CASA

WORKING PAPER SERIES

76
03 | 04

http://www.casa.ucl.ac.uk/working_papers/paper76.pdf

 ISSN: 1467-1298

# Design and anticipation: towards an organisational view of design systems

Theodore Zamenopoulos
Katerina Alexiou

www.casa.ucl.ac.uk

T +44 (0) 20 7679 1782 • T +44 (0) 20 7679 1813 • F +44 (0) 20 7813 2843 • E casa@ucl.ac.uk

Centre for Advanced Spatial Analysis • University College London • 1 - 19 Torrington Place • Gower St • London • WC1E 7HB

UCL

# Design and anticipation: towards an organisational view of design systems

Theodore Zamenopoulos and Katerina Alexiou

Centre for Advanced Spatial Analysis
1-19 Torrington Place
London WC1E 7HB
t.zamenopoulos@ucl.ac.uk, a.alexiou@ucl.ac.uk

## 1. Introduction

Although the first attempts to establish a systematic way of studying, teaching and practicing design go back to the 1960s (Asimov 1962; Alexander 1963; Archer 1965), we are today just as far from agreeing on a definition of design science, let alone a definition of design itself. And although we will be discussing here the concept of anticipation and its 'repercussions' for the understanding and development of design systems, we will find that this paper is about anticipation as it is about scientific explanation. This is because the concept of anticipation couples with a special kind of scientific enquiry and explanation: an approach to science which takes the view that complex realities (or systems) can be better understood by studying their organisational principles, rather than building descriptions of their structural components. Research in various fields such as systems theory, cybernetics, information theory, artificial intelligence, physics, biology, and complexity science, gave birth to concepts (such as entropy, complexity and requisite variety) that marked this transition from observations of the state space of a system, to the examination of the system itself in terms of its functional components and their relations (Rashevsky 1954; Thóm 1972; Rosen 1991). In parallel, the concept of anticipation has also emerged as an alternative type of causality or reasoning which seeks to replace the reactive, cause-and-effect, Newtonian universe with a proactive one, according to which anticipated future states of a system can affect its current states (final cause explanations) (Rosen 1985; Dubois and Resconi 1992; Glaserfeld 1998).

It is asserted here, that if transferred to the context of design, the idea of anticipation necessitates studying and explaining design abilities on the basis of the organisational conditions and characteristics that underlie design systems. Organisational explanations are a potentially attractive direction of research because they allow the definition and explanation of essential design phenomena (or abilities) irrespective of the specific structures we can use to realise them.

In the following will we first review some predominant approaches to anticipation, both in natural and artificial systems, and subsequently discuss the concept of anticipation in the domain of design with the intention to develop a foundation for a formal definition of design systems.

## 2. Approaches to anticipation in various fields

Robert Rosen may well be considered as the 'father' of anticipatory systems (1985), even though hints and discussions on the concept of anticipation can be found long before that, albeit primarily in the context of philosophy, psychology and cognitive science (e.g. James 1890; Tolman 1932; Skinner 1938; Piaget 1954; Kelly 1955; Neisser 1975). However, Rosen was a biologist - or better a relational biologist - who strove to study, understand and explain the very nature of life using the apparatus of mathematics (Rosen 1991; 1999), and who in the process also submitted a rigorous re-assessment of physics and science in general. His concept of anticipation is tied with the reinstatement of the 'lost' cause of Aristotelian logic: the *final cause*. Briefly, the old Aristotelian categories of causation (called material cause, formal cause, efficient cause and final cause), constituted four distinct ways of answering 'why-questions': the first related to matter, the second to form, the third to producer and the fourth to end. According to Rosen, explanations of the fourth kind have been excluded from science because they have been taken as a direct violation of the traditional notion of causality, by which 'in any law governing a natural system, it is forbidden to allow present change of state to depend upon future states' (Rosen 1985: 9). Yet, he suggested that finality was the only kind of explanation that could be offered to anticipatory behaviour (the ability to foresee or predict some future change and accordingly adapt the present state or course of action), which is manifested at all levels of biological organisation, from the molecular level up to the human level. For him, the idea of building and employing predictive models was also a fundamental aspect of science in general, perceived as a means to construct homologies between modes of social and biological organisation and sustain a general theory of planning, policy generation and decision-making (ibid: 6-7).

Rosen's characterization of an anticipatory system is built on the dual relationship between a dynamical system *S* (running in real time) and a model *M* of *S* – another dynamical system – that can 'go faster' that real time and therefore predict future behaviour. The classic quote from Rosen's writings specifies that an anticipatory system is one 'containing a predictive model of itself and/or of its environment, which allows it to change state at an instant in accord with the model's predictions pertaining to a later instant' (Rosen 1985: 339). Such a system is diagrammatically illustrated in Figure 1, where *E* is an effector via which *M* and *S* interact with each other. Through *E*, input information from *M* is converted to some specific modifications of the dynamics of *S*; but *E* is also the vehicle through which the states of *M* are updated to match future states of *S*. Note that this sort of reinforcement and utilisation of information about future states, akin to processes of adaptation and learning, is in fact crucial for the generation and guidance of the predictive model (see Rosen 1985: 385-390). Rosen noted that the predictions of the model *M* could not in general be perfect and that the discrepancy between actual (system) and predicted (model) behaviour, and its growth, corresponds to the discrepancy between an open and a closed system. For that reason concluded that a system like this should be named quasi-anticipatory.

