Chapter 14

Emergence in multi-agent systems: conceptual and methodological issues

14.1. Introduction

The notion of emergence has several meanings. In the vernacular language, emergence denotes both a gradual beginning or coming forth, or a sudden uprising or appearance; to emerge also means to become visible; for example, emergence may denote the act of rising out of a fluid. This latter sense is close to its Latin roots, where *emergere* is the opposite of *mergere*: to be submerged. In what follows, we relate the "act of rising out" to the arising of some phenomenon from a *process*, and note the fact that to become visible presupposes some observer. So, the common sense of emergence is linked with the meaning of a process that produces some phenomenon that might be detected by an observer.

The concept of Emergence first discussed in philosophy, is also widely used in complex adaptive systems literature, especially in computer sciences [HOL 98] and related fields (multi-agent systems, artificial intelligence, artificial life...) as well as in physics, biology, and cognitive sciences. In a pioneering book on artificial society and multi-agent simulations in social sciences, Nigel Gilbert put the emphasis on "emergence" as a key concept of this approach: "Emergence is one of the most interesting issues to have been addressed by computer scientists over the past few years and has also been a matter of concern in a number of other disciplines, from biology to political science" [GIL 95b, p.8]. Later, [GIL 02] discussed the varieties of emergent

Chapter written by Jean Louis Dessalles, Jean Pierre Müller and Denis Phan.

phenomena by developing Schelling's Model of residential segregation as a case study. He put the emphasis on methodological problems linked with emergence in artificial society, such as social recognition, construction of categories, second order emergence and downward causation. These latter terms concern the process through which the macro-level emerging social structure "feedbacks" into the micro-level by re-shaping the "elementary" behaviour (also called "immergence" by [GIL 95a]. According to [CAS 98a], [CAS 98b] such a downward social determination is linked with cognitive agents: the agents' minds: works as cognitive mediators for the social action [CAS 00]. From a methodological point of view, emergence provides therefore concepts that allow us to encompass both holism and a reductionist version of methodological individualism. For instance, Sawyer, starting from a discussion of the emergentist approaches in philosophy, psychology and sociology [SAW 01a], [SAW 02a] advocates a "non-reductive individualist" approach of emergence [SAW 02b], [SAW 03], that allows us to understand societies as complex systems [SAW 04], [SAW 05]. Unfortunately, these approaches remain mainly at the conceptual level and do not develop a formal framework both operative in multi-agent frameworks and that could be linked in some way with more traditional models. This latter requirement allows us to make comparative appraisal and to provide an account of how emergence could effectively break new ground and support new conceptual and formal advance. Such an assessment would be easier with the models of the economics side (*Agent based Computational Economics* - ACE) which can be more easily related with traditional formalisms. This is the case for instance of the ACE model of ([AXT 01] – hereafter: AEY) which is an extension of a population game by [YOU 98].

The first section raises critical questions about emergence from two paradigmatic examples of emergent phenomenon resulting from local social interaction (Schelling's model of Segregation and the RY model of emergence of classes). The second section briefly reviews and discusses some conceptual or formal definitions of emergence from both Philosophy and Computer sciences, with a special attention for multi-agent systems [MUL 04]. The third section provides a complementary definition proposed by [BON 97] and discussed in [DES 05], [DES 07] to be *operative* in multi-agent frameworks and to *make* sense from both a *cognitive* and a *social* point of view. This definition is coherent with important related features, like cognitive hierarchy, detection, and complexity

14.2. Emergence as a bottom-up process in a multi-agent framework

In ACE and Computational Social Sciences as well, emergence is strongly related to the *Santa Fe Approach to Complexity* (SFAC). SFAC calls "emergence" the arising at the macro level of some patterns, structures and properties of a complex adaptive system that are not contained in the properties of its parts.

Interactions between parts of a complex adaptive system play a key role in both complex dynamics and emergence. An interesting part of the emergence process concerns the forming of some *collective order* (coherent structures or patterns at the macro level) as a result of the agents' interactions within the system's dynamics. The two following cases are paradigmatic examples for the occurrence of a collective order from the agents' interactions. Both examples provide the basis for raising questions about the nature of emergence.

14.2.1. *What does emerge in Schelling's model of Segregation?*

Schelling's model of spatial segregation [SCH 69] [SCH 78] introduced by Gilbert in Chapter 5 and extensively studied by Daudé and Langlois in Chapter 13 is a pioneering example of an emerging phenomenon resulting from local social interaction. Schelling's aim was to explain how segregationist residential structures could spontaneously occur, even when people are not so very segregationist themselves. The absence of a global notion of segregationist structures (like the notion of ghettos) in the agent's attributes (preferences) is a crucial feature of this model. Agents do not choose between living or not living in a segregationist structure, but have only local preferences concerning their preferred neighbourhood. Moreover, people have only a weak segregationist behaviour, but the game of interactions generates global segregation.

Under Schelling's behavioural assumption, the "fully integrated structure" (where agents of different colours alternate in all directions) is an *unstable equilibrium*. A slight perturbation is sufficient to induce a *chain reaction* and the emergence of *local segregationist patterns*. Local interactions are sufficient to generate spatial homogeneous patterns. Then, spatial segregation is an *emerging property* of the system's dynamics, while not being an attribute of the individual agents. Sometimes, local integrated (non-homogeneous) patterns may survive in some niches. But such integrated structures could be easily perturbed by random changes, while homogeneous structures (frozen zones) are more stable[1]. Independently of the question of the empirical relevance of Schelling's model (see [SUG 02] and Appendix 2), this pioneering work is generally viewed as a paradigmatic example of the first generation of agent-based models, producing macro-social effects from the bottom-up [GIL 02]. However, some fundamental questions about the emergence properties of this model remain unresolved. We claim that the presence of an external observer being able to discern emergent phenomena and levels of organization is unavoidable. Accordingly, who is this

---

[1] Complementary theoretical developments on Schelling's model of segregation can be found in the growing literature on this subject (see [DES 05] for further references)

observer? From the point of view of Social sciences, what does the higher level of organization consist in? For whom does this level make sense? (Figure 14.1).
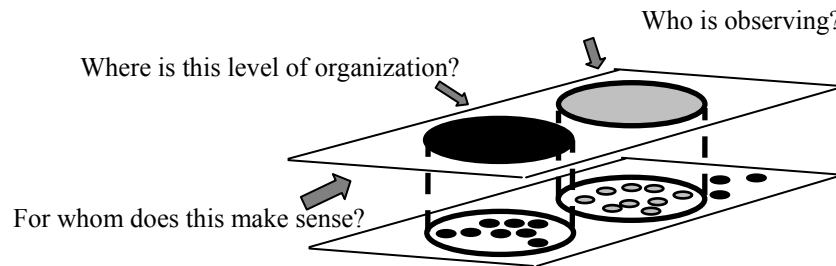
Source: [DES 05]

Figure 14.1: *The questions of emergence in the Social Sciences*

For the observer (i.e. the computational social scientist) this collective order makes sense by itself and opens up a radically new global interpretation, because this does not initially make sense as an attribute of the basic entities

14.2.2. *Emergence of "classes" in the ACE model of Axtell, Epstein and Young*

This model is a "random pair wise" type of population game [YOU 98], [BLU 97] with linear trembling hand. That is to say, Nash equilibrium can be reached without any assumption about common knowledge. Early analytical results can be found in [YOU 93]. During the game, agents are randomly paired and at each time step play a "one-shot" game with their opponent.

Agents choose the strategy which is their best response according to their beliefs (mixed strategy) about the behaviour of the others, drawn by induction from a distribution of observed strategies kept in a finite memory of size m. At each time steps, agents change partners and update their beliefs according to the result of the last meeting. Agents have a linear positive probability of deviation (trembling hand). The formal context is thus stochastic and the concept of stability used by the authors is due to Foster and Young [FOS 90] for stochastic evolutionary games.

The one-shot negotiation between pairs of agents is drawn from the one step Nash bargaining model. That is to say, each player tries to share a "cake" of size 100 with its opponents by opting for one of three possible strategies: "High" (H), "Medium" (egalitarian) (M), and "Low" (L). The corresponding percentages of the cake claimed by players can be fixed, without loss of generality, to 70%, 50% and 30% respectively. Only the couples of proposals which total is less or equal to 100 are accepted, other couples of strategies leading to null payoffs.

	S2 = H	S2 = M	S2 = L
S1 = H	(0,0)	(0,0)	(70,30)
S1 = M	(0,0)	(50, 50)	(50,30)
S1 = L	(30,70)	(30,50)	(30,30)

Player 1 in lines, player 2 in columns

Figure 14.2: *Nash bargaining game: best reply equivalent matrix for a bilateral game of agent*

The authors distinguish three situations. First, there are three situations where the agents' payoff is null and consequently highly inefficient (because they asked for more than 100% of the whole). Second, there are three situations where the whole cake is allocated: three Nash equilibria in pure strategy, with one equitable balance - (M, M) - and two inequitable ones (H,L), (L,H). Third, there are three inefficient outcomes where both payers have positive payoffs but not the whole cake is allocated.

The frequency of the strategies played within the population is given by a triplet $\sigma = (p, q, 1-p-q)$, and each agent infers its expected payoff from a historical sample of size m, as: $\sigma_{im} = (p_i, q_i, 1-p_i-q_i)$. At each time step, randomly paired agents play their best response against their own expected mixed strategy σ_{im} with a probability $(1-\varepsilon)$ and play at random with a probability ε linear trembling hand "*à la* Young". When playing, an agent observes the strategy of its opponent and updates its belief by removing from its memory the oldest value and by updating its list including the last strategy observed. The state of belief of an agent can be represented by a point on the simplex of size 2 used to represent the expected mixed strategies of this game (Figure 14.3).

The initial beliefs can be initialized in a random way, or in a targeted zone. But an initial form of heterogeneity of the beliefs is necessary to usefully explore the dynamic properties of this model. Indeed let us suppose that all the agents initially form the belief that their opponents play M. Their best reply, conditionally to this initial belief will precisely consist in playing M, which will reinforce the overall initial belief of these agents. The initial beliefs of an agent can be interpreted as its "cultural" heritage and its updated beliefs as the product of "the history" of its last meetings (a "historical" form of interactional heterogeneity, since agents' histories differ from one to another). Let us note that in the AEY model, the agents do not have common beliefs nor beliefs upon the beliefs of other agents, but only about the distribution of strategies. When the beliefs of the agents are located in the same zone of the simplex (for example, "M"), they are in a weak sense "shared beliefs", because their best response is the same "M", but this situation is not recursive and

agents have not common beliefs. Finally, using the results of Young [YOU 93], the authors show that the only stochastically stable solution [FOS 90] corresponds to a situation where almost all the agents play "M".
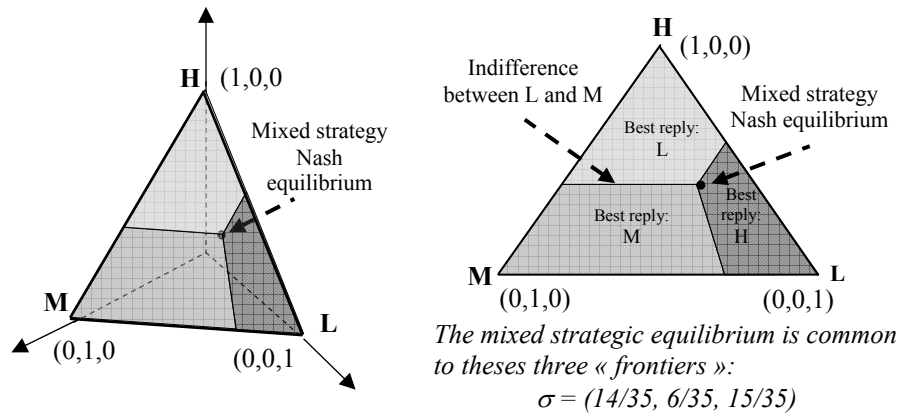

Figure 14.3: *Agent's belief representation in a simplex*

In a second time, the authors introduce two types of agents, differentiated by an observable external sign (a tag) which enables them to be identified (grey and black in our simulation). The authors assume that this sign does not have any intrinsic significance (completely meaningless).

However, the agents memorize the sign of the opponents whom they met and calculate the average behaviour corresponding to each type. There are thus two groups, determined beforehand by the tags, but this is not sufficient to cause a differentiated behaviour sensitive to tags, which could result in a shared belief on the behaviour of the members of these groups. However, in this model with two tag types, beliefs about the opponent's strategy may diverge depending on the opponent's tag, leading to between-types (grey against black, Figure 3, right) and within-type (grey against grey or black against black, Figure 3, left) responses. By definition, the formation of "classes" corresponds to the relative stabilization of distinct beliefs based on the group, leading to an equitable intra-group behaviour (within), and an unfair share between classes (the opposite case exists, but can be regarded as pathological). In the situation displayed on Figure 3 left, grey dots show equitable behaviour (they play M) when encountering agents of their kind (within = intra-group), whereas black dots do not (but they have moved closer to the zone of equity). The situation displayed on the right shows that black dots have the belief that grey dots adopt in the majority of cases a "dominated" behaviour (L) and their best response thus consists in claiming a large share (H). Conversely, grey agents have the belief that black agents preferentially show a "dominating" behaviour (H)

and their best response then consists in adopting a dominated attitude by accepting a small share (L). Therefore, both beliefs reinforce each other.
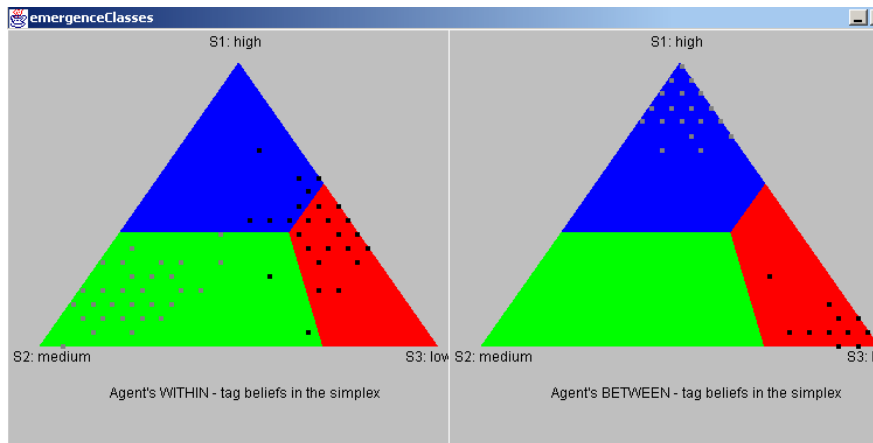

Figure 14.4: Simplexes Within (tag) - Between (tag) and emergence of classes
source: simulation on Moduleco-MadKit of [AXT 01]

In the model with tags, (as in the case without tags), the stochastic process is ergodic and the only stable regime is the "equitable" one: MM. More specifically, if the length m of the agents' memory and the ratio of the number of agents N to this length (N/m) are "sufficiently large" while the trembling hand effect remains "sufficiently weak", the ergodic (invariant) probability to be in the equitable area is high. However, if m is large and N small, the inertia of the system, (i.e. the time before reaching or leaving an area) can be very important ("broken ergodicity" see [PAL 89].). This is true in particular for the transition from the mode with class towards the equitable standard.

What does occur in the regime with class? The uneven distribution between groups is an unstable attractor in the sense of [FOS 90], but the system can converge towards this state and remain stable for a rather long duration (due to broken ergodicity). A typical stabilization corresponds to the situation where the players are equitable within their group and inequitable in relations between groups. Let us consider the situation represented on figure 14.3. When the black players are "aspired" in the basin of attraction of the equitable strategy, the agents play almost all "M" in the games within the group. Between groups, on the basis of their observation, the black agents have the common belief that the grey agents play mainly "L" Thus, their best reply is to play "H". The grey agents have the opposite beliefs. In all cases, these beliefs are not founded on belief in a common strategy of the members of the opposite group, identified by a specific tag, but on a statistical inference on the behaviour of a sample of agents characterized by their tag which

thus defines a group only as the set of its members and not as a "social object" [PHA 07]. The situation is (more easily) reversible precisely because the beliefs of the agents remain a sample of individual evaluations rather than on a representation of the behaviour of the group as such [DES 06].

