Chapter 3
History and Definitions

Giovanna Di Marzo Serugendo, Marie-Pierre Gleizes,
and Anthony Karageorgos

Defining and implementing Self-organisation and Emergence.

Objectives The objective of this chapter is to discuss the origins and definitions of self-organisation and emergence. When you will have read this chapter, you will:

- Obtain a clear understanding of the terms self-organisation and emergence and the pros and cons of the various definitions.
- Have been introduced to most common self-organising system properties.
- Understand technical and theoretical issues which are important for effectively using self-organisation and emergence in software applications.

3.1 Introduction

The issue of developing large-scale systems consisting of several (simple) components which, without explicit central control from some external authority, would

---

G. Di Marzo Serugendo
Birkbeck College, University of London, London, UK
e-mail: dimarzo@dcs.bbk.ac.uk

M.-P. Gleizes
IRIT, Université Paul Sabatier, Toulouse, France
e-mail: Marie-Pierre.Gleizes@irit.fr

A. Karageorgos (✉)
Technological Educational Institute of Larissa, Larissa, Greece
e-mail: karageorgos@computer.org

G. Di Marzo Serugendo et al. (eds.), *Self-organising Software*,
Natural Computing Series,
DOI 10.1007/978-3-642-17348-6_3, © Springer-Verlag Berlin Heidelberg 2011

collectively produce some complex global result has long been the subject of many research efforts. That global result, albeit complex, would not be chaotic or disordered, but instead it would be liable to some type of control laws, and it would demonstrate some observable properties. In other words, that result would be a desired identifiable organisation, for example a pattern, a process producing that pattern, or a global system function ensuring that the system would fulfil some specific requirements. Such a process and the resulting outcome are commonly referred to as *self-organisation* and *emergent phenomenon* respectively. The occurrence of an emergent phenomenon is commonly referred to as *emergence*. There has been recently a great interest in the study, analysis and design of systems capable of producing a collective outcome via self-organisation and emergence [14, 15, 29, 32, 72, 77]. In this context, of particular interest is the ability of self-organising systems to have a complex collective response arising from interactions among relatively simple individual components with limited abilities [54, 78].

The terms involved, for example organisation, self-organisation and emergence, have been used in various disciplines, such as mathematics, physics, biology and philosophy. These terms are increasingly used in computer science, and in particular they are widely met in the MAS research community. We need to establish a common, consensual and operational meaning of them to enable researchers to know if their artificial systems are self-organising or not and if there is emergent phenomenon or not.

To this purpose, Sects. 3.2 and 3.3 define self-organisation and emergence from the view of self-organising software, and in addition they provide an overview of the most common other definitions available in the literature.

3.2 Self-organisation

3.2.1 History

The concept of self-organisation has been first discussed in ancient Greek philosophy (see Haken in [41] referring to Paslack in [66]). In early modern times (18th and 19th centuries), self-organisation was discussed by the German philosophers I. Kant and F.W.J. Schelling [66]. Kant related self-organisation particularly with the formation of the planetary system, while Schelling's treatment of the subject was, as Haken states in [41], rather vague.

In more recent times the notion of "*self-organising system*" is traced back to the cybernetician W. Ross Ashby, who first used it in 1947 (see [41, 68] referring to [3]). By early 1950s, the term self-organisation had been well established in modern sciences by Nobel Prize winner Ilya Prigogine and his colleagues through studies in thermodynamics [35]. The concept of self-organisation introduced by Prigogine was essentially based on the assumption that open systems aim to decrease their entropy (order results from disorder) when some external energy is applied to them. The additional energy imposed often causes matter within the system to alter its organisational structure to facilitate the system reaching a new state with lower entropy.

The export of entropy does not directly explain how or why self-organisation does take place. Prigogine noted that self-organisation typically takes place in nonlinear systems which are far from their thermodynamic equilibrium state. That statement was supported by concrete observations of self-organisation phenomena which occur in physical systems, and it was accompanied with classical examples such as the Benard phenomenon[1] example. The observations of self-organisation cases in physical systems were complemented with more abstract and high-level analyses of self-organising behaviour in complex autonomous systems which were carried out by cyberneticians. These efforts resulted in the first conference on self-organising systems which was held in 1959 in Chicago and, as stated by Heylighen [43], was organised by the same multidisciplinary group of visionary scientists who had originally founded the discipline of cybernetics. Among the conference outcomes was the introduction of three fundamental principles concerning self-organising systems:

- According to the "*principle of self-organisation*", introduced by W. Ross Ashby, a dynamical system, independently of its type or composition, always tends to evolve towards a state of equilibrium. That evolution towards an equilibrium reduces the uncertainty regarding the state of the system, and therefore it decreases the system's statistical entropy. The resulting equilibrium can be interpreted as a state where the different parts of the system are mutually adapted. Such an equilibrium state has since been commonly known as an *attractor* [45, 52].
- Another cybernetician, Heinz von Foerster, formulated the principle of "*order from noise*". Foerster noted that, paradoxically, the larger the random perturbations ("noise") that affect a system, the more quickly it will self-organise (produce "order"). This can be intuitively explained as follows: the more widely a system is made to move through its state space, the more quickly it will end up in an attractor. If the system was not forced to move between states, no attractor state would ever be possible to be reached, and hence no self-organisation could take place.
- In addition to the above, Prigogine proposed the related principle of "*order through fluctuations*". According to that principle, nonlinear systems generally have several attractor states. When a system resides in between attractors, it will be in general a chance variation, commonly known as a "*fluctuation*" in thermodynamics [35], that will force it to move and stabilise in some particular attractor state.

At the same time further work on self-organisation was done in social systems. A representative example is the *theory of stigmergy* which was introduced by Grassé in 1959 after studying the social behaviour of termites [37]. Grassé summarised the theory of stigmergy in the phrase "*the work excites the workers*" meaning that the philosophy of stigmergy is to have members of a society attracted to places in the system where most activity has recently taken place. As a consequence, the activity at these points will be increased even further until the causes of the activity,

---

[1] The Benard phenomenon refers to the appearance of hexagonal cells or pattern rolls in liquids when heated from below. See Heylighen [43] for more detail.

for example the food availability when the goal of the societal activity is to collect food, at these points no longer exist. In this case, self-organisation refers to changes to the points where society members carry out their activities, which are dictated by attractions due to stigmergy. A main aspect of the theory of stigmergy is that interaction between society members can be effectively carried out through the environment and therefore direct communication is not mandatory. As a consequence, no direct interactions are necessary to coordinate a social group, and coordination and regulation tasks in a society can be realised based on information deposited into the environment without central control. In the case of ant colonies, stigmergy is realised by depositing chemical substances in the environment, which, as already mentioned in Sect. 2.1.2.3, are termed *pheromones*.

There is a fundamental difference between the concept of self-organisation introduced by Prigogine and the one suggested by Grassé. In the first case, self-organisation is the result of some external energy imposed to the system from its environment. In the second one, self-organisation is the result of internal actions of elements from within the system itself (for example in the case of an ant colony self-organisation actions initiate from the ants themselves which implement the stigmergy mechanism by depositing and following pheromones).

In further developments, Koestler established in late 1960s the definition of *holons* and *holarchies* [55]. As mentioned in Sect. 2.2.10.2 and detailed in Chap. 4, holons are both whole systems and parts of larger systems at the same time. Holarchies are hierarchies of such holons. Koestler presents a hierarchical view of self-organisation, which is applicable in a variety of cases ranging from enterprise organisations to the universe as a whole. The idea here is that order can result from disorder with progressive reorganisation of relations between complex structural elements at higher levels of abstraction. That high-level reorganisation results from statistically evolving simple relations between system elements at lower levels.

More specifically, Koestler defined holons as autonomous, self-reliant units that possess a degree of independence and handle contingencies without asking higher authorities for instructions. These holons are also simultaneously subject to control from one or more of these higher authorities. The first property ensures that holons are stable forms that are able to withstand disturbances, while the latter property signifies that they are intermediate forms, providing a context for the proper functionality for the larger whole. Holarchies consist of self-regulating holons that function first as autonomous wholes in supra-ordination to their parts, secondly as dependent parts in sub-ordination to controls from higher levels, and thirdly in coordination with their local environment.

In the 1970s Maturana and Varela, aiming to characterise the nature of living systems, introduced the notion of *autopoiesis* (literally meaning “auto (self)-creation”) as the process with which an organisation is able to reproduce itself [61]. Autopoiesis applies to closed systems, such as living organisms, consisting of autonomous components that interact and collectively (re-)produce higher-level components within the system. For example, the eukaryotic cell is made of various biochemical components such as nucleic acids and proteins, and is organised into bounded structures such as the cell nucleus, various organelles, a cell membrane and cytoskeleton. These structures, based on an external flow of molecules and energy, produce the components which, in turn, continue to maintain the organised bounded structure that gives rise to these components. An autopoietic system is to be contrasted with an allopoietic system, such as a car factory, which uses raw materials (components) to generate a car (an organised structure) which is something other than itself (the factory).

The notion of autopoiesis is often associated with that of self-organisation. The reason is that the dynamics of autopoietic systems resemble those of non-equilibrium systems; that is they include states, often called *dissipative structures* [43], which remain stable for long periods of time despite matter and energy continually flowing through them. However, an autopoietic system is autonomous and operationally closed, in the sense that every process within it directly contributes to maintaining the whole. Furthermore, autopoietic system elements are structurally coupled with their medium in a dynamic relationship that can be parallelised to sensory-motor coupling. In addition, autopoiesis presupposes an organised "self" [62], which, as Maturana argued, "*arises with the system*" and therefore "*cannot organise itself*" (see [21] referring to Maturana). Therefore, self-organisation is not autopoietic in its strict sense. Furthermore, since the mechanisms for self-organisation are inherent in the system itself and not modulated by an external organisation, it is commonly accepted that it is the system that organises itself [22].

In the 1980s the first attempts to apply self-organisation concepts to various domains were made. One such attempt is the development of "*neural networks*", which are simplified computer models (networks of artificial "neurons" linked with each other either directly or indirectly) mimicking the way brain neurons are linked and interact. There is no centralised control in neural networks, but they are capable of processing complex patterns of input. Another example is the production of laser light. Laser light beams result from the emission of photons of the same type, at the same time and at the same direction from atoms or molecules that are excited by an input of energy, and are synchronised under particular circumstances [43]. Based on observations of such collective results resulting from cooperation (or synergy) between system components, the German physicist H. Haken proposed the field of *Synergetics* to study such phenomena [41].

In the mid-1980s the proliferation of inexpensive and powerful computers enabled the use of computer simulation to explore self-organising system models of various degrees of complexity. Such systems typically included large numbers of interacting components and were difficult to be mathematically modelled. Based on this approach, a number of researchers associated with the Santa Fe Institute in New Mexico pioneered the domain of *Complex Adaptive Systems* (*CAS*) [45, 52]. CAS consist of many interacting components, which undergo constant change, both autonomously and in interaction with their environment. The behaviour of such complex systems is typically unpredictable, yet exhibits various forms of adaptation and self-organisation [47]. The components of CAS are typically considered to be agents [44, 47].

Examples of CAS include natural ecosystems comprising organisms which belong to many different species and interact with each other (compete or cooperate) and their shared physical environment. Additional CAS examples are markets, where different producers compete and exchange money and goods with consumers in a distributed and autonomous fashion. Despite markets being highly chaotic and nonlinear systems, they tend to converge towards approximate equilibria where conflicting goals of producers and consumers are mutually adapted and on average satisfied. This balancing mechanism, achieved by market dynamics, which Adam Smith, the father of economics, called "*the invisible hand*", can therefore be characterised as self-organisation [43].

The methods and tools used for the study of CAS enabled research in various related areas. The biologist S. Kauffman studied the development of organisms and ecosystems. Through simulations, he showed that sufficiently complex networks of chemical reactions will necessarily self-organise into autocatalytic cycles, the precursors of life [52, 53]. Another complexity theorist associated with the Santa Fe Institute, J. Holland, studied adaptation through natural selection of self-organising systems to a variable environment. By generalizing from the mechanisms through which biological organisms adapt, he provided the foundations of *Genetic Algorithms*[2] and developed models of cognitive, ecological and economic systems [45, 46]. Both Holland's and Kauffman's works have provided essential inspiration for the new discipline of *Artificial Life* (*AL*), which was initiated by C. Langton [56]. AL involves developing and executing software models mimicking lifelike properties, such as reproduction, sexuality, swarming and co-evolution, and subsequently concluding about real systems exhibiting these properties.

During the last 20 years, research in artificial systems has been oriented towards introducing self-organisation mechanisms specifically for software applications. These different works originate from different starting points including stigmergy [39, 59], autopoiesis [12] and the holon concept [11, 76]. Recently, in addition to reproducing natural system behaviour into artificial systems, latest research efforts have been oriented towards introducing self-organisation mechanisms specifically for software applications [29, 34]. Part II describes such mechanisms in more detail.

3.2.2 Definitions of Self-organisation

The concept of self-organisation has been defined in many disciplines, and it has been viewed from different angles. In the following the most common definitions found in the literature are presented, and a view of self-organisation from a software engineering perspective is provided.

---

[2] Genetic Algorithms is a general approach to computer problem solving which is based on mutation and recombination of partial solutions, and the subsequent selective reproduction of the most "fit" new partial solution combinations.

3.2.2.1 Overview of Self-organisation Definitions

Since the first appearance of the term, self-organisation essentially referred to a spontaneous, dynamically produced organisation in a system without external control. For example, as mentioned in the previous section, Kant used the term to refer to the autonomous formation of the planetary system without some central coordination point.

The formation of structures in physical systems was the basis of the view of self-organisation adopted later in thermodynamics as well. According to Prigogine, self-organisation resulted to alterations to structure of materials while energy was emitted to the environment in order to lower the entropy of the system [35]. A common example is *crystallisation*, the appearance of a symmetric pattern of dense matter in a system of randomly moving molecules [43], such as in the case of snow crystals [41]. Prigogine and his colleagues proposed four necessary conditions that must hold for self-organising behaviour to occur [65]:

1. The *Mutual Casuality* condition. A circular relationship must exist between at least two of the system elements resulting in mutual influence of each other.
2. The *Autocatalysis* condition: There exists at least one system component which is causally affected by other components, in a way that increases the quantity of the first component in a nonlinear fashion.[3]
3. The *Far-From Equilibrium* condition: the system imports a large amount of energy from outside the system, uses the energy to help renew its own structures (autopoietic behaviour) and dissipates, rather than accumulates, the accruing disorder (entropy) back into the environment.
4. The *Morphogenetic Changes* condition: The system must exhibit morphogenetic changes, namely the system components should be able to change their characteristics and behaviour [23], given appropriate conditions and stimulus.[4] To receive external stimulus, at least one of the system components must be open to external random variations originating in the system's environment.

Bonabeau et al. [10] provide a definition drawn from the behaviour of dynamically evolving natural systems. They consider self-organisation as a set of dynamical mechanisms whereby structures appear at the global level of a system as a result of interactions among its lower-level components. The rules specifying these interactions are executed on the basis of purely local information, without reference to a global pattern. Therefore, the resulting of structures is an emergent property of

---

[3] In Chemistry autocatalytic reactions are reactions in which at least one of the products is also a reactant. Such reactions are fundamentally nonlinear, and this nonlinearity can lead to the spontaneous generation of order. A dramatic example of this order is the one found in living systems. This spontaneous order creation initially seems to contradict the Second Law of Thermodynamics. However, this contradiction is resolved when the disorder of both the system and its surroundings are taken into account and it is recognised that energy is dissipated into the environment to lower the entropy of the system.

[4] The term *Morphogenesis* (from the Greek *morphe* = shape and *genesis* = creation) was originally introduced in biology to refer to a process that would cause an organism to develop its shape.

the system, rather than a property imposed upon the system by an external ordering influence. For example, in the case of foraging ants the emergent structures are the spatiotemporally organised networks of pheromone trails.

Bonabeau and colleagues identified four basic requirements for self-organisation:

1. *Positive feedback* (*amplification*) which involves simple behavioural "rules of thumb" that promote the creation of structures. Examples of positive feedback include recruitment to a food source and reinforcement to adopt a particular role observed in natural systems.
2. *Negative feedback* which counterbalances positive feedback and contributed to stabilising the overall collective pattern. Examples of negative feedback include satiation, exhaustion or completion. In the example of foraging ants in particular, negative feedback could result from the limited number of available foragers, the exhaustion of food source, and the crowding at the food source, or competition between food sources, respectively.
3. *Amplification of fluctuations* such as random walks, errors and random-task switching. In other words, small perturbations in system behaviour may increase and affect the collective system behaviour significantly. For example, when a new food source is found, the concentration of pheromones in the pheromone path formed increases significantly.
4. *Multiple interactions* among individuals which are the basis for the collective behaviour. More specifically, self-organisation generally requires a minimal density of mutually tolerant individual behaviours capable of using the results of both their own activities and the activities of others. For example, trail networks can self-organise and are formed by pheromones deposited by more than one ant.

Similar definitions drawn from natural systems are adopted by many authors, for example in [16].

In other domains self-organisation is typically defined as the evolution of a system into an organised form in the absence of external pressures [69]. This is explicitly stated by Haken [40], who considers a system as self-organising if "*it acquires a spatial, temporal or functional structure without specific interference from the outside*". As another example, Polani [68] considers self-organisation as a phenomenon under which "*a dynamical system exhibits the tendency to create organisation* '*out of itself*' *without being driven by an external system*".

Along the same line Heylingen [43] defines self-organisation as the "*spontaneous creation of a globally coherent pattern out of local interactions*". Because of its distributed character, this organisation tends to be robust, resisting perturbations. The dynamics of a self-organising system is typically nonlinear, because of circular or feedback relations between the components. Positive feedback leads to an explosive growth, which ends when all components have been absorbed into the new configuration, leaving the system in a stable, negative feedback state. Heylighen stresses that the basic mechanism underlying self-organisation is the (often noise-driven) variation which explores different regions until it enters a stable state (an attractor state).

In a similar manner Collier [22] defines self-organisation as "*a process by which larger scale* (*macro*) *order is formed in a system through the promotion of fluctu-*

ations at a smaller (*micro*) *scale via processes inherent solely in the system dynamics, and modulated by interactions between the system and its environment*". Self-organisation occurs when the properties of a system allow it to take on a more ordered state through the dissipation of energy (production of entropy) some of which goes into the newly formed structure.

However, not all definitions of self-organisation involve spontaneous creation of organisational structures. Maturana and Varela introduced *autopoietic systems* [61] where some form of organised system must pre-exist for subsequent changes to the organisation to take place, in a manner similar to *re-organising systems* discussed by Collier in [21]. Maturana and Varela view autopoietic systems as networks of interacting processes that are capable of producing (generating, transforming and destructing) other, possibly similar, processes and that satisfy the following requirements:

1. They are continuously regenerated and transformed through interaction and reproduction of its members processes.
2. They constitute a concrete unity in the space where the member processes exist and the topological domain of their associations is specified [75].

In other words, in autopoietic systems self-organisation refers to dynamic system regeneration and transformation to some known form.

Finally, in many computer applications self-organisation has been used to the process of changing the system's organisation based on some internal central point of planning or control within the system, as is the case for example in [1, 60]. This approach is followed in many cases in natural systems[5] and has certain advantages, such as better coordination efficiency and control. However, it can suffer from the known centralised point weaknesses such as robustness and performance.

In other cases, changes to the organisation of the system are the collective result of distributed autonomous points of control within the system. Examples include self-organised coordination [49], multi-agent system cooperation [71] and multi-agent system group formation [58].

3.2.2.2 Software Definition of Self-organisation

Self-organisation is an attractive approach to handle the dynamic requirements in software. Considering the definitions given in other domains, the term self-organisation can be used in software engineering to refer to a process where a software system changes its internal organisation to adapt to changes in its goals and the environment without any explicit external directing (command and control) mechanism. Therefore:

---

[5] For example, in societies of termites the queen selects a location and deposits pheromones symmetrically at equal distances resulting in nest arches being equally distanced from the queen location [10].

> *self-organisation* is the process enabling a software system to dynamically alter its internal organisation (structure and functionality) during its execution time without any explicit external directing mechanism.

Furthermore, due to the wide range of self-organising software systems, it is necessary to further distinguish them into those involving no explicit centralised control (either external or internal) and those based on some internal centralised coordination and planning point. By consequence the following definitions are given:

> *Strong self-organising systems* are systems where self-organisation process decisions are distributed locally among the system components without involving any centralised point of control (either internal or external).

> *Weak self-organising systems* are those systems where, from an internal point of view, self-organisation is internally administered by a centralised point of planning and control.

It is important to emphasise that the central point of control in a weak self-organising system is not visible from an external observer located outside the system itself. For example, the users of a cluster-based e-commerce web site that self-organises to balance the incoming user load will not notice any difference regardless whether load is balanced in a centralised or a distributed manner within the system.

Typical examples of strong self-organising systems are those implementing ant algorithms. In such systems there is no internal point of central control since they are built with the aim to mimic the behaviour of ant colonies when foraging. Representative examples of weak self-organising systems are those based on architectures involving centralised control such as general mediator systems, client-based systems and systems based on star and hierarchical architecture topologies.

Self-organisation can often result in emergent software behavioural patterns that can be either desirable or undesirable. Due to the dynamism and openness of contemporary software environments and the ever increasing distribution, complexity and dynamic changes in application requirements, understanding the mechanisms that can be used to model, assess and engineer self-organising behaviour in software is an issue of major interest.

3.2.3 Properties of Self-organising Systems

Self-organising systems are distinguished from traditional mechanical systems studied in Physics and Engineering by a number of mandatory and optional properties, which can be considered as part of what defines them.

3.2.3.1 Mandatory Properties of Self-organising Systems

Self-organising systems are characterised by the following fundamental properties:

- *Global Organisation.* The process of self-organisation brings the system into some ordered and relatively stable state, in which it can fulfil its function and purpose given the constraints imposed from its environment. For example, if we consider a system consisting of a light bulb and having the purpose of providing light to a warehouse, then satisfactory organised states for the system would be one where the light bulb would be switched off during the day and one where it would be switched on during in dark hours. The resulting system organisation can be *static*, where the organisational positions of the system components are fixed, as is the case in the light bulb example, or *stationary*. In the latter case, system components continuously change their organisational positions, but this is done in some ordered fashion and according to some stable organisational pattern, such as the Benard cells pattern (see Sect. 3.3.2.1). The Benard cells pattern refers to the formation of hexagonal cells which appear when a liquid is evenly heated from below and cooled on its surface [43]. In Benard cells there is an upward flow of liquid on one side and a downward flow on the other. Despite the continuous movement, the Benard cell pattern remains stable as long as the heating and cooling of the liquid are not altered.

- *Dynamic Adaptation*. Self-organising systems are capable of changing their organisation dynamically to adapt to changes in their intended function and the conditions of their environment. For example, insect populations behave differently when mating than when collecting food [10], and their behaviour further varies according to environmental conditions, such as ground and weather conditions, and time of day.

  A common question then is what behaviour can be considered as part of the dynamic reorganisation process that realises system's adaptation to the current environmental conditions, and what is considered as normal, possibly dynamic, application behaviour that realises the system's intended function. This question becomes hard to answer since reorganisation and application behaviours are largely interweaved, and the former commonly affects the latter, namely application behaviour is often modified as a result of reorganisation.

  In simple systems, dynamic reorganisation can be perceived and modelled as normal application behaviour, although this is not generally a good practice since it increases model complexity. For example, let us consider a sensor attached to a light bulb capable of emitting light linearly and in inverse proportion to the amount of natural light sensed. The emitted light will gradually increase when the amount of light in the surrounding space decreases, for instance if natural light is reduced due to cloudy weather. One could intuitively view the increase in light emittance level as system reorganisation, considering the level of emitted light as an organised system state and the actual light emittance as the functionality delivered by the system. One other could argue that the proportionally adjusted light emittance is simply predesigned application behaviour and that changes in natural light availability are simply external input and not changes in the system's environment. In such simple cases, where the possible reorganisation results and the stimuli that triggers reorganisation can be perceived or calculated in advance, reorganisation and application behaviour can indeed be a matter of definition and modelling perspective. For instance, in the above example the light emittance can be exactly calculated as a linear function of the absence of natural light, and hence it can be easily perceived as application behaviour.

In general, however, it is not always possible to exactly determine the system configuration that would result from the reorganisation process in advance. For example, it is difficult to predetermine the exact path that will be formed by foraging ants since the number of possibilities is too high. In these cases, dynamic reorganisation is considerably harder to be modelled as application functionality, since the adapted organisation that will realise application functionality is not known in advance. Therefore, it is a good practice to follow intuition and model reorganisation separately whenever possible because that increases modularity, ease of understanding and separation of concerns. For example, in an ant-based system it is easier to model the behaviour of searching for food and the behaviours of carrying and processing food separately.

- *Lack of External Control.* The dynamic reorganisation taking place in self-organising systems is executed without any external control, and the order produced is endogenous. For example, if we have some external operator switching the light bulb on and off, then the light bulb system would not be self-organising. On the other hand, if the system comprises both the light bulb and a sensor device capable of switching it on and off according to the light conditions, then that is a clear self-organising system example.

  Consequently, it is very important to define the system boundaries before characterising a system as self-organising or not. To clarify this with an additional example, if we consider a central light sensor connected to a computer that remotely operates light bulbs in different rooms, then each individual light bulb alone is not a self-organising system. However, if we considered a larger system comprising all light bulbs and the operating computer, then we would have a clear case of a weakly self-organising system.

3.2.3.2 Optional Characteristics of Self-organising Systems

In addition to the above mandatory properties that allow us to determine if a system is actually self-organising, there are several characteristics that can appear in self-organising systems, either natural or artificial, in various combinations.

- *Nonlinearity and Complex Interactions.* Systems showing instability are characterised mainly by nonlinear dynamics (small fluctuations when near some critical point can provide significant modifications of the system), by sensibility to initial conditions and parameter sensibility (small changes on a parameter produce different patterns). Thus the overall properties cannot be understood simply by examining separately the components.

  Self-organising systems commonly exhibit complex, nonlinear behaviour which enables them to adapt to a larger range of environmental conditions. Therefore, the more complex the overall system behaviour, the higher the adaptation capability of the system.

  Nonlinear system behaviour can trivially be the additive result of nonlinear behaviours of individual system components. For example, let us assume a number of pipes flowing water into an irrigation basin, and that the flow of each pipe is controlled by an intelligent sensor device. The device regularly senses a number of environmental parameters and controls the water flow according to some non-linear irrigation model by operating a valve attached on the pipe. Depending on the environment parameter values, the flow of water through each pipe will vary, and the overall water flow will be the sum of flows of all water pipes.

More complex collective system behaviours can be achieved when the number of system components increases. For example, let us assume that we have dozens of light bulbs distributed in some area. Each bulb is associated with a sensor and it is able to emit different amounts of light according to the level of natural light perceived by its sensor. Assuming that the levels of natural light are not the same in all locations within the designated area, that system will be emitting only the necessary light in each location, saving thus in energy as compared to a system having only one central strong light bulb. It is intuitively obvious that by increasing the number of light bulbs in the area we can achieve higher light emittance granularity and detail, and hence better energy savings. This is generally the case: the more individual components are included in the self-organising system, the more sophisticated overall collective behaviour can be achieved.

The complexity of the overall system behaviour increases even further when there are interactions between system components which affect their individual behaviours. For example, instead of having each pipe controlled by a sensor executing some independent irrigation model, sensors could interact with each other; for instance, they could be connected nodes in a neural network that would be exchanging data and influencing each other. That would result in more complex overall behaviour, such as implementation of a more complex irrigation model, since there would be interrelations between individual behaviours.

- *Decentralised Control*. Organised systems are commonly associated with some hierarchy of internal or external controllers which guide and direct them. The controller hierarchy is a physically distinct subsystem that exerts its influence over the rest of the system. In these cases the control is characterised as *centralised*. For example, most business organisations have leaders, boards of directors and unit managers that develop policies and coordinate business operations. An additional example is the human body whose actions are largely planned and controlled by the brain.

  In weak self-organising systems control is centralised. In strong self-organising systems however, control is distributed over the whole system. All system components contribute evenly to the resulting arrangement. For example, in spin magnetisation all spins influence each other to maintain the overall spin direction which has been reached and no spin or group of spins can deviate from it. Another example is the human brain which is organised over a network of interacting neurons. Different brain regions are specialised for different tasks, but no neuron or group of neurons has overall control. This is demonstrated in the cases where some brain parts get damaged, for example due to some accident, and the whole brain functioning continues unaffected.

  Centralised control offers advantages, such as more autonomy and stronger specialisation for the controller [43]. Furthermore, in decentralised control, units are prone to opposing actions, their behaviour may induce needless redundancies, and it is not guaranteed that the globally optimal solution will be found [25]. However, for systems that are complex and operate dynamic environments, the use of decentralised control offers significant advantages such as increased scalability and robustness and reduced communication and unit processing costs [13].

- *Simple Behaviours and Local Interactions*. In addition to acting autonomously in a decentralised manner, system components commonly have simple behaviours with limited perception abilities, and hence they do not have a global view of the system. For example, ants in an ant foraging colony can perceive the existence of pheromone or food only at a short radius. Spins in a piece of magnetised material exert non-negligible influence on their near neighbours. In the Benard example, liquid molecules influence only a few molecules that they collide with. Yet, most ants eventually move along an emerging path, the piece of material becomes magnetic as a whole, with a single North pole and South pole, and the liquid as a whole starts cycling through a sequence of rolls. In other words, despite the locality of interactions, the system reaches a global organised state where all components are significantly correlated.

  The locality of interactions implies that neighbouring configurations are strongly correlated, and this correlation diminishes as the distance between configurations increases. Furthermore, for an external influence to transfer from one region of the system to another, it must pass through all intermediate regions, and it will be affected by all fluctuations and activity that is taking place in them. For example, if we randomly place an obstacle on the foraging ants path, ants will progressively form a different path to avoid the obstacle, and this will eventually affect the majority of foraging ants. Ants closer to the point where the obstacle was placed will be affected more significantly, and the effects will be less as we move away from the obstacle.

  Simple interacting behaviours can still lead to quite complex patterns. Furthermore, they have the advantage that the information stored in the description of each behaviour is less than the information needed to describe the pattern: it is only needed to store information about how to produce the pattern, not the pattern itself.

- *Robustness, Resilience*. Self-organising systems consisting of large number of interacting components can be particularly robust, namely they can be relatively insensitive to errors or perturbations from the environment. For example, if we remove some bees from a bee swarm or an ant colony, the harvesting result will be the same. In another example, if some nodes or links of a neural network are removed, the task the network was trained to perform will be largely still carried out. A reason for robustness is the *redundancy* inherent in such distributed organisation systems, and the remaining components can cover for the removed ones.

  Furthermore, such systems are commonly resilient, namely they have a strong capacity to restore themselves. For example, an ecosystem that has been damaged from a fire will in general recover in a short time. In a magnetic material, if part of the spins are diverted from their alignment, the magnetic field produced by the rest of the spins will put them back. Apart from redundancy, another reason for resilience is the random fluctuations, or "noise". For example, foraging ants can move to some random direction with some small probability. This randomness allow them to initially discover the food source and repair the emergent path in case it gets damaged. Finally, a third reason for resilience is the stabilising effect of feedback loops which will be described below.

- *Feedback Loops*. In complex, nonlinear systems the relation between cause and effect is not straightforward, and small causes can have large effects and vice versa. For example, let us assume that a magnetised piece of iron is subjected to an external magnetic field with a direction different from its own field. In the beginning, a large increase in the external field has practically no effect until a threshold is crossed. After that, any small further increase suddenly reverses the polarisation of the whole system.

  This is due to the *feedback* relation which commonly holds between components of self-organising systems. Each component, for example a spin in a magnet, affects the other components, but these components in turn affect the first component. This *cause and effect* is circular. Any change in the first components is fed back via its effects to the other components to the first component itself.

  Feedback can be either *positive* or *negative*. In positive feedback the recurrent influence reinforces or amplifies the initial change. In other words, if a change takes place in a particular direction, the reaction being fed back takes place in that same direction. In negative feedback the reaction is opposite to the initial action, namely the original fluctuation is suppressed or counteracted, rather than reinforced.

  Positive feedback makes deviations grow in a runaway, explosive manner. It leads to accelerating development resulting in a radically different organisation. Negative feedback, on the other hand, stabilises the system by bringing deviations back to their original state.

- *Emergent Properties*. The resulting organisation can give rise to particular organisational constructs, such as structures, patterns or other system properties that cannot be reduced to the properties of its elements. The emergent outcomes exist and can be perceived only after self-organisation has taken place. For example, Benard cells are created after self-organisation of the liquid, and they are characterised by the direction of the molecules rotation, which is an emergent property. For an independent molecule, such rotation is not defined.

  A higher-level emergent property typically constrains the behaviour of the lower-level components. For example, the overall rotation characterising a Benard roll will force the liquid molecules to move in particular directions instead of others. This is called *downward causation*. Downward causation refers to a higher organisation level exerting influence downwards to a lower organisation level, causing in the Benard rolls example the molecules to move in a particular way.

  Random perturbations are crucial for emergence, since they enable discovery of new solutions, and fluctuations can act as seeds from which structures nucleate and grow. For example, foraging ants may get lost by following trails with some level of error, and this phenomenon may lead them in finding new, inexplored food sources and recruit nestmates to these food sources. In the general case, self-organisation can be witnessed without emergence [27]. For example, in the light bulb system there is no property that occurs only after system self-organisation. Further discussion on the emergence concept is provided in Sect. 3.3.

- *Symmetry Breaking*. Non-organised systems are originally in a disordered configuration meaning that possible states for the individual components have the same probability. Therefore, despite differences in component behaviours on the global, macroscopic level, the system is considered homogeneous and symmetric since every behaviour is represented at the same degree, and hence it will look "the same" from whatever direction it is observed.

  After self-organisation, however, the global organisation configuration dominates all others, and therefore, the symmetry in probability is lost. For example, the probability of spin direction in a magnet will be either one or zero. This is referred to as "*symmetry breaking*".

- *Bifurcations*. The evolution from disordered to ordered configuration is normally triggered by a change in the external situation, the *boundary conditions* of the system. That evolution can take place through different possible sequences of organisational states, the system may be settling in during self-organisation, which can be conceived as *evolution paths*.

  We can have system variables whose values are associated with the states the system would settle in. These variables are called *order parameters*. It is possible that when order parameters reach a particular value, any further increase in their value would result in system evolution along different paths that have been arbitrarily chosen when the particular value was reached. If we represent the possible system states vs the values of an order parameter in a diagram, we can notice one or more *clear branchings* or *bifurcations*.

  For example, the speed of molecules in a Benard cell can be increasing with increasing temperature to a direction arbitrarily chosen when temperature had reached a particular threshold (see also Fig. 3.1 taken from [51]).

- *Far-from-Equilibrium Dynamics*. In the absence of external perturbation, the system is expected to stabilise in some states in which emergent properties can be observed. This implies a kind of dissipation of some "energy", otherwise the system would be continuously changing.

  Self-organising systems can reach a static global organisation, or *equilibrium*, or stationary involving ongoing activity. Equilibrium is characterised by the absence or entropy production or, equivalently, by the fact that all excessive energy has been dissipated to the environment.

  When a constant input of energy is provided in the system, as for example in the Benard cell case, the system cannot reach equilibrium, and it constantly dissipates energy to its environment. Such systems are termed *Far-from-equilibrium* systems. A far-from-equilibrium system is in principle capable of producing a much greater variety of regulating actions leading to multiple stable configurations. However, due to the dependency on an external source of energy, such systems are more fragile and sensitive to changes in the environment and also more dynamic and capable to react.

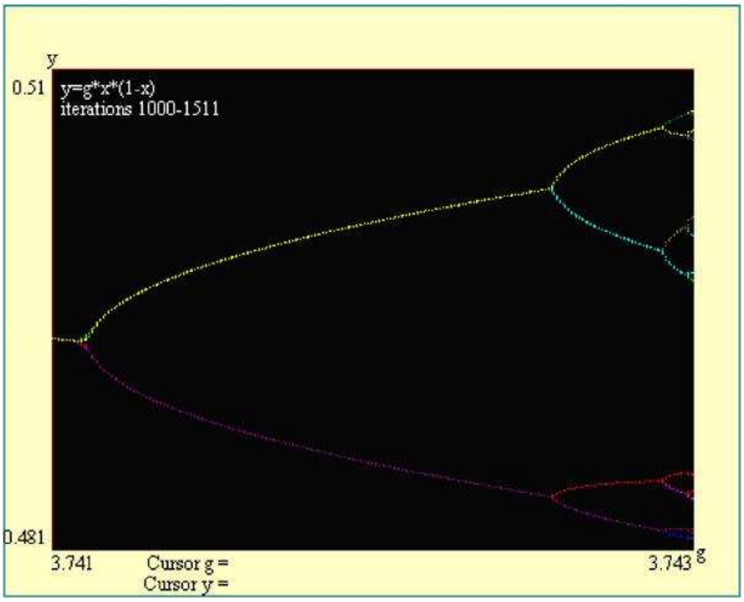


Fig. 3.1 Feigenbaum Cascade bifurcations

Complex self-organising systems tend to reside on the "*edge of chaos*". The "edge-of-chaos" is the narrow domain between equilibrium and turbulent, chaotic activity. The mechanism by which complex self-organising systems tend to maintain on this critical edge has been called "*self-organised criticality*".

The system behaviour on the edge of chaos is typically governed by a "*power law*" [43] which states that large adjustments are possible but they are much less probable than small adjustments.

- *Adaptability*. The problem of adaptability is to maintain a particular organisation in spite of environmental changes, and the question involved is which action to use in which circumstances. It can be modelled as a problem of regulation or control to achieve "*self-regulation as to operate the same in a changing environment*" [43]. Not all self-organising systems are adaptable. For example, magnets and Benard cells are adaptive but not adaptable since they do not maintain the same function for changing environments.

  In various disciplines such as organisational management and ecology, adaptability is described as the ability to cope with unexpected disturbances in the environment while maintaining a particular organisation. For example, the immune systems are adaptable systems since they maintain the organisation (health) of the living organism despite any perturbations organisms receive from their environment through contact with viruses.

  To find the right type of antibodies, immune systems simply produce a vast variety of different antibody shapes. However, only the ones that "fit" the invaders are selected and reproduced in large quantities. The aim is to minimise deviations from a goal configuration by counteracting perturbations before they become large enough to endanger the essential organisation.

  This means that:

1. The system must produce a sufficient variety of actions to cope with each of the possible perturbations.

2. The system must select the most suitable counteraction for a given perturbation.

  Variety can be fostered by keeping the system sufficiently far from equilibrium so that it has plenty of stationary states to choose from. Selectivity requires that these configurations are in sufficiently small numbers and sufficiently stable to allow an appropriate one to be selected without risk of modifying the overall organisation.

  Another typical example of simple variation and selective reproduction of fit components underlies economic and ecological adaptation mechanisms.

- *Organisational Closure and Hierarchy*. Self-organising systems at some point settle into a negative feedback regime or into an attractor state. Subsequently, they become relatively impervious to external disturbances and largely independent from their environment. Therefore, they can be considered that they are “closed” against relatively small influences from the outside.

  For the outside observer, closure determines a clear distinction between inside (components that participate in the closure) and outside (those that do not). In other words, closure defines a boundary separating the organisational structure from the environment, This boundary can encompass all components of the original system, for example in case of a magnet, or only part of them as in a Benard cell.

  A self-organising system may settle into a number of relatively autonomous, organisationally closed subsystems. These subsystems will continue to interact in a more indirect way, determining subsystems at a higher hierarchical level, which contain the original subsystems as components. For example, a cell is an organisationally closed system encompassing a complex network of interacting chemical cycles within a membrane that protects them from external disturbances. Cells are themselves organised in circuits and tissues that form a multicellular organism. Organisms are further connected in cyclical food webs, collectively forming ecosystems.

3.3 Emergence

Our aim is to analyse the emergence concept from two perspectives: One concerning the properties that when observed are sufficient to identify emergence, and another focusing on those system characteristics that when all present the system has the capability to produce emergence. Subsequently, we attempt to provide an operational definition of emergence.

3.3.1 History

Emergent properties were first studied in the Greek antiquity in the context of philosophical treatises [24, 36]. For example, while referring to the significance of “wholes” in the natural world, Aristotle wrote: “*…the totality is not, as it were, a mere heap, but the whole is something besides the parts …*”,[6] which is interpreted as “*the whole is something before, over and above its parts, and not just the sum of them all.*” [24] or simply “*…the whole is more than the sum of its parts …*” [27, 30].

The expression “*whole before its parts*” refers to the ontological distinction between parts and wholes and in particular to the explanatory precedence of a whole entity over the parts of which the whole is made up. This is exemplified in the response given by Aristotle to a famous paradox introduced by Zeno. As [36] details, Zeno argued that a distance of any length could be divided into an infinite number of shorter segments. Hence, covering the distance would require traversing an infinite number of shorter segments, which, as a result, would take an infinite amount of time. That was obviously paradoxical since we do cross distances in finite lengths of time. Aristotle’s response was that a length was first and foremost a *whole*. That whole might indeed be divided into an infinite number of parts; however, the whole was fundamentally irreducible to those parts, and in fact it was only due to the pre-existence of the whole that it could be traversed.

The pre-existence of a coherent whole that cannot be described merely as a sum of its constituent parts was later endorsed in the concept of *Gestalt*, which originated from theories of Johann Wolfgang von Goethe, Immanuel Kant and Ernst Mach [9, 30, 36] and refers to whole forms or global configurations. Initially, the term ‘Gestalt’ referred to secondary qualities that emerged from the parts constituting a whole. For example, Goethe considered a Gestalt to be “*a natural unity that was the endpoint of an entelechetic development out of primordial chaos*” (see [36] referring to [42]). This had some similarities with current notions of emergence, for example with the one adopted in complex systems where it is considered that order emerges from chaos [46, 53].

After various ramifications, the term Gestalt was accepted as a perceptually primary concept, defining the parts of which the whole is composed. For example, the founder of modern Gestalt psychology Christian von Ehrenfels illustratively remarked that perception takes place through recognising whole patterns: “*the whole is greater than the sum of the parts*” (see [36] referring to [42]). In modern Psychology the *Gestalt effect* refers to the form-forming capability of our senses, particularly with respect to the visual recognition of figures and whole forms instead of just a collection of simple lines and curves. A typical example given to demonstrate the Gestalt effect is that of a picture of a dog, which can be perceived as a whole at once, and it is not recognised by identifying its parts first (such as feet, ears, and nose), and only then inferring the dog from those component parts [74].

Neither Gestalt nor the Aristotelian view of emergence explain how the emergent outcome is produced. Furthermore, they both assume the existence of a pre-given whole, and therefore they capture emergence only in a static sense. The dynamic aspect of emergence is considered in the technical definition of the term which was

---

[6] Aristotle, *Metaphysics*, Book H 8.6.1045a:8-10.

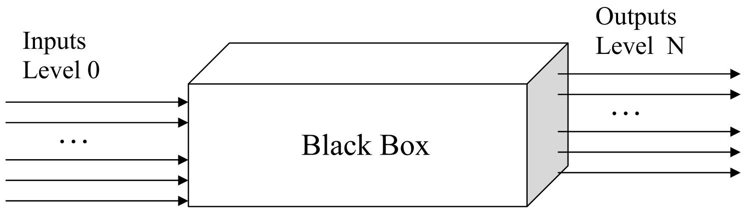


Fig. 3.2 The Proto-emergentist view of emergence

given by the English philosopher G.H. Lewes in 1875. Lewes was based on an earlier philosophical work of J.S. Mill concerning differentiation of types of causation, and he classified chemical compounds resulting from chemical reactions as either *resulting* or *emergent* (see [36] referring to [57]):

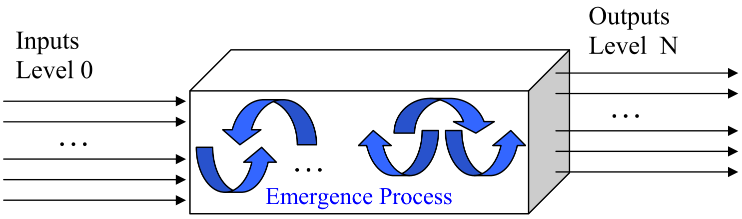
> although each effect is the *resultant* of its components, we cannot always trace the steps of the process, so as to see in the product the mode of operation of each factor. In the latter case, I propose to call the effect an *emergent*. It arises out of the combined agencies, but in a form which does not display the agents in action ... Every resultant is either a sum or a difference of the co-operant forces ... [and] is clearly traceable in its components ... the emergent is unlike its components insofar as these are incommensurable ... and it cannot be reduced either to their sum or their difference.

The newly defined concept was widely adopted in late 19th century, and it was placed in the core of a loosely joint movement covering social sciences, philosophy and theology which was originally termed *emergent evolutionism* (see [9] for history and review) and later known as *proto-emergentism* [27, 30, 36].

Despite that several proto-emergentists employed the term ‘Gestalt’ for describing emergent phenomena, the proto-emergentist view is radically different. The emergent process is viewed as a black box (see Fig. 3.2). A number of intermediate system organisation levels is considered, but only the inputs and outputs at the lowest and highest levels respectively can be realised, and the emergent outcome is formed dynamically without any knowledge of how the entries are transformed to outputs. Many well-known researchers including G.H. Lewes, C.L. Morgan, J.S. Mill, S. Alexander, D. Broad, W. Wheeler and A.N. Whitehead participated in that movement trying to explicit the characteristics of emergent phenomena. However, the lack of adequate explanations regarding the causes of emergence resulted in the deterioration of the movement in early 20th century.

After 1930 a different perspective started gaining ground, realised by a movement termed *neo-emergentism* [27, 30, 36]. Neo-emergentism had its roots in approaches followed to study phenomena and dynamics of systems in positive and computer sciences. In contrast to proto-emergentism, this movement aimed at understanding and reproducing the processes which lead to emergent phenomena. The view is that the unique features of emergence can be better apprehended by sketching out its scientific and mathematical sources, as is done in complexity theory for instance. Representative examples of neo-emergentist approaches include the works of H. Haken, J. Holland, S. Kauffman, C. Langton and I. Prigogine. Their efforts concentrated on developing theories, tools and methodologies which aimed at enabling the expression of emergent processes as less dense using more primitive models and consequently as less miraculous (see Fig. 3.3). The neo-emergentism movement is still

Fig. 3.3 The Neo-emergentist view of emergence



actively developing having so far shown important advances in modelling, engineering and controlling emergence [29, 32, 54, 79].

In the early neo-emergentist approaches developed in cybernetics, information theory and general systems theory, emergent phenomena were not the explicit focus of research per se. The reason is that these earlier approaches investigated systems which were simple, linear and equilibrium seeking [36]. In time the interest gradually shifted to complex, nonlinear and nonequilibrium systems, for example such as those targeted by complexity theory [45]. The basic idea in neo-emergentism is that the dynamical characteristics of emergence can be better understood by considering its association with the arising of attractor states during system operation which are not pre-given in the sense of a Gestalt. Attractors technically pre-exist of course; however, they are not possible to be known or accurately predicted in advance due to system complexity,[7] and they reveal themselves when a dynamical system bifurcates. The convergence to an attractor state signifies both a quantitative and a qualitative metamorphosis of the system. Upon their appearance, these "new" attractor states dominate the system and hence enable the emergence of something radically novel in respect to what existed before [46, 53, 64, 65].

An important strand in neo-emergentism is *connectionism* [70]. Connectionism is a set of approaches originated in the 1980s in the fields of artificial intelligence, cognitive psychology, cognitive science, neuroscience and philosophy of mind that models mental or behavioural phenomena as the emergent processes of interconnected networks of simple units. There are many forms of connectionism, but the most common forms use neural network models (see for example [48]). Although the basic idea of connectionism has received fierce criticism from the proponents of the symbolic AI school, connectionist models are believed to be a step in the direction toward capturing the intrinsic properties of the biological substrate of intelligence, since they have been inspired by biological neural networks and seem to be closer in form to biological processes. Furthermore, such models are capable of handling incomplete, approximate, and inconsistent information, as well as generalisation. Their main weakness however is the difficulty in development, which is still an open research issue [14, 27, 32, 79].

---

[7] Complexity can be simplistically perceived as the minimum amount of information necessary for system description. Various definitions of complexity exist; for example see [4].

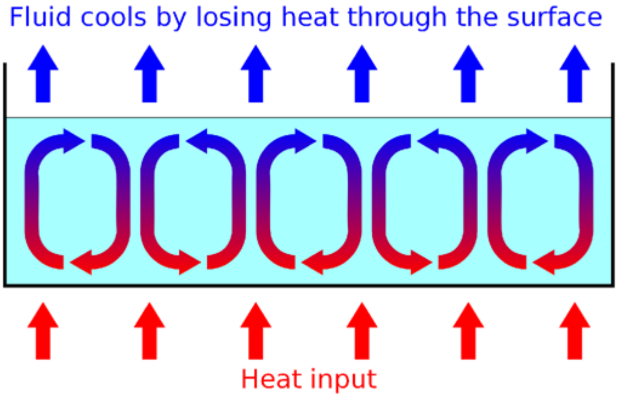


Fig. 3.4 Benard Cells appear at macroscopic level when temperature reaches a critical point

3.3.2 *Definitions*

3.3.2.1 Examples of Emergent Phenomena

Phenomena that can be characterised as emergent include organisational structures and frameworks, behavioural processes, particular system states, and even outcomes corresponding to functions[8] not explicitly implemented in the system.

The foraging ant path and Benard cells are representative examples of emerging organisational structures. As already mentioned in Sect. 3.2.3.1, Benard cells are convection cells that appear spontaneously in a liquid layer when heat is applied from below (see Fig. 3.4 taken from [2]). The setup includes a layer of liquid, for example water, between two parallel planes. Initially, the temperature of the bottom plane is the same as the top plane. The liquid then tends towards an equilibrium, where its temperature is the same as that of the surrounding environment. Once there, the liquid is perfectly uniform, and it appears the same from every direction. When the temperature at the bottom increases up to a critical point, then hexagonal convection cells appear (see Fig. 3.5). The microscopic random movement of the liquid molecules spontaneously becomes ordered on a macroscopic level, with a characteristic correlation length. The rotation of the cells is stable and will alternate from clock-wise to counter-clockwise as we move horizontally along the liquid. If the warmth increases, this phenomenon disappears.

Typical examples of emergent behavioural processes and system states can be found in the well-known *Game of Life*, which is a cellular automaton devised by the British mathematician J.H. Conway in 1970.[9] The game consists of a collection of cells which are placed on a board with a connexity of 8. Cells can have the status of ‘dead’ or ‘alive’, and at each game round they can change their status and affect the status of their neighbouring cells according to a few simple mathematical rules. Alive cells are coloured, and depending on the initial conditions, they can form

---

[8] The term *function* here refers to a process or operation and not to some direct input–output mechanism such as a mathematical function. For example, an foraging ant-colony can find the shortest path between food source and nest without it being an explicit function of the system.

[9] http://www.bitstorm.org/gameoflife/

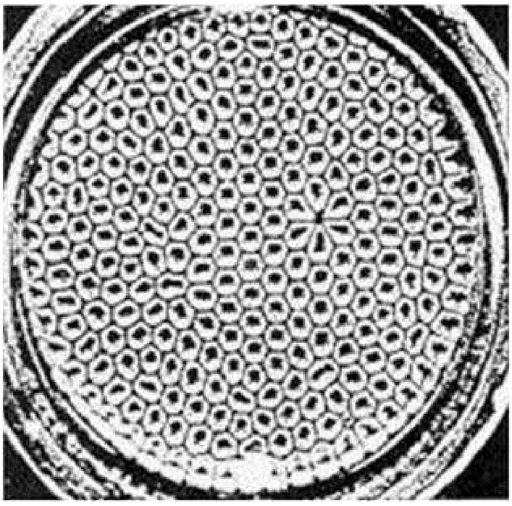


Fig. 3.5 Benard cells appearing in heated water

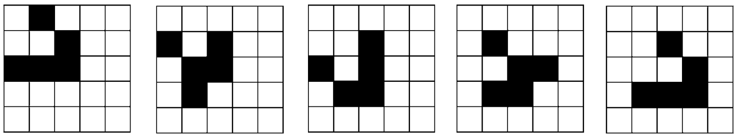


Fig. 3.6 Glider rotation in the Game of Life

various patterns throughout the course of the game. Experimentation with the game involves creating an initial cell configuration and observing how it evolves during game execution. Based on different initial configurations, several types of emergent states, patterns and behavioural processes can be observed [5, 7, 46].

A representative emergent pattern in Game of Life is a *glider* which appears to be moving across the board (see Fig. 3.6) diagonally at a speed of 1/4 of cell at each round. The glider pattern is often produced from randomly generated starting configurations [7]. Gliders are important to the Game of Life because they are easily produced, they can be collided with each other to form more complicated objects, and they can be used to transmit information over long distances. Other characteristic emergent patterns appearing in the game of life include “still lives”, oscillators and patterns that translate themselves across the board (“spaceships”) [5, 7].

A typical process that can emerge during the game of life is *Glider Gun* which produces gliders repeatedly (see Fig. 3.7). A glider gun is a pattern consisting of a main part that repeats periodically like an oscillator and emits gliders at regular time intervals. The discovery of the glider gun process eventually led to the proof that Conway’s Game of Life could function as a Turing machine. An example of emergent state in the Game of Life is the *R-predomino* pattern (see Fig. 3.8). R-predomino acts as a glider gun producing exactly six gliders in 1103 game rounds. The pattern subsequently settles down to a stable state, and any movements it makes on the board are restricted into a 51-by-109 cell region. This is a representative example of an emergent phenomenon that consists of an emergent state. There are numerous examples of function outcomes obtained indirectly through emergence, such as emergent construction of course timetables [67] and development of manufacturing schedules [64]. In these cases the system is modelled as several components which interact locally, for example along the lines of an ant colony. The outcome of interest, such as timetable, manufacturing plan, assembly movement sequence and

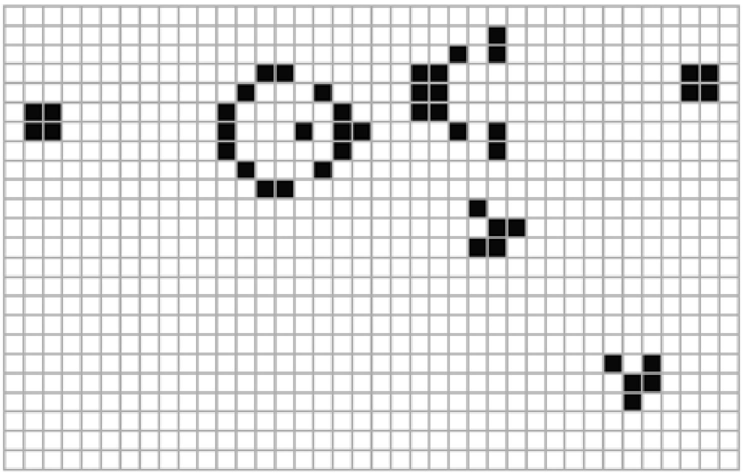


Fig. 3.7 Snapshot of a Glider Gun emitting gliders

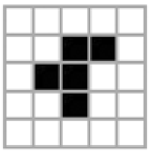


Fig. 3.8 The R-predomino pattern

machine configuration, corresponds to a global property which obtains the desired value as a result of the emergence process that takes place based on component interactions.

3.3.2.2 Overview of Emergence Definitions

Definitions of emergence can be broadly classified in two categories based on their view of *emergent properties* [50]. In the first category any system properties are considered as emergent when they cannot be analysed (or "reduced") through functional decomposition. In these cases it is said that, "... *the system is more than the sum of its component parts* ..." [5, 36, 44]. On the other hand, definitions belonging to the second category accept that "... *any unexpected properties are emergent* ..." [18, 19, 50, 73].

A representative example of the first category is the view of emergence introduced by proto-emergentists (also termed British emergentists). The proto-emergentist view was inspired by properties of living systems where "... *no mere summing up of the separate actions of system ingredients will ever amount to the action of the living body itself* ..." (see discussion about proto-emergentist view of emergence in [50]). According to proto-emergentists, emergent properties are produced as *heteropathic* system effects. Such effects appear in complex biological and chemical systems where the conjoint actions of system components cannot be characterised as the sum of any individual causes. For example, the addition of sodium hydroxide to hydrochloric acid produces sodium chloride and water. However, it is unclear how such a reaction could be characterised as the sum of individual chemical compounds. In contrast, any properties occurring as *homeopathic* effects are not considered as emergent. Homeopathic effects arise when system causes acting together are identical to the sum of the effects of those causes acting in isolation. For example, forces acting on an object can have the same effect when applied in combination or separately.

Proto-emergentists proposed a layered view of complexity in which the world is divided into different *strata*. At the bottom there are the fundamental physical laws. On this foundation we can observe chemical, biological, psychological and social interactions at ever increasing levels of organisational complexity. In upper levels, we can identify properties that cannot be understood in terms of the individual observations of underlying physical phenomena. These emergent properties can be influenced by behaviours at lower levels in this layered approach. Research in physical sciences, therefore, investigates fundamental properties and laws that are broadly applicable. The remaining "special sciences" focus on properties that emerge from complex systems.

Despite having irreducibility as the common basis, proto-emergentist definitions of emergence have considerable differences as to how irreducibility can be modelled. For example, Alexander argues that emergent properties are novel qualities and associated, high-level, causal patterns which cannot be directly expressed in terms of more fundamental system entities and principles (see Johnson in [50] for a discussion). A common example of such a property is consciousness which is intrinsically a system level property and quite distinct from the underlying physiology of lower-level system components [26]. Along this line, it makes little sense to talk of human cognition in terms of individual neurons. However, this clear separation of emergent properties and their causal processes contrasts other proto-emergentist definitions. For example, Mill accepts that strong emergent properties, although irreducible to physical laws of causal composition, can still be described in terms of other laws or patterns of behaviour. Therefore, we can still talk about causes of patterns of cognitive behaviour despite that we cannot explain in detail how cognitive behaviours relate to underlying electrochemical changes in the brain. In this view, some relation between the emergent phenomena and the underlying system layers can be accrued by experimental analysis, for instance using simulations and/or statistics. For example, this approach has been extensively followed by 19th century bridge builders who used to develop experimental models of bridge behaviour when they did not have a sufficient theory to describe why some bridges failed while others did not.

Proto-emergentists do not consider interactions of the system with its environment, and they do not provide any suggestions about design methods or development techniques [50]. However, the latter view of emergence provides some hope for 'controlling' or at least anticipating emergent properties by aiming to understand the source of any non-determinism and the resulting emergent phenomena, for example by studying the underlying properties of lower-level components within a system. In this respect, functional decomposition can still be considered as a possible tool for engineering emergence. For example, one could attempt to develop multivariate statistical models linking the electrochemical brain changes with different of cognitive behaviours and then aim to achieve particular behaviours by applying specific electrochemical brain variations.

The definition of emergence has been largely a philosophical issue not only in early and proto-emergentist approaches but also in recent times as well. For example Bedau proposes three types of emergence, *nominal*, *weak* and *strong* [5, 6]. Nominal emergence refers to systemic properties that the system parts cannot individually have. In this approach, a system comprising points that are equidistant from a fixed point of origin is considered to have the nominally emergent property of "circle". The reason is that being a circle is a property which none of the individual points can have. Furthermore, strong emergence refers to nominally emergent properties which are supervenient to the properties of the individual system components. In addition, strong emergent properties exhibit irreducible causal powers affecting component behaviour, something which Bedau refers to by the term *downward causation*. Typical examples of strongly emergent properties are phenomenal qualities, such as qualia and consciousness. Bedau further claims that the irreducibility of strongly emergent properties to properties of system components renders strong emergence scientifically irrelevant. Finally, a nominally emergent property is considered as weakly emergent if it can be derived but only by simulation "… *a nominally emergent property P possessed by some locally reducible system S is weakly emergent if and only if P is derivable from all of S's micro-facts but only by simulation* …". In other words, weak emergence is generally reducible, although only with considerable difficulty. Therefore, Bedau argues that weak emergence can be quite useful for scientific study of numerous real-world phenomena.

In the majority of neo-emergentist approaches, emergence refers to the process that gives rise to phenomena which are novel and unpredictable. For example, Goldstein in [36] defines emergence as the "… *arising of novel and coherent structures, patterns and properties during the process of self-organisation in complex systems* …". Goldstein considers emergent phenomena to be conceptualised at a systemic, *macro-level* in contrast to the *micro-level* components and processes out of which they arise. Along the same line, Johnson in [50] defines emergent properties as "… *unexpected behaviours that stem from interaction between the components of an application and the environment* …". Johnson emphasises that emergent properties can be beneficial, for example, if users adapt products to support tasks that designers never intended. However, they can also be harmful, for instance if they undermine important safety requirements. According to this view, unpredictability of emergent phenomena is caused by system non-determinism, which can largely result from system *underspecification*, for example when designers fail to correctly specify the behaviour of non-deterministic application processes. However, non-determinism can also be introduced by other factors, such as environmental conditions. In a layered system view, for example along the lines of stratified system view suggested by proto-emergentists, non-determinism can also stem from interactions between adjacent layers.

Some definitions cover both irreducibility and unpredictability of emergent phenomena by considering different emergence types. For example, Chalmers argues for the existence of three types of emergence, termed *strong*, *weak* and *intermediate* [18, 19]. In this view, high-level phenomena are characterised as strongly emergent with respect to a low-level domain if they are not deducible, even in principle, from truths in that low-level domain. However, although strongly emergent phenomena are not deducible from truths from lower-level domain, they are still correlated to them. Regarding irreducibility, this notion of strong emergence is similar to that introduced by proto-emergentists; however, Chalmers further argues that the only emergent phenomenon that is truly strong is consciousness. To provide an example of consciousness as a strongly emergent phenomenon, Chalmers assumes a colour-blind scientist that is given complete physical knowledge about brains. He then emphasises that the scientist would nevertheless not be able to deduce and obtain a conscious experience of the red colour. Weakly emergent phenomena are also defined similarly to the neo-emergentist approaches, since they are considered as unexpected given the principles governing the low-level domain. Chalmers describes formations in cellular automata as representative examples of weakly emergent phenomena. Finally, intermediately emergent phenomena are not deducible from low-level laws and initial conditions, but they would be only from low-level facts. This type of emergence involves *downward causation*, meaning that higher-level phenomena are not only irreducible, but they also exert a causal efficacy to the low-level components. Typical examples of such phenomena are the otherwise inexplicable complex chemical compounds, which, once created, further constrain the behaviour of the participating chemical elements.

The majority of emergence definitions pays emphasis to axiomatically defining the physical nature of emergent phenomena. For example, Stefan in [73] argues that *weakly emergent* phenomena must have three fundamental properties: (a) they must be instantiated by systems consisting exclusively of physical entities (physical monism), (b) they must be global, namely they should not occur in individual system parts (systemic properties), and (c) they should depend nomologically on the microstructure of the system (synchronic determination), namely changes in the status of emergent phenomena should necessarily result from changes in the properties of the system parts or their environment. Stephan further defines *diachronic emergence* as having the additional properties of genuine novelty in system evolution and structure unpredictability, and *synchronic emergence* which further has the property of irreducibility. These three types of emergence are analogous to the previously described nominal, weak and strong emergence types introduced by Bedau.

There are also cases where the emergent phenomenon is not necessarily global. For example, Haan [38] argues that certain local properties can emerge in parallel with global systemic ones, a phenomenon which he terms *conjugate*. More specifically, Haan considers three types of emergence:

- Discovery Emergence. In this case the emergent phenomenon is observed only by an external observer and no conjugate occurs.
- Mechanistic Emergence. Here the global emergent phenomenon is again identified by an external observer, but there is also local emergence (conjugate). Furthermore, the dynamics producing the local emergence actually depends on the global emergent pattern, due to downward causal powers (downward causation).
- Reflective Emergence. In this case no external observer is required. The system objects have some reflective capacity, which enables them to observe the emer-

gent phenomena, both global and local, they produce. This case also involves both conjugate and downward causation.

Furthermore, there are definitions of emergent phenomena which have been given specifically for artificial systems. For example, Forrest [33] defines an *emergent computation* phenomenon as having three fundamental properties: (a) System behaviour should be determined from a micro-level dynamic process resulting from interactions of a collection of individual agents, (b) an epiphenomenon, such as a particular system state, should be produced by the aforementioned process at the macro level, and (c) the phenomenon should be a natural *interpretation* of the produced epiphenomenon either as computation process or computation results.

Similarly, Muller [64] considers a phenomenon as emergent iff:

1. It occurs in a system of entities in interaction whose expression of the states and dynamics is made in an ontology or theory D.
2. The system dynamics produce a epiphenomenon, such as a process, a stable state, or an invariant, which is necessarily global regarding the system of entities.
3. The emergent phenomenon itself is identified by interpretation of the aforementioned global epiphenomenon, either by an external observer or by the entities themselves, via an inscription mechanism in another ontology or theory D'.

Muller claims that the nonlinearity of component interactions guarantees the irreducibility of D' to D, and he considers two types of emergence: (a) strong Emergence that is observed by an internal observer, for example a social structure in social systems and (b) weak Emergence that is observed only by an external observer, for example the path formed by foraging ants.

Finally, computational emergent phenomena have been operationally defined in IRIT [17] as having two fundamental properties: (a) they must be realisations of specific system goals, such as solving specific problems or providing a particular, possibly evolving over time, functionality that must be confirmed by relevant system users, and (b) the details of the phenomenon, such as the solution given or the functionality produced, should not have been exactly pre-engineered in advance by system designers, but they should have been adaptively produced as a result of system dynamics.

3.3.2.3 The Essence of Emergent Phenomena

To be able to engineer emergence to achieve intended results, several issues need to be clearly understood. One such important issue is the scope of emergent phenomena. The consensus is that emergent phenomena are *global* and they concern the system as whole despite being dependent on the individual system components [5, 18, 36, 44, 64]. This is aligned with the common view that systems producing emergent phenomena can be perceived as involving two levels (see [44, 73] for instance). The system component (*micro*) level, where the mechanisms producing the emergent phenomenon are executed, and the systemic (*macro*) level, where the emergent phenomenon is perceived. For example, in a foraging ant colony, the individual ant movements belong to the micro level, while the path that is formed emergently between nest and food source is considered to reside at the macro level.

The question then is who can perceive the emergent phenomenon. Considering their global scope, emergent phenomena can generally be identified by some observer located outside the system that produces them [36, 38]. For example, the ant path can only be identified by those examining the foraging ant colony from some distance. Ants cannot perceive the whole path since they can only perceive their surrounding environment. However there are cases where emergent phenomena are identifiable from system components as well [18, 36, 64]. Examples of such cases are phenomena that emerge in social systems such as emergent leadership and emergent work ethics. In both cases the emergent phenomenon is identified by all society members.

Upon arising the emergent, phenomena subsequently affect the behaviour of system components. In most definitions the view is that component behaviour is affected by downward causal powers appearing due to the emergent phenomenon. This downward influence is commonly referred to as *downward causation* [6, 36, 44]. In other cases it is suggested that emergence can also take place at the micro level giving rise to emergent component properties which evolve in parallel with the global emergent phenomenon. An example of such dual emergence is the *conjugate* proposed by Haan [38].

Depending on whether the emergent phenomenon is visible by the system components or not, its downwards effects are direct or indirect respectively. For example, once an ant path or a flock of birds has been shaped, then movements of individual ants or birds are aligned to follow the path or to not collide with each other in the flock respectively. This is done indirectly since neither the birds not the ants are aware of the existence of the whole flock or path respectively. In contrast, once a leader has emerged in a society, as is the case of a leader emerging in music quartets for instance (see Goldstein [36] for more details), all society members become aware of the leader and receive direct influence.

The predictability of emergent phenomena is another issue of concern. In fact, in many definitions of emergence any unpredictable/unexpected phenomena are considered to be emergent [18, 36, 44, 50, 73]. In a first glance this seems to contradict the view that emergence could possibly be harnessed and used in artificial systems for specific purposes. The obvious question arising would be that if some outcome is indeed unpredictable, then how can we be certain that it will happen at all. Following this line of thought, we would conclude that we cannot hope to engineer emergence as the solution to a given problem. However, this problem is mitigated by the clarification that *unpredictability* of emergent phenomena actually means *unpredictability in practice*. In other words, emergent phenomena can be considered as completely unpredictable only the first time ever they are perceived since they could not be predicted in advance. After initial discovery of emergent phenomena however, models of cause and effect (both statistical and simulation ones) can be established between component behaviour and emergent outcomes, and hence emergent phenomena can become predictable in principle. In fact many emergent phenomena are discovered by simulations, and that approach is also widely applied to confirm that certain initial conditions, environmental dynamics and selected component behaviours will lead to specific emergent outcomes [5, 8, 28]. However, only probabilistic estimations of the type and time of occurrence of emergent phenomena can generally be made, and in cases where the environment has stochastic behaviour an emergent phenomenon cannot be exactly calculated in advance not even by simulation.

In particular, some authors consider emergent phenomena as *radically novel* (see for example [36, 73]); however, this cannot generally be the case. Emergent phenomena are in most cases only *practically novel* since, given the system dynamics, it is not easy to accurately predict them in advance. For example, it is not easy to exactly predict a hurricane and its characteristics before it is shaped. However, the occurrence of a hurricane is not something radically novel, it is simply practically impossible to exactly determine in advance the detailed characteristics of the particular hurricane that finally occurred. As Chalmers in [18] argues, a Laplace machine fed with all necessary data would in principle be able to accurately predict any phenomenon except consciousness.[10]

Another issue of concern is the in principle derivability or *reducibility* of emergent phenomena. In many definitions, irreducibility characterises strong forms of emergence, while unpredictability is sufficient to justify weak ones [5, 18, 50]. However, irreducibility in principle has been strongly criticised. For example, in many cases irreducibility is considered as lack of knowledge, which an appropriate theory explaining how the currently irreducible emergent phenomenon is produced would cover [36, 50, 73]. In particular, apart from offering a means to conceptually link cause and effects in emergent phenomena that otherwise would have no concrete explanation, irreducibility does not assist in our understanding of emergence. For example, Bedau considers strong emergence as being *scientifically irrelevant* for the study of natural phenomena [5, 6], while, as mentioned above, Chalmers accepts only one case of strong emergent phenomenon, that of consciousness [19].

Another important issue in understanding emergence is how emergent phenomena occur. In many cases emergent phenomena are considered to pre-exist, for example as in the Aristotelian view of emergence. In other cases there is no explicit reference to the emergence process, for example as is done in proto-emergentist definitions where the production process is seen as a black box (for example, see Johnson in [50] referencing Alexander). Similarly, there is no reference to the mechanisms producing the emergent phenomena in definitions where emergent phenomena are considered as in principle irreducible to the properties and behaviour of the system components (for example, see [19, 73]).

In the majority of cases, however, the emergent phenomenon is considered to be produced by some process which is executed at the component level. That process is necessarily *dynamic* [5, 36, 38, 44, 46]. As Goldstein emphasises, "… *emergent phenomena are not pre-given wholes but arise as a complex system evolves over time. As a dynamical construct, emergence is associated with the arising of new attractors in dynamical systems …*". The dynamism of the micro-level processes is

---

[10] Chalmers argues that the only irreducible emergent phenomenon is consciousness and any other phenomenon can be in principle derived given initial conditions and computing power. See [18, 19] for more details.

mandatory for two main reasons: Firstly, our reasoning is done on the basis of *physical monism*,[11] according to which the scientifically relevant emergent phenomena under examination are created only by physical powers and lows. Secondly, we require that emergent phenomena be practically (for example analytically) unpredictable in advance.

The mechanisms that produce the emergent phenomenon are often considered to be based on interactions of individual components. The reason is that autonomy of interactions and a large number of components result in nonlinear system behaviour, which cannot be accurately predicted in advance. Therefore, several authors agree that component interactions in complex systems operating at the edge of chaos (namely exhibiting far from equilibrium dynamics) can lead to emergent outcomes [36, 44, 53]. In fact, this nonlinearity requirement has led many authors to propose that decentralised control, large number of components, and autonomous interactions are mandatory system properties in order to exhibit emergent phenomena [27, 31, 36, 44].

Nonlinearity is not caused only by local interactions among distributed components however. For example, in dynamical systems we can have behaviours based on *iterated functions*,[12] which can lead to nonlinear behaviour and in particular to interesting emerging results [63]. The behaviour of such systems is largely unpredictable since it is governed by *deterministic chaos*. Among other characteristics, such systems can exhibit the butterfly effect [43, 45], namely their trajectory through state space is sensitively dependent on the initial conditions, and hence unobservably small causes can produce large effects. This fact also relaxes the requirement for the system to necessarily have a large number of components, and it actually can allow having emergent phenomena in systems with just one component. Stephan [73] provides a good relevant example of nonlinear behaviour produced by logistic functions. The example involves an iterated function defined on the logistic function $y = mx(1 - x)$. Stephan demonstrates that for different values of m, there is different behaviour of the logistic function, for example, for $m = 1,5$, the function converges to a constant value depending on initial value of x.

Some authors have linked the appearance of emergent phenomena with self-organisation [31, 36, 43]. In particular, in many cases emergence is thought to occur only in self-organising complex systems operating in far-from-equilibrium dynamics and being in the *edge of chaos* [52]. In complex systems, the research very often centers on the emergent global dynamics of a whole system. It is usual in this approach to view the global properties of the system as emerging from the actions of its parts, rather than seeing the actions of the parts as being imposed from a dominant central source [45, 46]. Properly defined, however, there may be instances of self-organisation without emergence and emergence without self-organisation, and it is clear from the literature that the phenomena are not the same (see [27] for instance). For example, crystallisation or the creation of sand dunes do not require any form

---

[11] See Stephan in [73] for more details on physical monism.

[12] The values of iterated functions are calculated sequentially in a manner such that the output of each calculation is the input to the next one.

of self-organisation. Both phenomena result from a dynamic process which eventually stabilises giving rise to the emergent outcome. Furthermore, dynamic behaviour does not occur only in complex systems. For example, the Game of Life cannot be characterised as a complex system, and yet it exhibits emergent phenomena. Therefore, the link between emergence and self-organisation generally remains an active research question [27, 31].

Further to the characterisation of emergent phenomena, there is a common view that they need to exhibit a level of adaptability and presistence, which implies that the system will have reached an organisational regime having some degree of stability and resistance to perturbations [20]. In fact, it is this very resistance to perturbations that for some authors differentiates true emergent phenomena from epistemological artefacts appearing due to pure chance. For example, Goldstein [36] discusses the case of some hill in USA which, when viewed from a certain angle and at a certain time of day, resembles the profile of the American president J.F. Kennedy. Obviously that is an epistemological artefact which does not really exist. Such epistemological artefacts, for instance shapes formed from play of light on leaves in a breeze, are considered to include *serendipitous novelty* and are clearly separated from authentically emergent phenomena (see Holland [46] for a discussion). The stable organisational regime linked with occurrence of emergent phenomena can be an equilibrium or an organisationally closed set of states, which, as mentioned in Sect. 3.3.1, are commonly referred to with the term *attractor* [46, 53]. In this respect, self-organisation is quite relevant since it can cause a system to reach an attractor regime and hence create the conditions for emergent phenomena to be created.

3.3.3 Operational Definition of Emergence in Computer Science

Considering the definitions discussed above, the term *emergence* can be used in software engineering to refer to a dynamic process producing a global phenomenon that is practically unpredictable in advance. Such an *emergent phenomenon* will be dependent on system component functionality, and its unpredictability will be due to nonlinear system behaviour. This view of emergence includes both reducible and irreducible types of emergence described in the existing definitions, since operationally all cases of emergence in software systems are reducible in this respect. Therefore, we provide the following definition:

> *Emergence* is the process that causes a software system to produce an *emergent phenomenon*.

The emergence process can most commonly, but not necessarily, be a process of self-organisation. The emergent phenomenon can be anything perceived from the system's stakeholders based on the system organisational status, which was not exactly predicted in advance. For example, an emergent phenomenon can be a particular system response, a piece of information that can consist of a solution to a given problem, a particular system state, or a particular process that the system may become capable of executing. An emergent phenomenon is generated and becomes identifiable when the system reaches some organisational regime, for example an *attractor*, namely a particular state or a set of states the system convergently settles to or iterates within.[13] Therefore, an emergent result can be understood as an interpretation of a particular attractor the system has converged into. Hence, we can define an emergent result as follows:

> An *emergent phenomenon* produced by a software system is an interpretation of an attractor the system has converged into, which is practically unpredictable given the functionality of system components.

An attractor represents for a process a model of causal closure built with system states [44, 46]. When reaching the attractor, the process "closes in" and can not reach out except when significant perturbations are applied to the system. Attractors can have many different shapes, sizes and dimensions. The simplest one is a zero-dimensional point attractor which consists of a single state. This describes the situation where a system reaches an equilibrium. Furthermore, a quite common attractor is an one-dimensional limit cycle, where all states of the attractor are revisited at regular intervals. This describes certain far-from-equilibrium configurations where the system exhibits periodical behaviour, such as the Benard rolls. Other types of attractors, which are termed "strange", are characterised by a non-integer, fractal dimension. This is a representative characteristic of certain types of chaotic processes [44].

An emergent result is generally perceived and identified as such by some *observer* who is external to the system producing it. There are cases, however, where the system components can also perceive the emergent result as a whole once it occurs. As mentioned previously, the emergent result always affects component behaviour and hence depends on whether system components can also perceive the emergent result that influence, termed *downward causation* (see Sect. 3.3.3), can be characterised as *direct* and *indirect* respectively. In the first case, system components will be directly informed of and affected by the emergent outcome, for example as is the case of a multi-agent system where an agent emerges as coordinator at some point and subsequently all other agents start contacting it to receive coordination commands. In the second case the behaviour of the system components will be affected without them being aware of the global emergent outcome, similarly to foraging ants moving along an emergent path without being aware of its existence.

Hence we can define two types of emergent phenomena:

> *Weak emergent phenomena* are those that are identified only by an external observer and have an indirect influence on system component behaviour.

> *Strong emergent phenomena* are those that are identified both by an external observer and system components and have a direct influence on system components behaviour.

Although not exactly predictable in advance, emergent phenomena are in general not completely unexpected; if they were, they would not be particularly useful in software engineering. Totally unexpected emergent results are usually linked

---

[13] Heylighen [43] uses the term *organisational closure* to refer to the convergence of a system to a set of particular states as a result of a dynamic self-organisation process.

with undesired system behaviour, one that was not intended when the system was designed but occurred indirectly during dynamic system operation. Therefore, in artificial systems emergent results should generally be expected with some probability. In fact, in the extreme case that the system components do not function stochastically, emergent results should be able to be completely regenerated given the same initial conditions and system operation rules. For example, in a software system simulating a foraging ant colony where ants have deterministic behaviour, the same ant path will re-emerge if we re-execute the system with the same initial conditions. Therefore, the challenge for software engineers is to build appropriate functionality into system components, so that, given appropriate initial conditions and environmental input, desired emergent results will be produced.

3.3.4 Properties of Emergent Phenomena

In the various definitions of emergence, several properties have been attributed to emergent phenomena. Considering the definition given in the previous section, an emergent phenomenon is characterised by the following properties:

- *Global coherence*. The phenomenon must have global scope, and it be must coherent and dependent on the functionality of the system components while at the same time being clearly separated from the properties of the constituent parts of the system. Furthermore, it should be understood as being perceived at a *macro* level while the underlying processes will be executed at a *micro* level. Therefore, there is a strong dependency between the dynamics observed at both macro and micro levels.
- *Novelty.* Novelty refers to the fact that although the resulting phenomenon is derived by interpretation of some particular system organisation, reached by a process involving micro-level system parts, it is radically different from the individual properties of these parts, and it cannot be directly estimated from them. In particular, to identify the emergent phenomenon, different concepts and theories from those used to describe the micro-level activities are generally required. For example, in a foraging ant system the ant movements are described in terms of pheromones, while the emergent ant path is described in terms of concentration of ants to particular locations.[14] Therefore, emergent phenomena are viewed as interpretations of particular system organisational arrangements.
- *Unpredictability.* The emergent phenomenon is not practically predictable before its occurrence. Although emergent results appearing in software systems are derivable in principle, for example if resource and time consuming procedures such as simulations are applied, in practice it is not obvious when they will occur. Therefore in practice emergent phenomena can be anticipated only probabilistically. In fact, given that the exact phenomenon is not known in advance, the

---

[14] See [33, 64] for a discussion on how emergent phenomena are interpretations of epiphenomena produced from dynamic system operations.

user can actually anticipate a *class of emergent phenomena*, a particular instantiation of which will occur with a given probability based on initial conditions, micro-level processes and environmental factors. Therefore from a requirements fulfilment perspective, such as problem solving, the system designer should attempt to engineer a generic class of phenomena that will potentially provide the problem solution or fulfil the particular requirements and then probabilistically expect, given past experience, that a member of this class will instantiate and provide the desired result. However, the exact occurrence of this particular instance is not known in advance because it depends on many factors ranging from the inherent system dynamism to stochastic environmental dynamics. For example, a system designer modelling a problem solving process as ant colony optimisation can anticipate that a solution will eventually emerge given the initial parameters and environmental dynamics based on the known behaviour of ant colony optimisation algorithms. However, she will not be able to predict the exact solution in advance; if that was the case, then the problem would be have been solved in the first place. It is similar to anticipating, based on previous knowledge, that some ant path will be formed between a food source and the ant nest. However, we will not be able to directly estimate the exact path in advance.

- *Dynamic* (*nonlinear*) *system behaviour.* The emergent phenomenon is an interpretation of an attractor the system has converged to. Hence, the system where an emergent phenomenon is identified needs to exhibit dynamic nonlinear behaviour. A chain of linear activities enables explainability and predictability of a collective phenomenon. On the opposite, an emergent phenomenon originates from nonlinear activities at the micro-level, and is realised and becomes identifiable when the system reaches an attractor organisational regime.

3.3.5 Engineering Nonlinearity

To be able to engineer software systems capable of producing emergent phenomena that fulfil particular requirements, we need to identify what necessary characteristics such systems must have. The source of emergent phenomena is nonlinearity, and hence it needs to be inherent in the behaviour of such systems. Subsequently, system designers need to engineer appropriate behaviours into system's components, whose interaction will eventually produce some instance of a class of desired emergent phenomena with a given probability. Such component behaviours would be typically known from existing cause and effect models to generate particular overall system behaviours and produce emergent phenomena.

There are different ways to engineer nonlinearity in artificial systems. As mentioned in Sect. 3.3.2.3, nonlinearity can be mathematically modelled in the behaviour of certain system components, for instance by using iterated functions or centralised rule-bases that result in deterministic chaos. However, this approach is not particularly strong in addressing the requirements of contemporary software. For example, centralised solutions suffer from the well-known bottleneck problem.

Furthermore, they are more cumbersome as far as it concerns adapting to changing environmental conditions.

Distributed nonlinear systems, on the other hand, are flexible and can adapt to the changing environment while approaching the solution to the problem. For example, an ant colony will converge to a path linking food source and nest despite any environmental perturbations, such as stones thrown on the forming ant path during foraging. Therefore the most beneficial sources of nonlinearity in this respect are complex interactions in distributed systems, such as the complex adaptive systems. Such systems are based on distributed components acting autonomously in a decentralised manner aiming to achieve own goals. Interactions are local and normally simplistic, but overall they result to global nonlinear system behaviour which can give rise to emergent phenomena.

The resulting dynamic behaviour of such systems is typically self-organising. Furthermore, the mechanisms dealing with component interactions can be classified in two types: *external* and *internal*. The external mechanisms enable modification of the system's behaviour to be initiated by its environment, for example by imposing constraints, enforcing rules and delegating artifacts. The internal mechanisms are ways to change the interaction dimensions that are unfolded by processes within the system. Interaction is essential to this framework because the events of novelty and innovation within a system arise from the interactions of these agents with each other and with the environment.

The challenge for system designers therefore is how to specify local component behaviours so that they will overall form a complex adaptive system and particularly that they will produce a desired emergent outcome. More details about methodologies for engineering self-organisation and emergence are provided in Chap. 12.

3.4 Summary

The concept of emergence has been studied since the Ancient Greece times, and it appeared in various domains such as philosophy, mathematics, physics, thermodynamics, systemics and complex systems. Its description has often been resumed to the phrase *"a whole that is more than its parts"*, but a lot of different definitions of emergence currently exist. For artificial systems, a sufficient definition considers emergent phenomena as interpretations of a particular system organisational regime, which are not practically predictable in advance.

Emergence refers to the occurrence of some outcome which, despite not explicitly represented at a lower organisation level, appears at a higher level and cannot be understood by simply observing individual component behaviours. Emergent properties are intimately linked with dynamic self-organising systems having decentralised control and local interactions. Their source is commonly, although not necessarily, the nonlinearity arising from the locality of component interactions.

Hence we accept that self-organising systems have the capability to alter their organisation (which results to changing the functionality they deliver) to adapt to external influences. What is organisation and what is functionality can be a matter of definition (since in the end they both change to react to external stimuli). However, in complex self-organising systems, where the possible organisations are too many to be exactly calculated or observed, and often stochastically determined, self-organising systems are still capable of producing them and function as to fulfil their purpose.

Instead of attempting to eliminate emergent phenomena, it could be interesting to explore how this might be deliberately achieved and harnessed. That is, elaborate on how to engineer artificial systems with desirable emergent properties.

Key Points

- Most self-organising software systems currently mimic natural systems to implement self-organising behaviour.

3.5 Problems–Exercises

3.1 Provide examples of cases where we need strong and weak self-organisation:

(a) in the physical world;
(b) in software systems.

3.2 Discuss the definition and principles of self-organisation as these were introduced for dynamic systems. Are the effects of self-organisation permanent?

3.3 Describe the requirements and implement a proof of concept demonstrator of a software system that exhibits the mandatory self-organising system properties introduced in Sect. 4.3.2.1.

3.4 Justify whether it is possible to have self-organisation without emergence and provide examples.

3.5 Describe the main emergentist schools and discuss the advantages and disadvantages of each approach.

3.6 Provide a definition of the term emergence, suitable for software systems, together with examples. Is it always necessary to have a decentralised system consisting of interacting components for emergent phenomena to appear?

3.7 Provide a definition and examples of the term “downward causation”.

3.8 Provide a definition together with examples and discuss the main properties of emergent phenomena.

3.9 Describe an example and implement a proof of concept demonstrator of a software system capable of exhibiting emergent behaviour.

3.6 Further Reading

The Vision of Autonomic Computing. A foundational introduction to Autonomic Computing. (J.O. Kephart and D.M. Chess, 2003, IEEE Computer, 36(1):41–50.)

Advances in Applied Self-organizing Systems. A collection of papers describing software applications using self-organisation and emergence. (M. Prokopenko, (ed.), Advanced Information and Knowledge Processing series, 2008, Springer, London.)

Self-organization. An instructive review article on history and present status of self-organisation mainly in physical and biological systems. (H. Haken, Scholarpedia, 2008, 8(1), http://www.scholarpedia.org/article/Self-organization.)

Special issue on Self-organization in Distributed Systems Engineering. A special issue comprising papers on self-organising mechanisms and applications. (F. Zambonelli and O.F. Rana (eds.), 2005, IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, 35(3).)

The Science of Self-organization and Adaptivity. A foundational review of the self-organisation and emergence concepts and models as they appear in various scientific areas. (F. Heylighen, 2001, Knowledge Management, Organizational Intelligence and Learning, and Complexity, in: The Encyclopedia of Life Support Systems, EOLSS, L.D. Kiel (ed.) 2001, Eolss Publishers, Oxford, pp. 253–280.)

Swarm Intelligence: *From Natural to Artificial Systems*. A comprehensive book on self-organising algorithms drawn from natural systems. (E. Bonabeau, M. Dorigo and G. Theraulaz, 1999, Oxford University Press.)

Self-organization in Biological Systems. A detailed presentation of self-organisation mechanisms in biological systems. (S. Camazine, J.-L. Deneubourg, N.R. Franks, J. Sneyd, G. Theraulaz and E. Bonabeau 2001, Princeton University Press.)

The many Facets of Natural Computing. An up-to-date review article on recent advances in natural computing. (L. Kari and G. Rosenberg, 2008, Communications of the ACM, 51(10):72–83.)

Self-managed Systems and Services. An overview of the uses of self-organisation in implemented systems and services. (J.P. Martin-Flatin, J. Sventek and K. Geihs, 2006, Guest Editorial of special issue, Communications of the ACM, 49(3):36–39.)

That special issue includes additional relevant papers as well.

References

1. Abdallah, S., Lesser, V.R.: Modeling task allocation using a decision theoretic model. In: Proceedings of Fourth International Joint Conference on Autonomous Agents and Multiagent Systems, Utrecht, Netherlands (2005)

2. Administrator of English Wikipedia: Diagram of convection cells. Creative Commons Attribution, vol. 2011 (2007). http://en.wikipedia.org/wiki/File:ConvectionCells.svg
3. Ashby, W.R.: Principles of the self-organizing dynamic system. J. Gen. Psychol. **37**, 125–128 (1947)
4. Bar-Yam, Y.: Dynamics of Complex Systems. Perseus Books, Cambridge (1997)
5. Bedau, M.A.: Weak emergence. In: Tomberlin, J. (ed.) Philosophical Perspectives: Mind, Causation and World, vol. 11, pp. 375–399. Blackwell, Malden (1997)
6. Bedau, M.A.: Downward causation and the autonomy of weak emergence. Principia **6**(1), 5–50 (2002)
7. Berlekamp, E.R., Conway, J.H., Guy, R.K.: Winning Ways for Your Mathematical Plays, vol. 2, 2nd edn. AK Peters, Wellesley (2001)
8. Bernon, C., Gleizes, M.P., Picard, G.: Enhancing self-organising emergent systems design with simulation. In: O'Hare, G.M., Ricci, A., O'Grady, M.J., Dikenelli, O. (eds.) Engineering Self-Organising Systems. Third International Workshop, ESOA 2005, Utrecht, The Netherlands, 25 July 2005. Lecture Notes in Artificial Intelligence, vol. 4457, pp. 284–299. Springer, Berlin (2007). Revised Selected Papers
9. Blitz, D.: Emergent Evolution: Qualitative Novelty and the Levels of Reality. Episteme, vol. 19. Springer, New York (1992)
10. Bonabeau, E., Dorigo, M., Théraulaz, G.: Swarm Intelligence: From Natural to Artificial Systems. Oxford University Press, London (1999)
11. Brennan, R.W.: Holonic and multi-agent systems in industry. Knowl. Eng. Rev. **16**(4), 375–381 (2001)
12. Briscoe, G., Dini, P.: Towards autopoietic computing. In: Colugnati, F.A.B., Lopes, L.C.R., Barretto, S.F.A. (eds.) Digital Ecosystems. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol. 67, pp. 199–212. Springer, Berlin (2010)
13. Brueckner, S., Parunak, H.V.D.: Self-organising Manet management. In: Serugendo, G., Karageorgos, A., Rana, O.F., Zambonelli, F. (eds.) Engineering Self-Organising Systems, Nature-Inspired Approaches to Software Engineering. Lecture Notes in Artificial Intelligence, vol. 2977, pp. 20–35. Springer, Berlin (2004)
14. Brueckner, S., di Marzo Serugendo, G., Karageorgos, A., Nagpal, R. (eds.): Engineering Self-Organising Systems. Lecture Notes in Artificial Intelligence, vol. 3464. Springer, Berlin (2005)
15. Brueckner, S., Robertson, P., Bellur, U. (eds.): Second IEEE International Conference on Self-Adaptive and Self-Organizing Systems (SASO 2008), Venice, Italy. IEEE Computer Society, Los Alamitos (2008)
16. Camazine, S., Deneubourg, J.L., Franks, N.R., Sneyd, J., Théraulaz, G., Bonabeau, E.: Self-Organization in Biological Systems. Princeton University Press, Princeton (2001). 2nd edn. (2003)
17. Capera, D., Georgé, J.P., Gleizes, M.P., Glize, P.: Emergence of organisations, emergence of functions. In: AISB'03 Symposium on Adaptive Agents and Multi Agent Systems (2003)
18. Chalmers, D.J.: Varieties of emergence. Tech. Rep., University of Arizona (2002)
19. Chalmers, D.J.: Strong and weak emergence. In: Davies, P., Clayton, P. (eds.) The Re-emergence of Emergence, pp. 244–256. Oxford University Press, Oxford (2006)
20. Churchland, P.: Matter and Consciousness. MIT Press, Cambridge (1984)
21. Collier, J.: Fundamental properties of self-organisation. In: Arshinov, V., Fuchs, C. (eds.) Emergence, Causality, Self-Organisation, pp. 150–166. NIA-Priroda, Moscow (2003)
22. Collier, J.: Self-organization, individuation and identity. Rev. Int. Philos. **59**, 151–172 (2004)
23. Contractor, N.S., Seibold, D.R.: Theoretical frameworks for the study of structuring processes in group decision support system—adaptive structuration theory and self-organising systems theory. Hum. Commun. Res. **19**(4), 528–563 (1993)
24. Corning, P.A.: The re-emergence of "emergence": a venerable concept in search of a theory. Complexity **7**(6), 18–30 (2002)
25. Craig, W.R.: Flocks, herds and schools: a distributed behavioral model. SIGGRAPH Comput. Graph. **21**(4), 25–34 (1987)

26. Cunninham, B.: Capturing qualia: higher-order concepts and connectionism. Philos. Psychol. **14**(1), 29–41 (2001)
27. De Wolf, T., Holvoet, T.: Emergence versus self-organisation: different concepts but promising when combined. In: Brueckner, S., Di Marzo Serugendo, G., Karageorgos, A., Nagpal, R. (eds.) Engineering Self-Organising Systems. Lecture Notes in Computer Science, vol. 3464, pp. 77–91. Springer, Berlin (2005)
28. De Wolf, T.D., Holvoet, T., Samaey, G.: Engineering self-organising emergent systems with simulation-based scientific analysis. In: Brueckner, S., Serugendo, G.D.M., Hales, D., Zambonelli, F. (eds.) Engineering Self-Organising Systems. Third International Workshop, ESOA 2005, Utrecht, The Netherlands, 25 July 2005. Lecture Notes in Artificial Intelligence, vol. 3910, pp. 138–152. Springer, Berlin (2006). Revised Selected Papers
29. Di Marzo Serugendo, G., Karageorgos, A., Rana, O.F., Zambonelli, F. (eds.): Engineering Self-Organising Systems: Nature-Inspired Approaches to Software Engineering. Lecture Notes in Computer Science, vol. 2977. Springer, Berlin (2004)
30. Di Marzo Serugendo, G., Gleizes, M.P., Karageorgos, A.: Self-organisation and emergence in MAS: an overview. Inform. Slov. **30**(1), 45–54 (2006)
31. Di Marzo Serugendo, G., Gleizes, M.P., Karageorgos, A.: Self-organisation in multi-agent systems. Knowl. Eng. Rev. **20**(2), 165–189 (2005)
32. Di Marzo Serugendo, G., Martin-Flatin, J.P., Jelasity, M., Zambonelli, F. (eds.): First International Conference on Self-Adaptive and Self-Organizing Systems (SASO 2007). IEEE Computer Society, Cambridge (2007)
33. Forrest, S.: Emergent computation: self-organizing, collective, and cooperative phenomena in natural and artificial computing network. In: Proceedings of the Ninth Annual CLNS Conference. MIT Press, Cambridge (1991)
34. Gershenson, C.: Self-organizing traffic lights. Complex Syst. **16**(1), 29–53 (2005)
35. Glansdorff, P., Prigogine, I.: Thermodynamic theory of structure, stability and fluctuations. Am. J. Phys. **41**, 147–148 (1973)
36. Goldstein, J.: Emergence as a construct: history and issues. Emergence **1**(1), 49–72 (1999)
37. Grassé, P.P.: La reconstruction du nid et les coordinations interindividuelles chezbellicositermes natalensis etcubitermes sp la théorie de la stigmergie: Essai d'interprétation du comportement des termites constructeurs. Insectes Soc. **6**(1), 41–80 (1959). doi:10.1007/BF02223791
38. Haan, J.D.: How emergence arises. Ecol. Complex. **3**, 293–301 (2007)
39. Hadeli, K., Valckenaers, P., Kollingbaum, M., Van Brussel, H.: Multi-agent coordination and control using stigmergy. Comput. Ind. **53**(1), 75–96 (2004)
40. Haken, H.: Information and Self-organization: A Macroscopic Approach to Complex Systems. Springer, Berlin (1988)
41. Haken, H.: Self-organization. Scholarpedia **3**(8) (2008)
42. Harrington, A.: Reenchanted Science: Holism in German Culture from Wilhelm II to Hitler. Princeton University Press, New Jersey (1999)
43. Heylighen, F.: The science of self-organization and adaptivity. In: Kiel, L.D. (ed.) Knowledge Management, Organizational Intelligence and Learning, and Complexity, The Encyclopedia of Life Support Systems, EOLSS, pp. 253–280. Eolss Publishers, Oxford (2001)
44. Heylighen, F.: Complexity and self-organization. In: Bates, M.J., Maack, M.N. (eds.) Encyclopedia of Library and Information Sciences. Taylor & Francis, London (2008)
45. Holland, J.H.: Hidden Order: How Adaptation Builds Complexity. Addison-Wesley, Longman, Redwood City (1995)
46. Holland, J.H.: Emergence: From Chaos to Order. Oxford University Press, Oxford (1998)
47. Holland, J.: Studying complex adaptive systems. J. Syst. Sci. Complex. **19**(1), 1–8 (2006). doi:10.1007/s11424-006-0001-z
48. Hopfield, J.J.: Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. **79**(8), 2554–2558 (1982)
49. Ishida, T., Gasser, L., Yokoo, M.: Organisation self-design in distributed production systems. IEEE Trans. Knowl. Data Eng. **4**(2), 123–134 (1992)
50. Johnson, C.W.: What are emergent properties and how do they affect the engineering of complex systems? Reliab. Eng. Syst. Saf. **91**(12), 1475–1481 (2006)

51. Jones, J.D.: Bifurcations. Accessed on 24-07-2011, http://www.mcasco.com/Order/bifurcat.html (2011)
52. Kaufman, S.: The Origin of Order: Self-Organization and Selection in Evolution. Oxford University Press, New York (1993)
53. Kaufman, S.: At Home in the Universe: The Search for the Laws of Self-Organization and Complexity. Oxford University Press, London (1995)
54. Kephart, J.O., Chess, D.M.: The vision of autonomic computing. Computer **36**(1), 41–50 (2003). doi:10.1109/MC.2003.1160055
55. Koestler, A.: The Ghost in the Machine. Hutchison and Co, London (1967), the Danube edn.
56. Langton, C.: Artificial Life: An Overview. MIT Press, Cambridge (1997)
57. Lewes, J.: Problems of Life and Mind, vol. 2. Kegan Paul, London (1875)
58. Malvile, E., Bourdon, F.: Task allocation: a group self design approach. In: International Conference on Multi-Agent Systems, pp. 166–173. IEEE Press, New York (1998)
59. Mamei, M., Zambonelli, F.: Programming stigmergic coordination with the TOTA middleware. In: Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS'05, The Netherlands, pp. 415–422. ACM, New York (2005)
60. Maturana, F., Norrie, D.H.: Multi-agent mediator architecture for distributed manufacturing. J. Intell. Manuf. **7**, 257–270 (1996)
61. Maturana, H.R., Varela, F.J.: Autopoiesis and Cognition: The Realization of the Living. Boston Studies in the Philosophy of Science. Springer, Berlin (1991). 2nd edn. (1980), 1st edn. (1973)
62. Maturana, H.R., Varela, F.J.: The Tree of Knowledge: The Biological Roots of Human Understanding. Shambhala, Halifax (1992)
63. Meiss, J.: Differential Dynamical Systems. SIAM, Philadelphia (2007)
64. Muller, J.P.: Emergence of collective behaviour and problem solving. In: Omicini, A., Peta, P., Pitt, J. (eds.) Engineering Societies in the Agents World 4th International Workshop, ESAW 2003, London, UK, 29–31 October 2003. Lecture Notes in Artificial Intelligence, vol. 3071, pp. 1–20. Springer, Berlin (2004)
65. Nicolis, G., Prigogine, I.: Self-Organization in Non-Equilibrium Systems. Wiley, New York (1977)
66. Paslack, F.: Urgeschichte der Selbstorganisation. Vieweg, Braunschweig (1991)
67. Picard, G., Bernon, C., Gleizes, M.P.: ETTO: emergent timetabling by cooperative self-organization. In: Engineering Self-Organizing Applications—Third International Workshop (ESOA) at the Fourth International Joint Conference on Autonomous Agents and Multi-Agents Systems (AAMAS'05), July 2005, Utrecht, Netherlands. Lecture Notes in Artificial Intelligence (LNAI), vol. 3910, pp. 31–45. Springer, Berlin (2005)
68. Polani, D.: Foundations and formalizations of self-organization. In: Prokopenko, M. (ed.) Advances in Applied Self-Organizing Systems, pp. 19–37. Springer, London (2008)
69. Prokopenko, M.: Design vs. self-organization. In: Prokopenko, M. (ed.) Advances in Applied Self-organizing Systems, pp. 3–17. Springer, London (2008)
70. Rumelhart, D.E., McClelland, J.L. (eds.): Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Foundations, vol. 1. MIT Press, Cambridge (1986)
71. Shehory, O., Kraus, S., Yadgar, O.: Emergent cooperative goal-satisfaction in large scale automated-agent systems. Artif. Intell. **110**(1), 1–55 (1999)
72. Stassner, J., Dobson, S., Fortes, J., Gowsami, K. (eds.): International Conference on Autonomic Computing (ICAC 2008). IEEE Press, Chicago (2008)
73. Stephan, A.: Emergentism, irreducibility, and downward causation. Grazer Philos. Stud. **65**(1), 77–93 (2002)
74. Sternberg, R.J., Mio, J.S.: Cognitive Psychology. Wadsworth, Belmont (2005)
75. Varela, F.J.: Principles of Biological Autonomy. Elsevier, New York (1979)
76. Vrba, P., Tichy, P., Mařik, V., Hall, K.H., Staron, R., Maturana, F.P., Kadera, P.: Rockwell automation's holonic and multiagent control systems compendium. IEEE Trans. Syst. Man Cybern. Part C, Appl. Rev. **41**(1), 14–30 (2011)
77. Yousif, M., Rana, O.F., Fortes, J., Goswami, K. (eds.): International Conference on Autonomic Computing (ICAC 2007), Jacksonville, Florida, USA. IEEE Press, New York (2007)

78. Zambonelli, F., Parunak, H.V.D.: Signs of a revolution in computer science and software engineering. In: Petta, P., Tolksdorf, R., Zambonelli, F. (eds.) Engineering Societies in the Agents World 3rd International Workshop, ESAW 2002, Madrid, Spain, September 2002. Lecture Notes in Computer Science, vol. 2577, pp. 120–125. Springer, Berlin (2003)
79. Zambonelli, F., Gleizes, M.P., Mamei, M., Tolksdorf, R.: Spray computers: frontiers of self-organization for pervasive computing. In: Second International Workshop on Theory and Practice of Open Computational Systems (TAPOCS 2004) in 13th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE'04), Modena, Italy, pp. 397–402. IEEE Press, New York (2004)