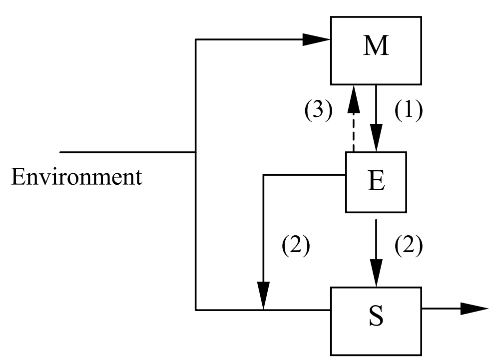



Figure 1 An anticipatory system as illustrated by Rosen (1985)

In fact Rosen's approach is developed in the context of the more generic paradigm of 'modelling relation', the establishment of relations between a natural system (an aspect, member, or element of the external world we wish to study) and a formal system (a system we create to represent, model and draw inferences about the natural system). The endeavour of modelling relation refers to the consistent encoding of a natural system into a formal one, so that the inferences developed within the formal system become predictions about the first. Consistency relates to the verifiability of the predictions in the natural world when these are decoded into relations in that world (Figure 2).

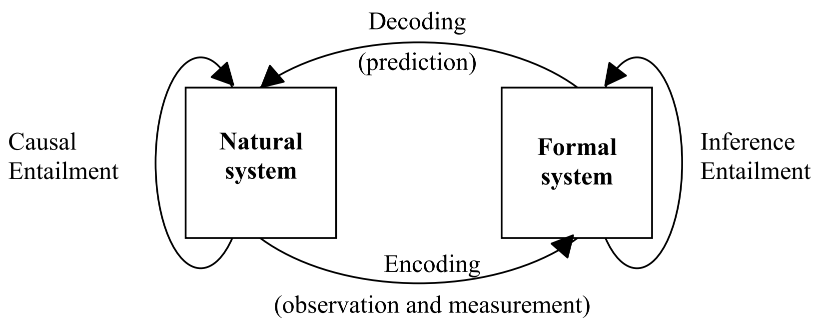



Figure 2 Modelling relation

Within this framework, we can now see *M*, *S* and *E* in Figure 1 as parts of the same system $S_2$ which we consider to be anticipatory. The environment is represented by another system $S_1$. Rosen (1985: 344) arrives at the formulation of five conditions summarised as follows: 'An anticipatory system $S_2$ is one which contains a model of a system $S_1$ with which it interacts. This model is a predictive model; its *present* states provide information about *future* states of $S_1$. Further, the present state of the model causes a change of state in other subsystems of $S_2$; these subsystems are (a) involved in the interaction of $S_2$ with $S_1$, and (b) they do not affect (i.e. are unlinked to) the model of $S_1$. In general, we can regard the change of state in $S_2$ arising from the model as an *adaptation*, or pre-adaptation, of $S_2$ relative to its interaction with $S_1$' (italics in original).

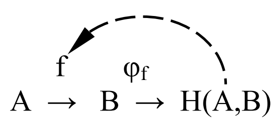
A pertinent example that epitomises such an anticipatory ability is a class of relational cell models called metabolism-repair (M, R)-systems. These systems are characterised by two fundamental biological qualities or components: a metabolic component, which can be represented as a set of mappings that convert inputs from the environment to outputs, and a repair component, which maintains and reconstitutes the metabolic activity. More formally, adopting Rosen's notation (1985), if $f\colon A \rightarrow B$, is a metabolic element in a metabolic system, then $f$ belongs to a larger set of physically realisable metabolisms the cell can display, denoted by H(A,B). Any process that can generate copies of $f$ must have its range in H(A,B): if $\varphi_f$ is the repair element, then this will be a mapping into this set. Hence the simplest (M, R)-system is given by:

$$\underbrace{A \xrightarrow{f} B}_{\text{metabolism}} \underbrace{\xrightarrow{\varphi_f} H(A,B)}_{\text{repair}} \tag{1}$$

Rosen showed that this formalism already contains what is needed for the operation of another decisive genetic component: a replication mechanism. This reproductive mechanism can be represented by a mapping $\beta_b$: H(A,B) → H(B, H(A,B)) (under the condition that there is an element $b$ in $B$ for which $\hat{b}^{-1} \equiv \beta$). An abstract (M, R)-system can therefore be expressed by the following diagram:

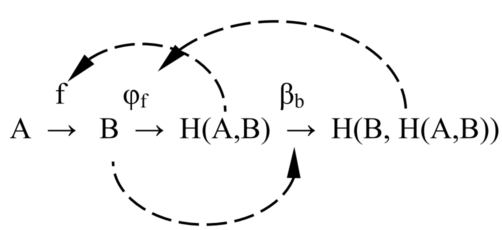
$$\underbrace{A \xrightarrow{f} B}_{\text{metabolism}} \; \underbrace{\xrightarrow{\varphi_f} H(A,B)}_{\text{repair}} \; \underbrace{\xrightarrow{\beta_b} H(B, H(A,B))}_{\text{replication}} \qquad (2)$$

This is the quintessence of a relational model in that the property of replication is entailed by the metabolism and replication mappings (or functions); it is constructed on the basis of these mappings alone, from the organisation of the system, and independently from any particular realisation or structure of the living cell. Let us briefly uncover the explanatory attributes of this diagram. For the simple mapping f: A→ B, the question 'why *B*' can only have two answers: in terms of material cause ('because *a*' - the value of *A*) and in terms of efficient cause ('because *f*'). But *B* has no explanation in terms of final cause – there is nothing to explain *B* by its effects in this diagram, or to endow it with a function (any finalistic answers have to pertain to the external environment of the system). On the contrary, *f* and *a* can only be explained in finalistic terms: because of their function in the diagram, that is to entail *B*. By extending the original mapping with the repair component the function *f* is efficiently entailed by $\varphi_f$:




(3)

Furthermore, by adding the replication mechanism the mapping $\varphi_f$ can also be entailed; but it is the fact that *B* can entail $\beta_b$ that eventually does the trick. In this last diagram everything is efficiently entailed and all final cause answers are found within the system. Note that only the initial input *A* originates from the environment.




(4)

This for Rosen also exemplified the essence of an organism, a system closed to efficient causation, which is characteristically non computable (contains non-simulable models). Nonetheless, Rosen's examination of anticipatory systems with its associated premises has not only been an inspiration for the study of biological worlds, but also for the scientific enquiry, understanding, modelling and control of complex systems in general. For example, Casti (1992; 2002) discusses Rosen's idea of anticipation in the context of control, highlighting that the crucial quality underlining an anticipatory control process, in contrast to classical (or reactive) control processes, 'is the notion of *self-reference*' (Casti 1992: 167). The idea of a system containing an internal model of itself is linked to the ability of a system to control not only its operation, but also its fabrication/structure and the way it evolves and adapts. In the above (M, R) formalism, the boundary conditions of the repair and replication mappings, serve as an internal self-model of the cell, allowing it to exert control over its own behaviour (Casti 2002: 12).

In the domain of intelligent systems, Tsoukalas with his colleagues has elaborated fuzzy logic and neural computing concepts and techniques for the development of anticipatory control systems (Xinqing et al 1996; Tsoukalas 1998). This approach to anticipation considers systems that can determine current action (or take decisions) on the basis of current as well as anticipated states. On the core of these systems is a predictive neural network model working in synergy with a fuzzy rule base used to determine the 'one-step-ahead' predictive control input. The overall architecture of such a control system is given in Figure 3, where $u(k)$ is the control input, $y_n(k+1)$ is the predicted neural output, $y_d(k+1)$ is the desired output and $y_p(k)$ is the plant output.

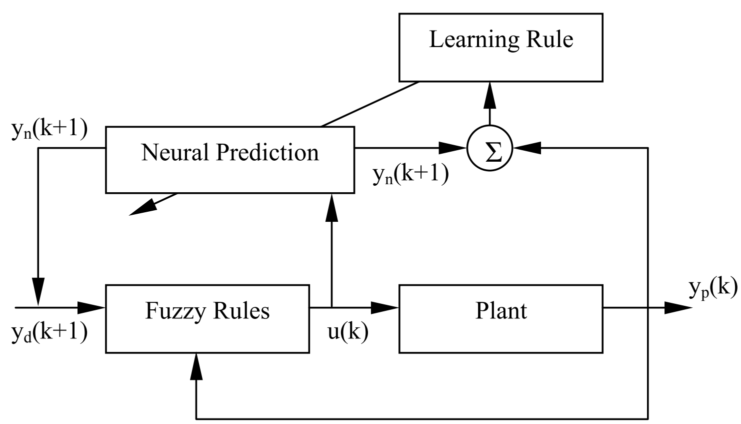



Figure 3 The structure of a neurofuzzy anticipatory control system as illustrated by Tsoukalas (1998)

It is worth mentioning here that Rosen's own view of the main difference between classical control and this kind of anticipatory control lies in the predictive character of the implicit model; while classical control rests on feedback regulation through the use of an error signal, anticipatory control rests on feedforward regulation through the use of model prediction (pre-adaptation) (Rosen 1985: 42). This has implications on the way final causation is expressed. In cybernetics, finality is connected to an externally defined purpose or a desired final state that the controller tries to reach or maintain. On the contrary, Rosen's idea considers finality without commitment to some ultimate goal - the existence of a goal is instead relative to the model (and the model alone) and hence is to a degree conjectural. For an overview and discussion on purpose, finality and teleological explanations see Boden (1972), George and Johnson (1985) and Stout (1996).

Another prominent approach to research in anticipatory systems is advocated by Dubois (Dubois 1997; Dubois 1998; Dubois 2000). Starting from a divergent position to that of Rosen, he suggests that anticipation is not a characteristic of biological systems alone (a trait of life), but is fundamentally present in all physical systems. In particular, he asserts that Rosen's notion constitutes a special (weak) form of anticipation as it is founded on model-based prediction, and discusses a formulation where anticipation as change of current state according to initial as well as final conditions is achieved at a system level. His alternative interpretation is based on the concepts of incursion (implicit recursion) and hyperincursion (incursion with multiple solutions).

More specifically, Dubois describes Rosen's anticipatory system *S* as a set of differential equations as follows (*M* denotes the predictive model):

$$\Delta S/\Delta t = [S(t+\Delta t) - S(t)]/\Delta t = F[S(t), M(t+\Delta t)] \quad (5a)$$

$$\Delta M/\Delta t = [M(t+\Delta t) - M(t)]/\Delta t = G[M(t)] \quad (5b)$$

His alternative proposition, where the future state of the system *S* <u>and</u> the model *M* at time $t+\Delta t$ is a function *F* of this system *S* at time *t* and of the model *M* at a later time step $t+\Delta t$ (Dubois 1998: 5), is written as follows:

$$\Delta S/\Delta t = [S(t+\Delta t) - S(t)]/\Delta t = F[S(t), M(t+\Delta t)] \quad (6a)$$

$$\Delta M/\Delta t = [M(t+\Delta t) - M(t)]/\Delta t = F[S(t), M(t+\Delta t)] \quad (6b)$$

The above system exemplifies the concept of incursion, where future state is computed in a self-referential manner. Dubois uses the concepts of incursion and hyperincursion as a method to investigate and develop a series of formal models, ranging from control of feedback and chaotic systems, to the generation of fractals from incursive automata and digital wave equations. The driving force for this programme seems to be the development of a new computational paradigm or, more accurately, a new abstract computational machine (a generalised Turing machine) that takes into account final causation.

As a matter of fact, Dubois's view summarises the main discussion points in research relevant to anticipatory systems: whether anticipation is a unique characteristic of biological systems or extends to all complex systems (biological, natural, and artificial), and whether anticipation can be realised computationally. In between Rosen and Dubois, there exists a multitude of views about anticipation and lively arguments about the conditions that enable anticipatory behaviour to arise (e.g. predictive power, entailment, the existence of an internal model, memory, language, learning, autonomy, adaptability etc) and the structures that can implement it. These questions are increasingly being debated in yet another field, that of adaptive systems, autonomous agents, and multi-agent systems.

Butz et al (2003) offer a classification of anticipatory mechanisms used in the context of artificial adaptive learning systems. In particular they distinguish four classes of mechanisms: implicitly anticipatory mechanisms in which no prediction is made but anticipations are built-in in the behavioural structure or algorithm of the adaptive system; payoff anticipatory mechanisms in which predictions are restricted to possible payoffs of future actions; sensorial anticipatory mechanisms in which a predictive model of the environment is used to predict input stimuli and influence sensorial processing; and state anticipatory mechanisms in which an explicit predictive model forms predictions about future states in order to directly influence current action and decision making. Inevitably the question about the applicability of an internal predictive model is of central importance here, as it also related to the debate about the suitability of reactive versus cognitive/deliberative approaches in artificial intelligence. Most researchers however are increasingly inclined to develop hybrid approaches that combine reactive behaviour with higher-level learning and reasoning abilities.

But let us have a more detailed look at some indicative approaches to anticipation in the field of autonomous agents. For example Butz, Stolzmann and colleagues (e.g. Butz et al 2002; Gérard et al 2002) have developed a class of mechanisms called anticipatory learning classifier systems. These systems are informed by the psychological learning theory of anticipatory behavioural control (Hoffmann 1993) and contain an explicit prediction component. The predictive model consists of a population of classifiers, each representing a condition-action-anticipation rule. The learning process starts with most general rules and gradually generates more specialised classifiers thus building a complete internal representation of the environment. The learning process compares the anticipation produced by each classifier in an action set (i.e. the prediction produced about the effects of an action) with the real next situation, and either modifies the classifier, or generates a new one. In this framework both generalisation and direct specialisation mechanisms are used.

Another approach is exemplified in the framework discussed by Ekdahl and his colleagues (Ekdahl et al 1995; Ekdahl 2000) who see anticipation as a characteristic of the situations in which agency occurs. They take a linguistic approach to anticipatory agents and propose a distinction between descriptive and model-based anticipation; while description-based anticipatory agents are based decisively on Rosen's model, model-based agents also have a meta-level component, which allows them to interpret the predictive model (see Figure 4).

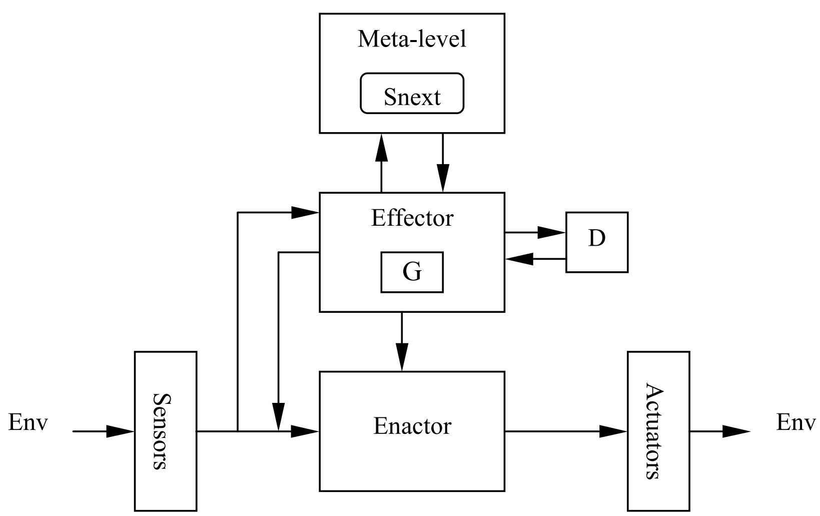



Figure 4 The basic architecture of an anticipatory agent by Ekdahl et al (1995)

The authors recognise that such an approach may lead to infinite regress as Rosen had pointed out, but they suggest that the problem can be disregarded because of time restrictions that exist in situations involving agents. Nevertheless, they see this hybrid architecture as a promising framework for examining issues of self-reference and autonomy in general, while being cautious not to neglect linguistic restrictions posed for the realisation of anticipatory agents in computer systems.

From a more philosophical standpoint, Christensen and Hooker (2000) discuss their view of agency as something deeply rooted in the organisational characteristics of life, and of which autonomy, adaptiveness and anticipation are three integral aspects. According to this view, intelligence lies in the organisational conditions of complex systems and is taken to emerge through increasing self-directedness. The notion of self-directedness refers to the ability of a system (agent) to adapt on the basis of formed anticipations and shape system-environment interaction in ways that are beneficial for maintaining autonomy. This is a constructivist approach to agency, in that it highlights the importance for agents to be able to interact with the environment, evaluate this interaction and modulate their future behaviour and action, including changes in the interacting parts of the environment. Reflecting on Rosen's original formulation of the conditions for anticipatory systems (see above) they suggest that their proposed framework may provide an additional account for aspects of autonomy such as the formation and pursuit of goals, as well as the use of feedback to guide adaptation and learning.

Christensen and Hooker are in fact part of a constructivist/dynamic systems 'school' within artificial intelligence that emphasises the importance of taking into account the dynamics of system-environment interaction (Beer 1995; Pfeifer 1995; Smithers 1995; Prem 1997; Edmonds 1999; Quick et al 2000). These approaches are tightly linked with the study and development of embodied agents. For example, Pfeifer (1995) is one of the few who looks particularly at anticipatory behaviour, in a robot developed based on his 'distributed adaptive control' architecture. Pfeifer concludes that such behaviour can only be seen as a product of the interplay between computational mechanisms, environmental features, and physical attributes of the robot (Pfeifer 1995: 56). He additionally suggests that a sufficient account of this behaviour would have to include factors that go beyond the information processing paradigm. Another example is the work of Quick et al (2000) who go on to devise a definition of embodiment based on the idea of mutual perturbation between system and environment, such that can enable structural coupling between the two. The notion of structural coupling is taken from Maturana and Varela (1980) and stands in line with the view that behaviour can be observed in the interplay between system and environment. Furthermore, this definition suggests that higher-level cognitive capabilities (i.e. such as intentionality, learning and anticipation) could also be explained on the basis of the relationship between agent and environment, the sensorimotor repertoire of the agent and the properties of its structural coupling with the environment (Dautenhahn et al 2002: 408).

Riegler (2001; 2003), another 'constructivist', involved in the investigation of the role of anticipation in cognition, agrees with Dubois in regarding Rosen's approach as a weak form of anticipation. He alternatively suggests that anticipation, and the process of creating an individual world-view as a means for understanding and acting in the world, is indeed a fundamentally constructive process - but one that characteristically does not lie beneath conscious mechanisms. Drawing from evidence related to 'stubbornness' in reasoning in humans, or to kinds of inborn and emotional (instinct-driven) anticipation in animals, he reflects the internal model hypothesis, proposing instead that anticipation is caused by 'canalization'. According to Riegler, canalization is a consequence of system organisation, and denotes constraints imposed on individuals from the way various elements of experience are interlocked in a constructed hierarchy of schemata, which act as internal hypotheses to guide future action. In contrast to the anticipatory classifiers discussed above, these schemata do not have a component responsible for formulating expectations but are only composed by a condition-sequence of actions duplet.

Additionally to the concepts of intelligence, autonomy, adaptability or embodiment in single agents, anticipation has also been studied in relation to concepts of coordination and cooperation in groups of agents. For example Davidsson (1997) presents a framework for linearly anticipatory agents based directly on Rosen's formulation. Each agent consists of a reactive component, a world model and an anticipatory component. The anticipatory component uses the world model to make predictions and guide agent behaviour by adapting the reactive component, which carries out a rotating sequence of perception (including other agents), action selection and execution. The experiments describe the use of this architecture to guide learning in simple competitive and cooperative multi-agent tasks. Another example where reactivity and anticipation are combined in a multi-agent setting is described by El hadouaj et al (2000). They discuss a multi-agent simulation of road traffic where the ability to look ahead in space is linked to the ability to look ahead in time and forecast future conflicts. The most 'famous' example however, has to be the work of Veloso et al (1999) who used multi-agent anticipation in their winning team of agents competing in the RoboCup-98 tournament. Anticipation is seen as a means to achieve better social utility and maximise the probability of future collaboration among agents. This is achieved with the combination of two key elements: a decision theoretic selection of the active agent using a single-step look ahead algorithm, and a dynamic strategic positioning of the passive agents who take into account both the location of the other agents and the attacking goal to anticipate their future contributions to the team. It is interesting to note here that the connection between expectation formation and concepts of cooperation, conflict, and coordination, goes back to the theory of games first introduced by John Von Neuman and Oskar Morgensten in 1944. Game theory is generally concerned with modelling and predicting rational decision making in groups of individual players who seek to employ appropriate strategies in order to satisfy their goals and achieve preferred outcomes. Preferences are typically captured by a measure of expected utility, and players are assumed to act so as to maximise this utility through a process of assessing the future choices of others. The classical framework generally suggests that agents act perfectly rationally, they have consistent preferences, and complete access to information about future actions/choices. The recognition of the fact that information cannot always be readily and fully available to agents, as well as that agents do not always behave consistently, nor do they possess infinite information processing capability, led to the formalisation of the bounded rationality hypothesis (Simon 1957). The idea of bounded rationality has a significant effect on the way expectation formation is approached: it implies that the process is not only affected by 'objectively' defined external factors, but also by factors relative to the way the agents perceive their environment and themselves acting within it. This is closer to the view of anticipation as involving the construction of an internal model. An interesting approach to expectation formation given the effects of bounded rationality is expressed by Arthur (1994) who links expectation formation to hypothesis generation and evaluation by considering decision making as an inductive reasoning process that involves forming temporal expectations.

Let us now summarise the different approaches to anticipation[1] and draw attention to some key issues, especially with reference to Rosen's paradigm. In general terms, we can discern from the examples above two major classes of approaches: one that considers anticipation as a sophisticated cognitive capability which needs to be studied at a high representational and/or semantic level; and one that considers anticipation as an intrinsic characteristic which underlines complex behaviour in all kinds of dynamic systems, natural and artificial. Rosen considered anticipation as a characteristic attribute of biological systems and stressed the existence of a parallel between natural languages and organisms, in the sense that they both possess semantic modes of entailment that cannot be encapsulated in a syntactic formalisation, a formal system, or a machine (Rosen 1991: 247). His premise was that traditional reactive approaches were inadequate to deal with anticipation; but this has often led to a common perception of the internal predictive model (what is an indispensable condition for anticipation in Rosen's view) as a necessarily cognitive one. This is somewhat paradoxical if we consider the primary examples Rosen chose to illustrate his point: the M-R model of a living cell (as discussed above), or the model of a biosynthetic pathway with a forward activation step (see Rosen 1985: 349-354). Rosen's point was instead to illustrate that anticipation is a corollary of complexity, which pertains to the modelling enterprise itself and is crucially linked to the existence of an observer. In this fashion he associated complexity to the concept of error and related the appearance of bifurcations with our ability to produce enough independent encodings so as to fully describe a given natural system (Rosen 1985: 83). Thus, his concept of anticipation served a dual purpose: first, to formulate a relational definition of organisms as a general class of systems superseding traditional views of biological systems as mechanisms, and second, to devise an alternative approach to complexity that captures our intuitive understanding of it as something related to the appearance of bifurcations and the emergence of unexpected behaviours.

Although the question as to whether an anticipatory system could in principle be realised computationally is still open, it is nevertheless true that both dynamic systems perspectives and artificial intelligence have been offering a fertile ground for investigation of anticipation and related notions. Amongst the various arguments that seek to establish links between anticipation and concepts such as complexity, autonomy, adaptability, embodiment or coordination, there is one that seems particularly worthwhile our investigation. This corresponds to the ever-sustained duality between system and environment. Rosen had detected this duality in the machine metaphor of biological systems and the restrictions that posed to the understanding of final causation, and so strove to dissolve it by offering a formalism of organisms as systems whose fabrication process is entailed within them. However, as we have seen, many others have highlighted the importance of taking into account the mutual affects of system-environment interaction in the scientific investigation of agency in natural and artificial worlds. For example, Maturana and Varela's (1980) idea of structural coupling, the focus on establishing mutual perturbatory channels between robots and their environment (Quick et al 2000, Dautenhahn et al 2002) and the more explicit request for including interacting aspects of the environment in the process of achieving and sustaining organisational closure (Christensen and Hooker 2000), all disclose a need for re-evaluating the relation between internal and external, observation and action, or interaction and creation.

[1] Although this section is aimed to provide an overview of the main approaches to anticipation, it is not exhaustive. For a full view of the work in the field see the Computing Anticipatory Systems (CASYS) Conference Proceedings, that appear annually since 1997 and are edited by Dubois, as well as the book edited by Butz et al (2003).

Obviously anticipation seems to be an important aspect for understanding, modelling or even constructing the world (any world), and is also strongly related to scientific investigation per se. But how does anticipation relate to design in particular, and what is the specific answer design science has to offer (if it has) in relation to the notion of anticipation? We will explore these questions in the following section.

## 3. Design and anticipation

Design is recognised as a fundamental instrument of change in the world. It is considered to be a natural human activity but it is also often seen as a characteristic sign of intelligence. Not surprisingly then, the study of design has been approached from many different angles and disciplines ranging for example from theology, to philosophy, biology, cognitive science, social science, computer science and engineering. Although there is no generally accepted definition of design, it is usually perceived as a purposeful, constraint decision-making process, which aims to synthesise descriptions of artefacts to be realised or implemented (Braha and Maimon 1997; Gero 2000). Seeing design as a purposeful activity implies explaining the final configuration by way of the (desired) function it fulfils; namely, providing final cause explanations seems to be one of the most distinguishing characteristics of design 'worlds'. Additionally, Smithers (2002: 7) notes that in the core of what is design(-ing) lies an apparent paradox: designing has to do with arriving at a solution to a problem which is not a-priori specified. In other words, although design is driven by a need or goal, this goal (i.e. the final cause) is actually constructed by the very process of designing. This attribute has been increasingly receiving attention in design studies, and corresponds to views of design as exploration or co-evolution where emphasis is put on the integration of problem finding and problem solving (Smithers 1998; Maher 2000; Dorst and Cross 2001). From all that we have discussed in the previous section, it seems now rather natural to associate design with anticipation; design seems to be intrinsically linked with a process of adapting to or moving towards *internally* 'constructed' expectations about future changes (note: the word internally here is used to suggest that the construction of expectations is an integral part of designing and that final cause explanations should be given from within the design process or system). This association has been hinted in various studies as we shall see below, but it has never been explicitly asserted and/or put in context. Yet, doing so would allow us to not only arrive at a potentially overarching view of design, but more importantly to provide a generic account of design as science. This has a number of upshots for the understanding and consequent development of design systems. But let us first investigate in more detail how different design studies have incorporated different notions or aspects of anticipation.

One of the most fundamental subject matters in design research is creative design, which is often linked with the phenomenon of emergence. Is there a notion of anticipation related to this fundamental aspect of design? The answer is yes, although at first sight the relation seems to be a negative one. Emergence is commonly associated with a spontaneous discovery of some new attribute (form, structure, or function) of the design description or artefact, which has not been expected or anticipated. However, this view of emergence is increasingly being challenged within the design community. For example, looking at visual emergence in design, Oxman (2002) puts forward a view of 'anticipated emergence' that contradicts the traditional definition. She suggests that the emergence of shapes is due to a process of resolution of shape ambiguities that relies both on perception, as well as cognition and the ability to 'think with shapes'. According to her approach, emergence is therefore not accidental but it is canalised by high-level cognitive schemata, which guide the resolution of shape ambiguities. Knight (2003) also examines the link between emergence and unpredictability and talks about the classification of emergent shapes in shape grammars into three classes: anticipated, possible and unanticipated. According to her view, anticipated emergence constitutes a key to analysis applications of shape grammars where the emergence of shapes is carefully predicted, whereas possible emergence involves the formation of conjectures about the emergence of shapes from (again intentionally) applied rules. Finally, unanticipated emergence is considered to play an important role in conceptual and creative stages of design and involves the generation of shapes that are not premeditated in any way. Interestingly, Knight highlights that the classification of shapes into the aforementioned categories is 'relative to the knowledge and eye of the author or user of grammar' (ibid: 135), and further elaborates ways by which emergent shapes can be used as input for further computations. Likewise, starting from a classical example of visual emergence, Brown (1998) notes that the appearance of a new shape is something that occurs even if we do have prior knowledge of the phenomenon: that is, emergence can be expected. On this basis he distinguishes between two types of emergence: the first, directly identifiable emergence, occurs when the identification of a new property can be traced back in the existing knowledge of a person, whereas the second, indirectly identifiable emergence, is linked to a discovery process of setting apart a property as interesting and hence worth remembering and classifying. Brown further suggests that identification of an emergent property comes about by way of analogical and/or functional reasoning, the latter being concerned with the use, or purpose, of a design artefact.

Working along similar lines, Gero (1998; 2003) links the notion of emergence to that of situatedness, proposing that what one 'sees' is affected by both the situation he or she is in (what is 'out there') and the previous knowledge available which guides what one is 'looking for'. A fundamental characteristic of his approach, which draws on situated views of cognition and intelligence, is the notion of constructive memory. The driving idea is that memory is not only constructed by experience, but it is also re-interpreted and re-constructed in the light of the present situation. Gero further exploits the concepts of situatedness and constructive memory to devise a model of designing that can form the basis for the development of situated design agents (Gero and Kannengieser 2002). The model considers situation as something that incorporates three different kinds of environments, the external world, the interpreted world and the expected world, linked to one another with the processes of interpretation, focussing and action. While the external world consists of representations outside the design agent, the interpreted world consists of internal representations of the part of the external world that the agent interacts with. The expected world is also formed within the agent and constitutes the environment where the effects of actions are predicted according to current goals and interpretations of the external world. The process of focussing works as a link between the interpreted and the expected world as it distinguishes aspects of the interpreted world to be used as goals in the expected world and suggests actions to be executed so as to produce states that reach these goals (ibid: 93-94). Notably, in this framework the formation of expectations is considered fundamental for both the formation of internal representations and the construction of memories. The differentiation between external, interpreted and expected world is tellingly reminiscent of the most classical and fundamental perception of anticipation, which considers that a model of the external world (or environment) constructed within a system (here design agent) could be used to form expectations about future changes and guide current action.

The ability of agents to interpret and act in the external world by constructing internal representations of this world based on memories, experiences and expectations, is generally associated with the ability to reason reflectively about the situation they find themselves in (in line with Schön's (1983) arguments about reflection in action). While Gero focuses primarily on reflective reasoning as a characteristic of individual cognition, others emphasise more the need for reflective reasoning as a consequence of the limited and partial knowledge available to agents in distributed design settings (in line with the bounded rationality hypothesis discussed above). For example, Brazier et al (2001: 137) argue that in distributed design where multiple agents need to combine their efforts to achieve a design solution, agents should be endowed with the ability to reason reflectively about additional elements, such as design partners, design culture and shared understanding of the design problem. The authors suggest that reasoning about the knowledge and experience of other agents and their expected actions and results, as well as reasoning about the types and content of interactions, involves reasoning with imprecise and incomplete knowledge and it is therefore to some extent hypothetical. Another example, where the focus on distributed design brings about the need for expectation formation, is found in Grecu and Brown (2000) The idea of anticipation is again put forward (although the word expectation is used instead) in relation to the necessity for agents to learn and evaluate advantages and disadvantages of a decision by predicting conflict situations or potential future goals and functions. The role of expectations is to compensate for restrictions related to time and information, which produce difficulties in establishing causal relationships between conditions that underline an event or action, and its results (deductively derived knowledge) (ibid: 656-657). The authors hence suggest the use of two types of expectations to guide agent learning: expectations that act as substitutes for precondition information and expectations that act as tools for inferring the effects of decisions. Finally, Zamenopoulos and Alexiou (2003) also discuss a formalisation of distributed design as a coordination problem, which takes into account the need for knowledge construction and co-evolution of the problem and solution spaces, but also explicitly suggests that this constructive process entails formation, evaluation and re-interpretation of (expected) future design descriptions (what the authors call the 'memory of the future', ibid: 193).

This brief review of studies, which establish links between anticipation and design, reveals some interesting points. Clearly, anticipation emerges as an important characteristic of designing decisively linked to phenomena such as creativity and emergence, and design 'abilities' such as generation of problems and solutions, formation of expectations, prediction and evaluation, and construction of goals and functions. However, as we mentioned at the beginning of this section, the way anticipation is perceived and expressed plays a critical role in the way design in general, and design systems in particular, are studied. The typical association of anticipation with designing as a cognitive process generally indicates a focus on individual design agents as the unit of analysis for design systems. The main assumption holding is that design abilities are fully embedded and embodied in a human or artificial design agent. That is, design abilities are coupled with, but external to the design situation or environment. Indeed, the studies mentioned above recognise the significance of reflective action, the constructive role of the situation, or the social and distributed character of design knowledge and abilities, and take steps towards a re-evaluation of the boundaries between system (design agent) and its environment (external world). Nonetheless, the explanation of design phenomena derives predominantly from the investigation of structural components of design systems – these may be human designers, design artefacts, or computer constructs

– and their interaction. Design abilities are in turn directly attributed to and explained on the basis of these structural components. Such studies are undoubtedly invaluable and have proved to be successful in many respects, but they offer no explanations as to how design abilities emerge, or how a design system is generated. Seemingly, alternative explanations are needed that can potentially go beyond the bounds of structural description and shed some light into these questions. This refers to a route of research (in line with Rosen's paradigm) that strives to offer functional explanations of design phenomena by alternatively focusing on the organisational level of design systems. Let us now examine how such a methodology could be pursued to develop a framework for a formal definition of design systems based on the concept of anticipation.

## 4. Towards an organisational view of design systems

Adopting an organisational view of design and design systems implies focussing on a fundamentally different research problem for design science. Instead of investigating the kinds of structures that can embody design abilities, we will be focusing here on the question of *how to decide if a given system is a design system*, or consequently, if an observed world is a design world. In the following we will seek to define the organisational conditions that enable design abilities to emerge in systems and then describe a specific formalisation such that any of its realisations would in principle constitute a design world.

From the studies presented above we can distil some critical requirements or conditions for encompassing anticipation in design systems. Put succinctly, a design system needs: a) to have a way to express design problems and establish relations between design variables – that is, include a language by which to carry out the generation of possible worlds where solutions can be found, b) to entail the conditions for its operation – that is, entail a way to reformulate the set of possible problems and solutions according to evidence in the actual world, and c) to entail final cause explanations – that is, construct internally future purpose(s) or function(s) for the design artefact. These conditions suggest that the realisation of design systems requires the composition of actual, possible and design worlds in a specific organisation. Before we attempt to formalise this organisation, we will first make some notes about traditional views of design systems and investigate whether they fulfil these conditions. We will argue that the majority of studies in design have implicitly or explicitly embraced the machine paradigm in science, which has restricted our ability to define design systems and understand design worlds. Let us first introduce some general notation and briefly present the main underlining assumptions behind the machine paradigm.

Adopting a system-theoretic terminology, we can generally assert that the fundamental aspiration of science is to establish causal relationships between observables (abstract equivalent classes) established within the world of interest. In this sense, an observed world *W* can be expressed by a mapping, or function, that transforms inputs to outputs (causes to effects), and can be represented as follows:

$$f: A \rightarrow B \qquad (7a)$$

We can rewrite (7a) as (7b) to distinguish between *F* that represents the binary relation between *A* and *B*, and *T* that represents the ternary relation between *F*, *A* and *B*:

$$(F, a) \xrightarrow{T} B \tag{7b}$$

In more detail, we can consider that *F* constitutes a language by which the relation between *A* (initial conditions or premises) and *B* (effects or theorems) is formalised. The definition and use of a language *F* is of critical importance in our attempt to model the world *W*: language is the tool by which we can define possible states of the world. The definition of *T* is also critical as it is somewhat responsible for the evaluation, or execution of *F*. The mapping (7b) can take many different forms. To give an example, in the field of computing *A* and *B* would represent the sets of input and output symbols respectively, *F* would represent the hardware and *T* would represent the computation process itself: the running of the program to calculate outputs *B*.

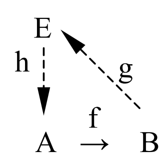
Having this notation we can now discuss the machine paradigm. In abstract terms, the machine view of the world proposes that the main question or problem to solve in order to understand, describe and generate possible worlds, is to find an effective procedure *T* (a computable or describable process) that can compute *B* given a programme or equation of states *F* and some initial conditions $a \in A$. In more logic theoretic terms, the world *W* and the effects *B* are presumably deduced by simply knowing the history of the system and the conditions *a*. The claim put forward is that there exists a describable process *T* such that can compute (or simulate) any formal description of the world *W*, namely the machine is the largest model of the world. Based on this paradigm a plethora of different abstractions has been developed, which in general either refine or introduce more structure (restrictions) to the machine formalism. For example, the paradigms of self-reproduction and evolution narrow down the machine assumptions by imposing restrictions related to complexity and hence the nature of the effect *B*: Self-reproduction assumes that there is an automaton *T* that can produce *B* such that the complexity of *B* is equal to the complexity of *F*. In this sense the world *W* is able to entail structures of the same complexity and hence reproduce itself. Evolution assumes that there is a process *T* that can generate another automaton *B* of greater complexity. In this sense the world *W* is able to increase its complexity. The proposed model of variation-selection-permutation is a specification of how this structure can be realized. Other abstractions impose some kind of a meta-structure on the machine metaphor: In cybernetics the assumption holds that there is a formal process *Fc* within a machine *T* that can reduce the complexity of the environment (expressed by *A*) by producing the appropriate *B*. In other words the effect *B* can be endowed with a purpose or function; a characteristic that was missing from machine structures. Finally, artificial intelligence generally assumes that the world *W* cannot be simply described by deductive steps: there is no such process *T* that can deductively prove *B*. The assumption put forward instead is that there is another process *Fi* at a meta-level able to define an *F* that can heuristically prove *B*. The assumptions and fundamental questions posed by the machine paradigm and its different 'offspring' are summarised in Table 1 below.

As we mentioned, design studies have generally assumed these paradigms as a way to understand and formalise design worlds and thereafter develop design systems. For example, many studies have assumed the machine metaphor to develop various algorithms and grammars to realise design worlds (Stiny and March 1981; Stiny 1991), while others have developed and used genetic and evolutionary algorithms to solve optimisation problems or generate and evolve novel forms (Frazer 1995; Rosenman 1996; Bentley 1999). Cybernetic and systemic approaches have influenced studies on design methods during the '60s (Asimov 1962; Archer 1965) and have also been implicitly or explicitly assumed in mathematical studies of design (Braha and Maimon 1998; Suh 2001). Finally, the paradigm of artificial intelligence has largely influenced the study of design as a problem solving activity formalised as search and (latter) exploration (Goel and Pirolli 1989; Chandrasekaran 1990; Smithers 1998; Maher 2000), and has generally inspired the formation of cognitive approaches to design processes and the development of knowledge based systems (Gero 1990; Coyne et al 1990). Now the question here arises: do these paradigms entail the critical conditions we considered to be necessary for encompassing anticipation in our view of design systems? If we examine the machine formalisation in more detail we will see that these fundamental conditions are in fact violated.

| | | |
|---|---|---|
| **Machine paradigm** | $(F, a) \xrightarrow{?T} B$ | Find T such that $\forall F\ C(T) \supset C(F)$ |
| **Self-organisation** | $(F, a) \xrightarrow{?T} B$ | Find T such that $C(B) = C(F)$ |
| **Evolution** | $(F, a) \xrightarrow{?T} B$ | