Temes de Disseny*, *18*, 170–204.

Ribas, J. I. (2010). Interactive discourse in cultural dissemination: Innovation and electronic readers. *Hipertext.net2, 8*. Retrieved from http://www.upf.edu/hipertextnet/en/numero-8/discourse_e-readers.html

Rinaldo, K. (2000). Autopoiesis: Artificial life robotic sculpture installation. Retrieved September 03, 2014, from http://www.ylem.org/artists/krinaldo/works/autopoiesis/autopoiesis.html

Rinaldo, K. E. (1998). Technology Recapitulates Phylogeny: Artificial Life Art. *Leonardo*, *31*(5), 370.

Riskin, J. (2003a). Eighteenth-Century Wetware. *Representations*, (83), 97–125.

Riskin, J. (2003b). The defecating duck, or, the ambiguous origins of Artificial Life. *Critical Inquiry*, *29*(4), 599–633. doi:10.1086/377722

Rocha, L. M. (1998). Selected Self-Organization And the Semiotics of Evolutionary Systems. In S. Salthe, G. Van de Vijver, & M. Delpos (Eds.), *Evolutionary Systems: Biological and Epistemological Perspectives on Selection and Self-Organization* (pp. 341–358). Dordrecht, Holland: Kluwer Academic Publishers.

Roland, E. M. A., Sipper, M., & Capcarrère, M. S. (1999a). Design, Observation, Surprise! A Test of Emergence. *Artificial Life*, *5*, 225–239.

Roland, E. M. A., Sipper, M., & Capcarrère, M. S. (1999b). Testing for emergence in artificial life. In D. Floreano, J. Nicoud, & F. Mondada (Eds.), *Advances in Artificial Life: Proceedings of the 5th European Conference on Artificial Life (ECAL'99)* (pp. 13–20). Heidelberg: Springer-Verlag. Retrieved from http://link.springer.com/chapter/10.1007/3-540-48304-7_5

Rosen, R. (1985). *Anticipatory Systems. Philosophical, Matehmatical and Methodological Foundations.* New York: Pergamon Press.

Rosen, R. (2000). *Essays on Life Itself*. New York: Columbia University Press.

Rosen, R. (2012). *Anticipatory Systems. Philosophical, Matehmatical and Methodological Foundations. 2nd Edition.* (2nd ed.). New York: Springer.

Rosenblueth, A., Wiener, N., & Bigelow, J. (1943). Behavior, Purpose and Teleology. *Philosophy of Science*, *10*(1), 18–24.

Salen, K., & Zimmerman, E. (2004). *Rules of Play - Game Design Fundamentals*. Cambridge, Massachussetts: MIT Press.

Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, *3*(3), 210–229.

Samuel, A. L. (1967). Some Studies in Machine Learning Using the Game of Checkers. II-Recent Progress. *IBM Journal of Research and Development*, *11*(6), 601 – 617.

Schmidhuber, J. (2012). A Formal Theory of Creativity to Model the Creation of Art. In *Computers and Creativity* (pp. 323–338). New York: Springer.

Schneiderman, B. (1993). *Designing the User Interface*. Reading, Massachussetts: Addison-Wesley.

Seaman, B. (2003). Recombinant Poetics - Media-Element Field Explorations. In N. Wardrip-Fruin & P. Harrigan (Eds.), *First Person: New Media as Story, Performance, and Game* (pp. 223–227). Cambridge, Massachussetts: MIT Press.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, *3*(3), 417–457. doi:10.1017/S0140525X00005756

Seelig, M. (2008). Interactivity on Traditional Media Web Sites. *Journal of Radio & Audio Media*, *15*(2), 231–248. doi:10.1080/19376520802397433

Seevinck, J., & Edmonds, E. (2008). Emergence and the art system "plus minus now." *Design Studies*, *29*(6), 541–555.

Seevinck, J., & Edmonds, E. (2009). Open in art, nature and emergence. In *Proceedings of the 21st Australasian Computer-Human Interaction Conference*.

Shanken, E. A. (1998a). Life as We Know It and / or Life as It Could Be: Epistemology and the Ontology / Ontogeny of Artificial Life. *Leonardo*, *31*(5), 383–388.

Shanken, E. A. (1998b). The House That Jack Built. *Leonardo Electronic Almanac*, *6*(10), 1969–1971.

Shanken, E. A. (2002). Cybernetics and Art: Cultural Convergence in the 1960s. In B. Clarke & L. D. Henderson (Eds.), *From Energy to Information* (pp. 155–177). Palo Alto: Stanford University Press.

Sharp, H., Roger, Y., & Preece, J. (2007). *Interaction Design. Beyond human-computer interaction*. West Sussex, England: John Wiley and Sons.

Shaw, J. (1994). The Legible City. Retrieved November 11, 2014, from http://www.jeffrey-shaw.net/html_main/frameset-works.php

Sierra Díaz, A. (2003). Lexicón Teofilolingüístico. Bogotá, Colombia: Publicaciones Cristianas. Retrieved from http://cristiania.net/LECTURAS/ARCADIO SIERRA/Lexicon.pdf

Simon, H. A. (1996). *The Sciences of the Artificial*. Cambridge, Massachussetts: MIT Press.

Sims, K. (1991a). Artificial Evolution for Computer Graphics. *Computer Graphics*, *25*(4), 319–328.

Sims, K. (1991b). Evolved Images. Retrieved September 01, 2014, from http://www.karlsims.com/genetic-images.html

Sims, K. (1993). Genetic images. Retrieved October 09, 2014, from http://www.karlsims.com/genetic-images.html

Sims, K. (1994a). Evolved Virtual Creatures. Retrieved March 02, 2010, from http://www.karlsims.com/evolved-virtual-creatures.html

Sims, K. (1994b). Evolving 3D Morphology and Behavior by Competition. In M. Brooks & P. Maes (Eds.), *Artificial Life IV Proceedings* (pp. 28–39). MIT Press.

Sims, K. (1994c). Evolving Virtual Creatures. In *Computer Graphics (Siggraph '94 Proceedings)* (pp. 15–22).

Smith, H. (2001). The Future of Game Design: Moving Beyond Deus Ex and Other Dated Paradigms. Retrieved from http://www.witchboy.net/articles/the-future-of-game-design-moving-beyond-deus-ex-and-other-dated-paradigms/

Sobecka, K. (2009). Sniff. Retrieved September 01, 2014, from http://www.gravitytrap.com/artwork/sniff

Solé, R. (2012). *Vidas Sintéticas*. Barcelona: Tusquets.

Solé, R., & Goodwin, B. (2001). *Signs of Life. How Complexity Pervades Biology*. New York: Basic Books.

Soler-Adillon, J. (2010a). Emergence and interactivity: A-Life Art as a paradigm for the creation of experiences in interactive communication. *Hipertext.net,* (8). Retrieved from http://www.upf.edu/hipertextnet/en/numero-8/a-life_art.html

Soler-Adillon, J. (2010b). Emergence in interactive media. An exploration of complex interactions in Artificial-Life Art. PhD Thesis proposal.

Soler-Adillon, J. (2011). Creating Black Boxes: Emergence and Interactive Art. In *Proceedings of the 17th International Conference on Electronic Art*. Retrieved from http://isea2011.sabanciuniv.edu/paper/creating-black-boxes-emergence-interactive-art

Soler-Adillon, J. (2012). Principios de diseño de interacción para sistemas interactivos. Retrieved from http://repositori.upf.edu/handle/10230/21513

Soler-Adillon, J., Freixa i Font, P., Sora, C., & Ribas, J. I. (2014). Towards a conceptualization of the emergent properties of Twitter as an information and news sharing environment. In *XX Congreso Internacional de la Sociedad Española de Periodística*.

Soler-Adillon, J., & Penny, S. (2014). Self-organization and novelty: pre-configurations of emergence in early British Cybernetics. In *Norbert Wiener in the 21st Century (21CW), 2014 IEEE Conference on* (pp. 1–8). Boston, MA: IEEE. doi:10.1109/NORBERT.2014.6893946

Sommerer, C., & Mignonneau, L. (1995). Phototropy. Retrieved November 11, 2014, from http://www.interface.ufg.ac.at/christa-laurent/WORKS/CONCEPTS/PhotoConcept.html

Sommerer, C., & Mignonneau, L. (2006). Life Writer. Creating life through the act of writing. Retrieved September 03, 2014, from http://www.interface.ufg.ac.at/~christa-laurent/WORKS/CONCEPTS/LifeWriterConcept.html

Sommerer, C., & Mignonneau, L. (2009). A-Volve. In G. Stocker, C. Sommerer, & L. Mignonneau (Eds.), *Christa Sommerer and Laurent Mignonneau. Interactive art research* (pp. 72–87). New York: Springer.

Sommerer, C., & Mignonneau, L. (2012). A-Volve exhibitions. Retrieved September 02, 2014, from http://www.interface.ufg.ac.at/christa-laurent/EXHIBITIONS/A-VolveExhib.html

SPECS. (2008). The Synthetic Oracle. Retrieved November 21, 2014, from http://specs.upf.edu/installation/549

Standage, T. (2002). *The Turk: The Life and Times of the Famous Eighteenth-Century Chess-Playing Machine*. New York: Walker Publishing.

Steels, L. (1995). The Artificial Life Roots of Artificial Intelligence. In C. G. Langton (Ed.), *Artificial life. An Overview1* (pp. 75–110). Cambridge, Massachussetts: MIT Press.

Steuer, J. (1992). Defining Virtual Reality: Dimensions determining telepresence. *Journal of Communication*, *42*(4), 73–93.

Stromer-Galley, J. (2004). Interactivity-as-Product and Interactivity-as-Process. *The Information Society*, *20*(5), 391–394. doi:10.1080/01972240490508081

Sundar, S. S. (2004). Theorizing Interactivity's Effects. *The Information Society*, *20*(5), 385–389. doi:10.1080/01972240490508072

Sundar, S. S., & Kim, J. (2005). Interactivity and Persuasion. *Journal of Interactive Advertising*, *5*(2), 5–18. doi:10.1080/15252019.2005.10722097

Tarbell, J. (2003). Substrate. Retrieved September 01, 2014, from http://www.complexification.net/gallery/machines/substrate/

Thompson, A. (2002). Notes on Design Through Artificial Evolution: Opportunities and Algorithms. In I. C. Parmee (Ed.), *Adaptive Computing in Design and Manufacture V* (pp. 17–36). London: Springer.

Todd, S., & Latham, W. (1992). *Evolutionary Art and Computers*. Waltham, Massachusetts: Academic Press.

Trask, M. (1971). *The Story of Cybernetics*. London: Studio Vista.

Turing, A. M. (1936). On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society, 42*, 230–265.

Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind, 59*, 433–460.

Umpleby, S. A. (2005). A History of the Cybernetics Movement in the United States. *Journal of the Washington Academy of Sciences*, *91*(2), 54–66.

Umpleby, S. A. (2009). Ross Ashby's general theory of adaptive systems. *International Journal of General Systems*, *38*(2), 231–238.

Van Riel, R., & Van Gulick, R. (2014). Scientific Reduction. In *The Stanford Encyclopedia of Philosophy (Summer 2014 Edition), Edward N. Zalta (ed.)*. Retrieved from http://plato.stanford.edu/archives/sum2014/entries/scientific-reduction/

Varela, F. J., Thompson, E., & Rosch, E. (1993). *The Embodied Mind. Cognitive Science and Human Experience*. Cambridge, Massachussetts: MIT Press.

Vaughan, M. (2007). Introduction: Henri Bergson's Creative Evolution. *Substance*, *36*(114), 7–24.

Ventrella, J. (1994). Explorations in the Emergence of Morphology and Locomotion Behavior in Animated Characters. In R. Brooks & P. Maes (Eds.), *Artificial Life IV Proceedings* (pp. 436–441).

Viveros-Fauné, C. (2001). Still Misbehaving. Perry Hoberman's Medium-Tech, Hi-Fi Excorcism of Technoculture. In P. Henry (Ed.), *Symptomatic. Recent Works by Perry Hoberman* (pp. 42–53). Bradford: University of Bradford.

Von Foerster, H. (1960). On Self-Organizing Systems and Their Environments. In M. C. Yovits & S. Cameron (Eds.), *Self-Organizing Systems* (pp. 31–50). London: Pergamon Press.

Von Neumann, J., & Burks, A. W. (1966). *Theory of self-reproducing automata*. Champaign, IL: University of Illinois Press. Retrieved from http://dl.acm.org/citation.cfm?id=1102024

Wadrip-Fruit, N., & Montfort, N. (2003). *The New Media Reader*. Cambridge, Massachussetts: MIT Press.

Waldrop, M. M. (1992). *Complexity. The Emerging Science at the Edge of Order and Chaos*. New York: Simon and Shuster.

Walter, W. G. (1950). An Imitation of Life. *Scientific American*, *182*(5), 42–45.

Walter, W. G. (1951). A Machine That Learns. *Scientific American*, *185*(2), 60–63.

Walter, W. G. (1953). *The Living Brain*. London: Duckworth.

Walter, W. G. (1963). The development and significance of cybernetics. *Anarchy 25. A Journal of Anarchist Ideas, 3*(3), 75–89.

Watz, M. (2010). Ralf Baecker – The Conversation. Retrieved September 01, 2014, from http://www.generatorx.no/20101130/abstrakt-abstrakt-ralf-baecker-the-conversation/

Weber, R., Behr, K.-M., & DeMartino, C. (2014). Measuring Interactivity in Video Games. *Communication Methods and Measures*, *8*(2), 79–115. doi:10.1080/19312458.2013.873778

Whitelaw, M. (2006). *Metacreation. Art and Artificial Life*. Cambridge, Massachussetts: MIT Press.

Whitlaw, M. (2013). Proto-computing. An interview with Ralf Baecker. Retrieved September 01, 2014, from http://teemingvoid.blogspot.com.es/2013/06/proto-computing-interview-with-ralf.html

Wiener, N. (1989). *The Human Use of Human Beings. Cybernetics and Society*. London: Free Association Books.

Wilson, S. (1994). The aestheteics and practice of designing interactive computer events. In *SIGGRAPH '94 Visual Proceedings*. ACM Press.

Wolfram, S. (1983). Statistical Mechanics of Cellular Automata. *Reviews of Modern Physics*, *55*(3), 601–644.

Wolfram, S. (1984). Universality and Complexity in Cellular Automata. *Physica D: Nonlinear Phenomena*, *10*(1-2), 1–35.

Wolfram, S. (2002). *A New Kind of Science*. Champaign, IL: Wolfram Media.

Wu, G. (1999). Perceived interactivity and attitude toward web sites. In *Proceedings of the 1999 Conference of the American Academy of Advertising*.

Wu, G. (2005). The Mediating Role of Perceived Interactivity in the Effect of Actual Interactivity on Attitude Toward the Website. *Journal of Interactive Advertising*, *5*(2), 29–39. doi:10.1080/15252019.2005.10722099

Yun, G. W. (2007). Interactivity Concepts Examined: Response Time, Hypertext, Role Taking, and Multimodality. *Media Psychology*, *9*(3), 527–548. doi:10.1080/15213260701283145