Exercise 14.1. *Build on the space (m, ε) a statistical distribution of the transition duration towards a stable state around the attractor "MM" from the 2 perspectives (between, within), with initialization on the axis H-L.*

In these two cases, we referred to the emergence of some order which was not initially and explicitly contained within the functioning of the system: segregation in the first case, dominant/dominated categories by the agents themselves in the second case. Although, the term *emergence* is used in both cases, it is not clear at that stage whether it is really emergence, because in a sense the system has been made for such an order to occur, or whether it is the same kind of emergence. These questions shall be given some answers in the next sections.

14.3. Characterizing emergence

In Philosophy, emergence has a long history, from Mill's chapter: "Of the Composition of Causes" in its *System of Logic* (1843) to the contemporary debates about the philosophy of mind, known as "the mind - body problem". For a synthesis, see among others: [BEC 92], [MCL 92], [MCL 97]. Philosophical emergentism deals with the questions of novelty, unpredictability, reductionism and holism. Lewes [LEW 75] for instance places emergence at the interface between levels of organisation. For descriptive emergentism, the properties of the "whole" "cannot even in theory, be deduced from the most complete knowledge of the properties of [the parts] in isolation" [BRO 25, Chapter 2]. This definition is usually taken as a characterization of complex systems. Emergent properties result therefore also from the *relations* between parts, and in some cases from some irreducible macro causal power from the system itself (i.e. downward causation). We notice that the relational properties that structure the system are neither at the level of the whole nor at the level of the parts, but are constitutive of both. In agreement with the questions raised in Figure 1, one can view emergence as a process at the interface between two levels of organization (micro / macro). Thus, Bedau distinguishes two hallmarks of how macro-level emergent phenomena are related to their micro-level bases. Emergent phenomena are dependent on (constituted by, and generated from) underlying processes, and are (somehow) autonomous from these underlying processes [BED 97, p. 375].

14.3.1. *Varieties of Emergence: Some conceptual issues from the philosophic debate*

For the purpose of this paper, we rely on the distinction, introduced by [LAB 96], [FER 97] and developed by [MUL 04] in the field of multi-agent systems, between "weak" and "strong" emergence. The latter refers to a situation in which agents are able to witness the collective emergent phenomena in which they are involved, which opens the road for both upward and downward causation. We shall explore the various meanings that can be given of this distinction between weak and strong emergence.

In the philosophical debate around the definition of emergence, all authors consider *downward causation* and *irreducibility* as necessary conditions for strong emergence. According to downward causation, the behaviour of the parts (down) may be determined by some properties or behaviour of the whole (top). For instance, parts of the system may be restrained in conformity with some rules given at the system level. Causation would come "downward" according to a holistic principle rather than "upward" (from the bottom up). The debate is about reductionism. Some authors consider irreducibility as a necessary condition of emergence, other do not. As a consequence, there is no unified definition of weak emergence.

As an example, Bedau [BED 97],[BED 02] proposed to distinguish different kinds of emergence: "nominal", "weak" and "strong", and Gillet [GIL 02a],[GIL 02b] distinguishes "weak", "ontological", and "strong" emergence. For Bedau, the broader (weaker) form of emergence is called "nominal". *Nominal emergence* concerns the existence of some macro-property that cannot be a micro property. Each level has its specific distinct role and properties: "macro-level emergent phenomena are dependent on micro-level phenomena in the straightforward sense that wholes are dependent on their constituents, and emergent phenomena are autonomous from underlying phenomena in the straightforward sense that emergent properties do not apply to the underlying entities" [BED 02]. Under this latter condition, *strong emergence* is the opposite of nominal emergence, as in this case, emergent properties have irreducible causal power on the underlying entities: "macro causal powers have effects on both macro and micro levels, and macro to micro effects are termed downward causation" [BED 02]. For Bedau, *weak emergence* is a subset of nominal emergence for which the emergent phenomenon is not easy to explain, according to Simon: "given the properties of the parts and the law of their interactions, it is not a trivial matter to infer the properties of the whole" ([SIM 96], p. 184 quoted by [BED 02]). Accordingly, for Bedau, weakly emergent phenomena are those which *need to be simulated,* to be revealed: "Assume that (a macro-state) P is a nominally emergent property possessed by some locally reducible system S, then P is weakly emergent if P is derivable from all of S's micro facts but only by simulation". According to the non-trivial (surprising) dimension of

emergent phenomena, the need for simulation seems to be a *transitory* epistemic criterion only. If, *in a context of discovery*, computer simulation reveals some new emerging patterns, this is not a sufficient condition to have no other way forever. Later justification by some explanatory formalism is a possible outcome. Thus, could a surprising (weak) emergent phenomenon become a "simple" (nominal emergent) one?

Stephan [STE 02a], [STE 02b] proposes an interesting discussion on the difference between weak and strong forms of emergence in a larger framework, using the difference between "synchronic" and "diachronic" emergentism (see also [RUE 00]). In *synchronic emergence*, a macroscopic emergent phenomenon can be explained by the current (synchronic) interactions of the interrelated microscopic entities. In *diachronic* emergentism, the emergent phenomenon appears across time by observing in a diachronic perspective the sequential adaptation of microscopic entities. The centre of interest is then the evolution of macro structures, and not only the occurrence of a particular structure. As underlined by Stephan [STE 02b] synchronically emergent properties include also diachronically emergent ones, but not conversely.

For instance, in a Wolfram's one-dimensional network of automata [WOL 84], a specific configuration of the network emerges at each step from the value of the automaton and the structure of their relations at the previous steps (synchronic emergence). In some cases, identified by both [WOL 84] and [LAN 84], the existence of an attractor drives the system towards a particular stable configuration (fixed point regular cycle, see Chapter 10 section 2.2.). In some others cases, called by Langton "the edge of Chaos" [LANG 98], the evolution of the states of the systems, step by step, generates a particular structure, such as the Sierpinsky's triangular structures (Chapter 10 Figure 10.7). This structure is only observable from a diachronic perspective, and results from the *succession* of synchronic emergence of macrostructures due to the local interaction of microstructures (namely, automata) within the specific one-dimensional nearest-neighbour interaction.

	Synchronic determination	Diachronic determination
Weak (reducible)	weak emergentism	weak diachronic emergentism
Strong (irreducible)	(strong) synchronic emergentism	strong diachronic emergentism

Irreductibility ↓

— Novelety →

Figure 2: *Irreducibility and novelty in emergent phenomenon for Stephan (adapted from [STE 02b])*

For Stephan [STE 02a] the weaker version of emergentism (*weak emergence*) can be characterized by three features. First, following the physical monism thesis, only *physical entities* can bear the emergent properties or structures. Secondly, emergent properties or structures are attributes of the system itself, and cannot be attributed to some system's part. Thirdly, the principle of *synchronic determination* implies that all properties of the system *nomologically depend* on its micro-structures, namely, the parts and their relations. Stephan [STE 02a],[STE 02b] claims that numerous formal approaches to complex systems, connectionism and cognitive science can be related to weak emergentism. Figure 2 shows a cross perspective on Synchronic/Diachronic - Weak/Strong emergence, adapted from ([STE 02b]- without the case of unpredictability).

To summarize, in the philosophic literature, while both downward causation and irreducibility are generally considered as necessary conditions for strong emergence, the definition of weak emergence remains unclear, and depends on the author. For some of them, a necessary condition of emergence is irreducibility, and for [STE 02a], [STE 02b] reducibility corresponds to the case of weak emergence. The relevant question seems then to be: what is the framework of reference for the reducibility criterion?

According to Bunge [BUN 77], [BUN 04], a formal definition of emergence allows us to answer to this question. Unfortunately, various attempts have been made to define emergence in an "objective" way, driving to various criteria. Some definitions refer to self-organization [VAR 91], to entropy changes [KAU 90], to non-linearity [LANG 89], to deviations from predicted behaviour [ROS 77], [ROS 78], [ROS 85], [CAR 91] or from symmetry [PAL 89]. Other definitions are closely related to the concept of complexity [BON 95b], [CAR 91], [KAM 91a], [KAM 91b], see [BON 95a] and [DEG 06] for a survey).

Although these definitions make use of concepts borrowed from physics and information science, they all involve inherently contingent aspects, as the presence of an external observer seems unavoidable. To go beyond this problem the definition of emergence needs to be related both to a particular formalism, and to a particular system of observation. This allows us to encompass the problematic definition of weak and strong emergence, by defining emergence as an irreducible phenomenon from the standpoint of a particular formalism, including a particular observation system, and to define strong emergence by the downward causation inside this framework.

14.3.2 *Emergence as a phenomenon related to an observer (1) formal definition*

The unavoidable presence of an observer does not preclude the possibility of extending the definition of emergence to include non-human observers or observers that are involved in the emerging phenomenon. In our quest for "strong emergence", we wish to assign the role of the observer to elements of the system itself, as when human individuals become aware of phenomena affecting the whole society. This kind of self-observation is only possible because what is observed is a simplified state of the system. *Emergence deals* precisely *with simplification.*

Ronald and Sipper [RON 01] introduce a new approach called "emergent engineering", in order to get a controlled notion of the above-mentioned concept of "surprise". This approach opposes the classical engineered automation, based on unsurprising design, and the biologically inspired automation system, which allows the possibility of "unsurprising surprise". Many engineered emergent systems are based on this concept (e.g. [VAA 94], [MUL 04]). We do not deal directly with emergent engineering, but we discuss the framework used by this author, based on a specific formal test of emergence, previously presented in [RON 99]. This test of emergence involves *two functions*, which can be assumed by the same individual or by two different persons: (1) a *system designer* and (2) a *system observer*. An emergent phenomenon can be diagnosed by combining the three following conditions (from: [RON 01], p.20)

1 – *Design*. The system has been constructed by describing local elementary interactions between components (e.g. artificial creatures and elements of the environment) in a language L_1.

2 – *Observation*. The Observer is fully aware of the design, but describes global behaviour and properties of the running system, over a period of time, using a language L_2.

3 – *Surprise*. The language of design L_1 and the language of observation L_2 are distinct, and the causal link between the elementary interactions programmed in L_1 and the behaviour observed in L_2 is *non-obvious* to the observer- who therefore experiences surprise. In other words, there is a cognitive dissonance between the observer's mental image of the system's design stated in L_1 and its contemporaneous observation of the system's behaviour stated in L_1.

The question is then how easy it is for the observer to bridge the gap between L1 and L2. The authors use artificial neural network classifiers to evaluate this gap. Within this framework, an "unsurprising surprise" can be defined as an "expected" surprise. This question is exemplified later, within the [BON 97] framework of emergence as reduction of complexity within the observation system.

The framework of [RON 99],[RON 01] together with Forrest's definition of emergent computation [FOR 90] allow us to define emergence in SMA as occurring

between two organization levels [MUL 04], distinguishing the process itself and the observation of that process. The process concerns the evolution of a system formed by entities in interaction using a language L_1. These interactions may generate *observable epiphenomena* either stationary or dynamical, taking into account respectively the synchronic and diachronic dimensions of emergence. At the observation level, *epiphenomena are interpreted using another language* L_2. Finally, emergence is defined as a particular relationship between the two languages where L_2 is not compositionally reducible to L_1 in the sense of [BUN 77]. A good example is the game of life where the cells of a cellular automaton have two possible states (dead or alive) and the evolution depends on the neighbours' states (L_1). Some regular phenomena like the "glider" can appear as objects moving on the grid with their own laws (L_2). Notice that there are no objects moving anywhere in the cellular automaton but just fixed cells changing states. L_2 is neither predictable nor reducible to L_1 in any way. This account is stronger than the notion of *weak emergent phenomenon* in the sense of Bedau [BED 97], [BED 02] by adding the intrinsic irreducibility of the two description languages to the necessity of simulating,. For example, according to [BUN 77], the temperature is not emergent from molecular movements because it is reducible to the average kinetic energy of the system. Accordingly, most phenomena described by statistical mechanics are not emergent in the sense of Bunge. For us *weak emergence* arises when the observer is external to the system, while *strong emergence* arises when the agents involved in the emerging phenomenon are able to perceive it. In this latter configuration, the identification of epiphenomena by agents interacting within the system involves a *feedback from the observation to the process*. There is a coupling between the process level and the observation level through the agents because the agents are using both L_1 and L_2. This form of strong emergence is thus *immanent* in such a system. This definition is directly related to the role of cognitive agents for downward social determination as pointed out by [CAS 00]. In order to characterize this explicit description of how downward causation may occur, we call "M-Strong" the strong emergence in this sense [MUL 04].

To summarize, if there is *M-Strong emergence*, the system becomes reflexive, through the mediation of the agents.

(A) Agents are equipped with the capacity to observe and to identify a phenomenon in the process which represents the evolution of the system in which they interact. This capacity of observation and the target of such observation must then be sufficiently broad to encompass the phenomenon as a global one.

(B) The agents describe this epiphenomenon in a language other than the one used to describe the process.

(C) The identification of an emergent epiphenomenon by the agents involves a change of behaviour, and therefore a feedback from the level of observation to the process.

Additionally, [MUL 04] insists on the role of the environment as a mediation between the process and the observation of the process. Trivially, an observation is based on perception and perception is always perception of physical phenomena or states (even sounds). Downward causation is only conceivable when the emergent phenomena generate physical changes which can influence the components of the system. This account further distinguishes two kinds of downward causation, a local one and a global one. In the first case, a global phenomenon is produced but only a part of it is locally perceived, e.g. the global structure of a termite's nest which only locally influences movements and work sharing among the termites. It is a case of weakly emergent downward causation. In the second case, the global phenomenon as such is globally perceived as it would be the case if each termite had a map of the whole nest. Only the latter case is considered as strong emergence. Further more, [LAB 96], describes how an interaction structure can actually elaborate information about a global state of affairs towards a single agent. Accordingly, it is a strong emergence case because there is a real feedback of the global state on the local behaviour. However, only an external observer can actually interpret the local interaction as carrying information about some state of affairs. In other words, the *agents are not aware* of the downward causation. There is strong emergence but no reflexivity. As the feedback is external to the agent but internal to the collective this allows the emergence of fully "social" relationships (see Appendix 2, § A2.3.1*)*

14.3.3. *Emergence as a phenomenon related to an observer (2) quantitative definition*

This section proposes to seek a formal definition of emergence. The aim is not to encompass all the richness of the concept in a single formula, but at least to capture the gist of what is common to all preceding notions that pertain to emergence.

Emergence concerns phenomena that become suddenly perceptible, that are novel, that elicit surprise, that undergo a change of nature, that were unpredictable in the first place and must be discovered. Implicit in all these aspects is the existence of *two successive observation times*. The temporal change may be *physical*, as when the system's evolution undergoes a transition phase; the change may result from a *perspective* shift, for instance when one goes from the single agent's behaviour to the behaviour of the collective; the change may be *epistemological*, when the observer adopts two successive descriptions of the same collective phenomenon. A system that shifts into a chaotic behaviour for a definite value of the control parameter is an example of the first case. The formation of a price in economy, or the regular 2-D pattern formed by a 1-D cellular automaton are examples of the second case; emergence that takes the form of a discovery or that requires a change of description language can be seen as examples of the third, epistemological, case.

Emergence is also supposed to be a sharp rather than a fuzzy phenomenon. We may give an account of this by saying that the two observational times may be arbitrarily close to each other. If we note them t_1 and t_2, then their difference Δt may be arbitrarily small.

The next question is to know what happens between these two successive observational times that characterize emergence. If the question is to know *what* emerges, the most general answer is *a structure*. The only thing that cannot be said to emerge is noise. The system at date t_2 is characterized by some structural aspect that was absent from the same system at date t_1. The Sierpinsky's triangular pattern of Figure 10.7 (Chapter 10) is absent from the one-dimensional state and from the rule of the cellular automaton, whereas it is obvious in the diachronic perspective in which all successive states are displayed one after the other. Not all suddenly appearing structures are emergent, however. Every time a new picture appears on a computer or a TV screen, new structures appear that would hardly constitute examples of emergence.

Another necessary ingredient of emergence is the pre-existence of a *collective*: molecules in a gas, economic actors, agents in a simulation, pixels in an image, or starlings in a flock. This requirement is inherent in the observation that emergence must occur at the interface between two organization levels. The emerging structure is therefore composed of a number of identical elements. If we refer to Michael Leyton's theory of shape [LEY 01], such a structure results from the *transfer*, through a group of transformations, of a lower-level structure. According to Leyton, our visual system analyses a cloud of regularly spaced out starlings, by transferring one single starling through a group of integer translations through space. Transfer would be the only way for us to perceive a set of objects as an integrated structure.

A suddenly appearing collective structure still does not make up a case of emergence. The trivial cellular automaton that merely copies its current state produces a two-dimensional picture of parallel stripes that is highly predictable and thus can hardly be seen as emerging. *Unexpectedness* is a key ingredient of emergence. We capture this idea with the following definition of emergence [BON 97]:

$$E = C_{exp} - C_{obs} \qquad\qquad [14.1]$$

Emergence E is quantitatively defined as the difference between the expected complexity C_{exp} of the system and its actual, observed complexity C_{obs}. Complexity here means cognitive complexity [DES 06], i.e. the size of the smallest cognitive description of the system that is available to the observer. This concept is an instantiation of the general definition of the Chaitin-Kolmogorov complexity [LIV 93]. The cognitive complexity of a given structure may be assessed, for instance, by

recursively summing up the complexities of the different transfer groups that, in Leyton's account, make up that structure. Highly structured scenes have therefore lower complexity. The fact that some structure σ is emerging translates into the fact that C_{obs} is small. The fact that σ is unexpected means that E is positive. Emergence demands several additional properties: that σ be a collective structure, that C_{exp} be extrapolated from the complexity of the system at time t_1 while C_{obs} is its complexity at time t_2, and that $\Delta t = t_2 - t_1$ can be made arbitrarily small.

As we see from (1), the intensity of emergence rises with the amplitude of the complexity drop. In the case of the trivial cellular automaton, there is no emergence: the observed pattern is simple and highly structured, so C_{obs} is small; but C_{exp} is small too, thanks to the simplicity of the null rule. Both C_{obs} and C_{exp} amount to the complexity of the initial one-dimensional pattern plus the complexity of the translation group, and E is zero. The recursive triangular pattern of Figure 1, by contrast, as far as the observer cannot anticipate it from the initial state and from the rule, is highly unexpected. An observer who is not used to cellular automata may anticipate a visual complexity which is $C_{exp} = C_0 + nC_R$, where C_0 is the complexity of the initial pattern, C_R is the complexity of the rule and n is the number of time steps. The recursive triangular pattern has a much lower complexity, though. It results from the copy, on each scale, of the same triangular shape, and its complexity is the complexity of the simple recursive procedure that generates it in a two-dimensional plan. The contrast between C_{obs} and C_{exp} is thus high and emergence can be said to occur for that observer.

Emergent phenomena can be usefully described in a two-level architecture, where elements of the collective form the lower level of organization and the collective itself is detected at the upper level. These levels have no objective existence, as they pertain to the same physical system. They result from the detection abilities of the observer. The detected object at the upper level is composed of objects of the first level. Correspondingly, the upper level detector is triggered by the activity of lower-level detectors. A human observer may see a set of starlings as a moving flock because he or she can detect isolated starlings and because he or she is able to detect their coordinated movement.

Let us call s the emerging phenomenon, $\{d_i\}$ the set of lower-level detectors and D a higher-level detector. Before emergence occurs, the expected complexity may be written:

$$C(s \ \& \ \{d_i\}) = \Sigma_i \ [C(d_i) + C(s|\{d_i\})]$$

The notation $C(a|b)$ means the complexity of a when the description of b is available. In our example, the $\{d_i\}$ may refer to the detection of single starlings and of their movement, in which case $C(s|\{d_i\})$ is zero, as the $\{d_i\}$ are sufficient to

account for the ornithological scene. Let us suppose that a new detector is taken into account. The expected complexity becomes:

$$C_{exp} = C(s \ \& \ D \ \& \ \{d_i\})$$

Suppose the scene is described using D first. Then, the actual complexity becomes:

$$C_{obs} = C(D) + \Sigma_i \ C(d_i|D) + C(s|D \ \& \ \{d_i\}) \qquad [14.2]$$

Most of the time, $C_{obs} = C_{exp}$, which means that the complexity of the new detector compensates for what is gained by using it. In our example, looking at a set of a few starlings flying independently and considering it as a whole leads to no complexity decrease. If, however, D subsumes some of the d_i, then $C(d_i|D)$ becomes small or even zero, and C_{obs} gets significantly smaller than C_{exp}. This is where emergence occurs. In our example, D may be the ability to perceive the coordinated movement of the flock. When D is active, most of the starlings' individual movements become predictable. This sudden upper-level pattern recognition lowers the overall complexity according to the preceding formula, giving rise to computable amplitude of emergence.

Note that formulas [14.1] and [14.2] make a prediction that is not acknowledged in most models of emergence. The emerging characteristic must be *simple*. The simpler it is, the more significant the emergence. In formula [14.2], it is important that $C(D)$ be small, as a large value would ruin the emergence effect. As soon as the relative movement of the starlings becomes too complex, emergence vanishes. Conversely, those who can witness the collective reaction of a starlings flock to the attack of a hawk are struck by the emergent evasive manoeuvres of the group of birds that moves as a malleable entity that seems to cleverly dodge the predator's assaults. Such a scene is much more economically described at the collective level, a fact that the simple short-sighted self-preserving reactions of individual birds did not allow to anticipate.

The requirement that the emerging property be simple seems to be verified in all examples of emergence to be found in literature. This statement may be surprising at first sight. On certain occasions, emergence seems indeed to involve an increase rather than a decrease of complexity. Examples such as bifurcation into chaos come to mind: the apparent complexity of chaotic behaviour (for whom ignores the simple process that generates it) is huge, compared to usual deterministic trajectories. For such a situation to be considered a case of emergence, the chaotic outcome must be unexpected, *i.e.* unique in the set of anticipated outcomes. This uniqueness is what makes it simple: a minimal description allows isolating it among alternative

outcomes. Chaotic behaviour emerges only in such a context, as one remarkable instance among a large set of alternatives.
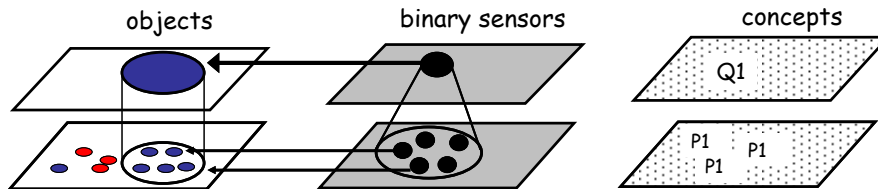

Figure 3 - *Parallelism between hierarchies: level of description, level of observations (detectors) and conceptual level (association concepts- detectors) [DES 05]*

We may wonder how the preceding definition of emergence as a complexity shift relates to other definitions reviewed in this chapter. As shown in [DEG 06], the change of description language invoked by Müller or by Ronald, Sipper and Capcarrère amounts to taking new detectors into account. This language change is captured by *D* in the preceding formula. The 'non-obvious' character of the behaviour described in the upper-level language, as invoked by [RON 99] corresponds in our framework to the unexpected complexity shift.

Definition [14.1] may also be applied to cases of diachronic emergence. The fact that a given structure can only be detected by comparison between successive states of the system may be merely ignored when considering complexity shifts. Structure and thus unexpected simplicity is discovered in the set of successive time slices. Diachronic emergence, according to definition [14.1], occurs whenever the complexity of this set turns out to be simpler than anticipated.

14.4. The road for strong emergence

The interest of the definition of emergence as reduction of the relative complexity in the system of observation does not prejudge the nature of the observer: this one can even be an artificial mechanism, equipped with hierarchical detectors. Thus, as Dessalles notices [DES 92], a system of observation of road traffic, can "see" the emergent phenomenon that human beings will name traffic jam, or accidents, simply because it will be equipped with suitable software detectors, able to locate these phenomena starting from the positions and speeds of the vehicles (this presupposes that the regulator system has an adequate model of such phenomena). The fact that an artificial observer can detect an emergent phenomenon allows that (1) the system itself or (2) some elements which constitute this system are able to retroact on the process. The first case arises each time we detect our own mental states. As a neuronal community, our nervous system is able to detect the activation of several classes of particular populations of its own

neurons. The second case arises when the system consists of cognitive agents. Those are then able to locate certain system requirements in which they are immersed (strong emergence in the sense of Müller, immergence in the sense of Gilbert). To make this conceptualization operational, it is necessary that the agents have sensors likely to identify correctly the phenomenon. The definition of the detectors and the construction of the hierarchy of levels for the analysis presuppose that the modeller designs an integrating model likely "to absorb" the emergent phenomenon.

14.4.1. *Introducing strong emergence in the Axtell, Epstein and Young model*

In the AEY model, the beliefs of the agents limit themselves to build an estimator of the expected behaviour of the others by means of the statistical sample σ_{im}, grounded on the observation of the strategies played by other players in the past. The agents have neither belief on the beliefs of the other agents nor belief on the behaviour of the group as such. According to Orléan [ORL 04] it is possible to qualify as "social beliefs" the beliefs on a regular behaviour of the members of the group identified by a tag [PHA 07]. The emergence of such "social beliefs" can takes the form of "awareness" by the agents of the existence of a common behaviour, characteristic of the members of each group, only founded on their membership identified by a tag. This regularity justify the subsumption under a general category – say the tag-group, of the inference of the agents' behaviour by means of the statistical sample σ_{im}. Such a subsumption results in a reduction of the perceived diversity of the agents to their only membership of a collective entity identified by a tag supposed to describe them in a sufficiently relevant way to rationally base their strategy on a rule of best reply to the opponent's strategy, supposed to be associated with the tag.

In this section, we suppose, as AEY, that the tag-groups pre-exist and we focus the discussion on this problem of subsumption, which results in a qualitative change of inferential indicator. The question is the emergence of a belief on a sufficiently regular behaviour of the members of the group as such, so that it is superfluous (for the rationale for the action) to wonder more about the effective diversity of the individual behaviour. For this goal, according to [DES 05], [DES 06], [PHA 07] it is possible to formalize this emergence phenomenon as a multilevel cognitive problem, conceptualizable as a specific form of "cognitive hierarchy" problem [PHA 04b], [DES 07].

14.4.2. *Interpretation: emergence of a "social belief" as a phenomenon of subsumption in the system of observation hierarchy of reflexive MAS*

In [DES 06], the agents minimize a procedural cost of informational treatment and it is thus an instrumental design (cost / benefit) which controls the "internal" agents' behaviour of subsumption. Other approaches of this mechanism are possible, but the core problem is elsewhere. In this approach, the agents are endowed at the beginning with a minimal capacity of categorization by the design of the model which provided explicitly the possibility for the agents to identify the emergent phenomenon previously studied by AEY. There is not thus a "surprise" in the sense of [RON 99]. The design and implementation of a "pre-wired" multilevel solution rather than a generative solution refer to the correspondence between the hierarchy of levels of complexity and the hierarchy of encompassed models in [BON 97]

In the simplest case, the model designer expects the set of relevant structures for the problem at the initial stage of design, which concerns the level of the formalization of the perceptive and cognitive functions, or at the level of the social organization as well. We are thus in an epistemic configuration of the type highlighted by Piaget [PIA 70] for so called *genetic structuralism.* The knowledge of the two structures (of departure and of arrival) is necessary to explain the transition between these structures. In genetic psychology, this form of structural emergence does not raise any problem of principle, since one knows the various cognitive structures which are connected. In economy, this constraint is not penalizing for *retrodiction* of always known emergent phenomena, but can be a problem for prediction. Another approach would have however been possible, but we did not seek yet to explore it. It consists in envisaging generative and flexible mechanisms of observation designed right from the start to categorize and simplify the practical resolution of the problems to treat, by integrating the possibility of having a "surprise effect".

14.4.3. *Towards cognitive agents*

M-Strong emergence provides a useful operationalizable semantics to model artificial societies [GIL 95a]; [CAS 98a], [CAS 98a]. This is the case also even if there is a mix of strong and weak emergence in most multi-agent based social simulation [DRO 01], [DRO 94] The reflexivity meditated by the agents' "consciousness" and/or "awareness" appears to be a determinant characteristic that distinguishes systems involving human agents from systems made of non-conscious or material entities. In this latter domain, it is interesting to distinguish a hierarchy in the cognitive capacity of agents. According to [BOU 93],[PHA 04b] distinguishes the following level of *cognitive rationality* to be used in Agent-Based modelling:

(1) A Reactive Agent (RA) has a pre-determined response for a given state of its environment. Such an agent may represent both a material entity (such as a "spin" in a ferromagnetic set) or all living entities (from the point of view of their pre-determined patterns of behaviour - the Darwinian creatures in the hierarchy introduced by Dennett, 1996[2]). The level of cognitive rationality of such an agent is null, and it can be characterized by its "situatedness", i.e. its immediate response to environmental events.

(2) A Behavioural Agent (BA) may modify its behaviour for a given state of its environment according to the observed (historic) payoff obtained from its past actions, following a Skinerian reinforcement learning (e.g. Q-Learning, classifiers, etc.). The level of cognitive rationality of such an agent is low, and it can be characterized by its adaptability: or "consciousness" in [BOU 93], [PHA 04b].

(3) An Epistemic Agent (EA) uses a model of its environment to pre-select actions. Its level of cognitive rationality may be medium or high, depending on the scope of its belief, and on the sophistication of its cognitive tools. So, in the hierarchy introduced by (Dennett, 1996) both Popperian and Gregorian are different cognitive levels of epistemic agents, and can be characterized by "awareness".

Both Behavioural and Epistemic agents have specific status with respect to the cognitive capability to process available information. Both are observers of the process which they take part of. A behavioural observer only takes into account visible characteristics of its environment that are known to have an effect on some personal criteria (viability, hedonic index…). An epistemic observer model in some way has a representation of this process and can simulate it, by means of symbolic tools. Let us remark that, from this point of view, an external observer (the modeler, the experimentalist) have some common features with an Epistemic agent. If we take into account the case of downward causation as described by [LAB 96], any kind of agent can produce strong emergence as well as weak emergence. The real question in this hierarchy is the question of awareness, i.e. of reflexivity on the relationship between the local perceptions and some state of affairs. In this direction, a Reactive agent has no awareness at all. A Behavioural agent has at least a feedback on whether its behaviour is correct (or how much it is) regarding some survival

---

[2] Dennett [DEN 96] represents the hierarchy of cognitive capacity in the phylogeny of living creatures by his "Tower of Generate-and-Test". At each phylogenic stage, a qualitatively new cognitive capacity comes to enhance the existing ones, inherited from lower-level stages. At the lowest stage, he places the Darwinian creatures, which have a rigid phenotype. At the second stage, Skinerian creatures have a phenotype adaptable through reinforcement-learning capabilities. At the third stage, Popperian creatures have some capability to pre-select actions, given the available information, coming from inheritance and/or acquisition. At the last stage, Gregorian creatures can enhance their individual performances through the use of "tools". Among the tools, language and models are special kinds of mental (symbolic) tools, very important for cultural transmission.

conditions. Only Epistemic Agents not only behave but have a representation of how they behave, hence the possibility to simulate it. Anyway, an important feature is the availability of the inferior level of cognition for higher level agents. In other words, an Epistemic agent can behave sometimes like a Behavioural agent or like a Reactive agent. Another important feature for the emergence in society of the cognitive agent, as we discuss it in section 4, is the notion of "*social intelligence*" used in particular by [CON 99] as a property of *socially situated* agents. Such agents are subject to a double-level hierarchy, and therefore a double reflexivity, both with the social and the cognitive dimension, based respectively on the awareness of the collective actions and of their own actions. Taking into account the cross-feedback between social and cognitive dimension of agents at the first level, the second level in the cognitive and social dimension can be viewed as an emergent phenomenon grounded in the first level interactions. However it is an open question which one comes first as illustrated by the still actual debate between [PIA 70] where individual thought comes first and Vygotsky [VYG 78] where social reflectivity comes first.

14.4. Conclusion

Emerging phenomena in a population of agents are expected to be richer and more complex when agents have enough cognitive abilities to perceive the emergent patterns and, even before, if the agents are endowed with perception capabilities for filtering information at various hierarchical levels as described in 3.1. Such feedback loops between emerging collective patterns and their cognitive components clearly occur among agents in human societies. They may obey laws that are still to be understood. Our aim here is to design a minimal setting in which this kind of strong emergence unambiguously takes place. In this context, the definitions of emergence as sketched in this chapter, both formal and quantitative, are essential steps towards such an aim

14.5. References

[AXT 01] Axtell R. Epstein J.M., Young H.P. "The Emergence of Classes in a Multi-agent Bargaining Model" in Durlauf, Young eds., *Social dynamics*, Cambridge Ma.. The MIT Press, 2001, p. 191-212.

[BEC 92] Beckermann A., Flohr A.H., Kim J., eds. *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism*. Berlin & New York: Walter de Gruyter, 1992.

[BED 97] Bedau, M.A. "Weak Emergence", Supplement: *Philosophical Perspectives*, 11, Mind, Causation, and World, Noûs, 31, 1997, p. 375-399.

[BED 02] Bedau, M. A. "Downward causation and the autonomy of weak emergence", special issue on Emergence and Downward Causation*, Principia* 6-1 June 2002, p. 5-50

[BLU 97] Blume L.E. "Population Games", in Arthur, Durlauf, Lane eds., *The Economy as an Evolving Complex System II* ; Santa Fee Institute, Studies on the Sciences of Complexity, Reading Ma , Addison-Wesley Pub.Co, 1997 p. 425-460.

[BON 97] Bonabeau E., Dessalles J-L. "Detection and Emergence", *Intellectica*, 25, 1997, p.89-94.

[BON 95a] Bonabeau, E., Dessalles, J.-L., Grumbach, A. "Characterizing emergent phenomena (1): A critical review", *Rev. Int. Syst.* 9, 1995, p.327-346.

[BON 95b] Bonabeau, E., Dessalles, J.-L., Grumbach, A. "Characterizing emergent phenomena (2): a conceptual framework." *Rev. Int. Syst*. 9, 1995, p.347-369.

[BOU 93] Bourgine P. "Models of autonomous agents and of their coevolutionary interactions" *Entretiens Jacques Cartier*, Lyon, 1993.

[BOU 04] Bourgine P., Nadal J.P. eds. (2004) *Cognitive Economics, An Interdisciplinary Approach*, Berlin-Heidelberg, Springer-Verlag.

[BRO 25] Broad, C.D. *The Mind and Its Place in Nature*, London: Routledge & Kegan Paul, 1925

[BUN 77] Bunge M. "Emergence and the mind", in *Neuroscience* 2, 1977, p. 501-509

[BUN 04] Bunge M. "Clarifying Some Misundetstandings about Social Systems and their Mechanisms", *Philosophy of the Social Sciences*, Vol 34-2 June, 2004, p. 371-281.

[CAR 91] Cariani, P. "Adaptivity and emergence in organisms and devices", *World Futures* 31, Gordon & Breach Science Publishers, 1991, p. 49-70,

[CAS 98a] Castelfranchi C. "Simulating with Cognitive Agents: The Importance of Cognitive Emergence", in Lecture Notes in Computer Science vol 1534, *Multi-Agent Systems and Agent-Based Simulation*, Berlin-Heidelberg, Springer, 1998, p.26-44.

[CAS 98b] Castelfranchi C. "Emergence and Cognition: Towards a Synthetic Paradigm in AI and Cognitive Science", in Coelho H., *Progress in artificial intelligence*, Springer, 1998, p. 13-26 reed.. in Lecture Notes in Computer Science vol 3187, *Multi-Agent Systems and Agent-Based Simulation*, Berlin-Heidelberg, Springer 2004, p.26-44.

[CAS 00] Castelfranchi C."Through the agents' minds: Cognitive mediators of social action", *Mind & Society*, Volume 1, Number 1 / March, 2000, p 109-140.

[CON 99] Conte R. (1999) "Social Intelligence Among Autonomous Agents" Computational & Mathematical Organization Theory 5:3, p. 203–228

[DEG 06] Deguet J., Demazeau Y., Magnin L."Elements about the Emergence Issue A survey of Emergence definitions", to appear in *ComPlexUs, International Journal on Modelling in Systems Biology, Social, Cognitive and Information Sciences*, Karger, Basel, Spring 2006.

[DES 92] Dessalles J.L. "La perception de l'émergence", *Proceedings of the Workshop on Emergence in Cognitive Science*, ENST Paris, 1992.

[DES 05] Dessalles J.L., Phan D. "Emergence in multi-agent systems: cognitive hierarchy, detection, and complexity reduction part I: methodological issues", in Mathieu, Beaufils,

Brandouy (eds.), *Agent-Based Methods in Finance, Game Theory and their Applications*, Series: Lecture Notes in Economics and Mathematical Systems, Vol. 564, Berlin, Springer, 200, p. 147-159.

[DES 06] Dessalles J.L., Galam S., Phan D. "Emergence in multi-agent systems part II: Axtell, Epstein and Young's revisited" 12th *International Conference on Computing in Economics and Finance*, June 22-25, 2006, Limassol, Cyprus.

[DES 07] Dessalles J.L., Ferber J., Phan D. (2007) "Emergence in Agent based Computational Social Science: conceptual, formal and diagrammatic analysis" in Shan Y. & Yang A. Eds., *Intelligent Complex Adaptive Systems*, Idea Group Inc, forthcoming

[DEN 96] Dennett D.C. *Kinds of Minds*, New York, Brockman, 1996.

[DRO 91] Drogoul A, Ferber J. (1991) "A Behavioural Simulation Model for the Study of Emergent Social Structures", *European Conference on Artificial Life*, Paris.

[DRO 94] Drogoul A, Ferber J., Cambier C. (1994) "Multi-agent simulation as a tool for analysing emergent processes in societies", *Proceedings of Simulating Societies Symposium*, University of Surrey, Guildford, 1992. Re-published in Gilbert N. Doran J. eds. *Simulating Societies: the Computer Simulation of Social Phenomena*, Londres, UCL Press 1994, p127-142.

[FER 97] Ferber, J., Labbani O., Müller J.-P., Bourjault. A. (1997) "Formalizing emergent collective behaviours: Preliminary report", *International Workshop on Decentralized Intelligent and Multi-Agent Systems*, St. Petersburg, p. 113-122.

[FOR 90] Forrest S. (1990) "Emergent computation: Self-organizing, Collective, and Cooperative Phenomena in Natural and Artificial Computing Networks", introduction to the proceedings of the ninth annual CNLS Conference', in *Emergent computation*, MIT Press, Cambridge, M.A., 1990, p.1-11.

[FOS 90] Foster, D.P., Young H.P. "Stochastic Evolutionary Game Dynamics, *Theoretical Population Biology* 38, 1990, p. 219-232.

[GIL 95a] Gilbert N. (1995) "Emergence in social simulations", in Gilbert, Conte (1995). op.cit. p. 114-156.

[GIL 95b] Gilbert N., Conte R. eds. *Artificial Societies : The Computer Simulation of Social Life*, London, UCL Press; 1995.

[GIL 02] Gilbert N., "Varieties of emergence", Paper presented at the *Agent 2002 Conference: Social agents: ecology, exchange, and evolution*, Chicago, 2002

[HOL 98] Holland, J.H. *Emergence From Chaos to order*, Readings, Ma., Helix Books Addison-Wesley pub, 1998.

[KAM 91a] Kampis, G. (1991a) *Self-Modifying Systems in Biology and Cognitive Science* Pergamon Press.

[KAM 91b] Kampis, G. (1991b) "Emergent computations, Life and Cognition", *World Futures* 31, Gordon & Breach Science Publishers, p. 33-48.

[KAU 90] Kauffman S. (1990) "The Sciences of Complexity and 'Origin of Order' ", *PSA: Proceeding of thr Biennal Meeting of the Philosophy of Sciences Association*, Vol II Symposia and Invited Papers p. 299-322.

[KIM 92] Kim, J. (1992) "Downward causation", in Beckermann, Flohr and Kim, eds. (1992), op cit. p. 119-138

[KIM 99] Kim, J. (1999) "Making sense of Emergence", *Philosophical Studies*, 95 p.3-36.

[LAB 96] Labbani O., Müller J.P.Bourjault A. (1996) "Describing collective behaviours" in *Workshop ICMAS*, 09-13 Dec., Kyoto, Japan

[LAN 84] Langton C.G. (1984) " Self-Reproduction on a Cellular Automaton", *Physica D*, 10, p. 135-144

[LAN 89] Langton C. (1989) "Computations at the edge of chaos: Phase Transitions and Emergent Computation", *Physica D*, 42, p. 12-37

[LIV 93]. Li M. Vitanyi P.M.B. *An Introduction to Kolmogorov Complexity and Its Applications*, New York, Berlin, Springer-Verlag, 1993.

[LEW 75] Lewes, G. H. Problems of Life and Mind, Rinehart & Winston, 1875.

[LEY 01]. Leyton, M. *A generative theory of shape*, New York: Springer Verlag, 2001.

[MCL 92] McLaughlin B. P. "The rise and fall of British emergentism", in Beckermann, Flohr and Kim, op cit. p. 49-93.

[MCL 97] McLaughlin B. P. "Emergence and Supervenience", Intellectica 25, 1997, p. 25-43

[MUL 04] Müller JP. (2004) "Emergence of Collective Behaviour and Problem Solving", in Engineering Societies in the Agents World IV, Fourth International Workshop ESAW-2003, Revised Selected and Invited Papers, *LNAI 3071* Springer Verlag, p. 1-20.

[ORL 04] Orléan A. "What is a Collective Belief" in Bourgine, Nadal eds. [BOU 04], op.cit., p.199-212.

[PAL 89]. Palmer, R. (1989) "Broken Ergodicity", In *Lectures in the Sciences of Complexity*, D. Stein, ed. Volume 1, Santa Fe Institute Studies in the Sciences of Complexity. Addison-Wesley, Longman: Reading, Ma.

[PHA 04a] Phan D. "From Agent-Based Computational Economics towards Cognitive Economics », in Bourgine, Nadal, (2004), op. cit. p. 371-398

[PHA 04b] Phan D. "Hierarchy of cognitive interactive agents and statistical mechanics: how Object Oriented Programming highlights the connection", in Coelho, Espinasse. Seidel .eds., *5th Workshop on Agent Based Simulation*, Lisbon, Portugal, May 2004 SCS Pub. House, Erlangen, San Diego, 2004, p. 69-76.

[PHA 07] Phan D., Feber J. "Thinking the Social dimension of the Artificial World: Ontological status of Collective Beliefs", *International Transactions on Systems Science and Applications* vol 3, no 1, 2007, forthcoming.

[PIA 70] PIAGET J. *Structuralism.* New York: Harper & Row.1970.

[RON 01] Ronald E., Sipper M. Capcarrère M.S. "Design, observation, surprise! A test of Emergence", *Artificial Life* 5, 2001, p. 225-239.

[RON 99] Ronald E.M.A., Sipper M. "Surprise versus unsurprised: Implication of emergence in robotics", *Robotics and Autonomous Systems*, 37, 1999, p. 19-24.

[ROS 77] Rosen, R. "Complexity as a system property", *International Journal of General Systems* 3, 1977, p. 227-232

[ROS 78] Rosen, R. *Fundamentals of measurement and representation of natural systems*, North Holland, 1978.

[ROS 85] Rosen, R. *Anticipatory Systems: Philosophical, Mathematical and Methodological Foundations*, New York: Pergamon Press, 1985.

[RUE 00] Rueger A. "Physical Emergence, Diachronic and Synchronic", *Synthese*, 124, 2000, p. 297-322.

[SAW 01a] Sawyer, R. K. "Emergence in sociology: Contemporary philosophy of mind and some implications for sociological theory", *American Journal of Sociology*, 107 (3), 2001 p. 551-585.

[SAW 01b] Sawyer, R. K. "Simulating Emergence and Downward Causation in Small Groups", in Moss and Davidsson eds., Multi Agent Based Simulation Berlin, Springer-Verlag, 2001, p. 49-67

[SAW 02a] "Emergence in psychology: Lessons from the history of non-reductionist science", *Human Development*, 45(2), 2002, p.-28.

[SAW 02b] "Nonreductive individualism, Part I: Supervenience and wild disjunction", *Philosophy of the Social Sciences*, 32(4), 2002, p. 537-559.

[SAW 03] Nonreductive individualism, Part II: Social causation. Philosophy of the Social Sciences, 33(2), 2003, p. 203-224

[SAW 04] Sawyer, R. K. "The Mechanisms of Emergence", Philosophy of the Social Sciences, Vol 34-2 June 2004, p. 260-282.

[SAW 05] Sawyer, R. K. *Social emergence: Societies as complex systems*. Cambridge University Press, 2005.

[SCH 69] Schelling T.S. "Models of Segregation", *American Economic Review, Papers and Proceedings*, 59 (2), 1969, p.488-493.

[SCH 78] Schelling T.S. *Micromotives and Macrobehaviour*, New York, W.W. Norton and Co, 1978.

[SIM 96], Simon H.A. *The Science of the artificial*, Cambridge, MIT Press, 1996.

[STE 02a], Stephan A. "Emergentism, irreducibility, and downward causation", *Grazer Philosophiche Studien*, 65, 2002, p.77-93

[STE 02b], Stephan A. "Emergence" in: Nadel (ed.) *Encyclopedia of Cognitive Science*. Vol. 1., London: Macmillan, 2002, p.1108-1115.

[SUG 02] Sugden R. "Credible World: the status of theoretical models in economics", in Maki ed., *Fact and Fictions in Economics*, *Models, Realism and Social Constructtion*, Cambridge Ma., Cambridge University press, 2002

[VAA 94], Vaario A, Hori K., Ohsuga S. Towards Evolutionary Design of Autonomous Systems. *The International Journal in Computer Simulation*, 5, 1994, p.187--206

[VAR 91] Varela F., Thompson E., Rosch E. *The embodied mind*, Cambridge, MIT Press, 1991.

[VYG 78] Vygotsky, L. *Mind in Society: The Development of Higher Psychological Processes*, Cambridge, MA: Harvard University Press, 1978.

[WOL 84] Wolfram S. "Universality and Complexity in Cellular Automata", *Physica D.*, 10, 1984, p. 1-35.

[YOU 93] Young H.P. "An Evolutionary Model of Bargaining", *Journal of Economic Theory* 59 (1), 1993, p. 145-168.

[YOU 98] Young H.P., *Individual Strategy and Social Structure*, Princeton, NJ, Princeton University Press, 1998.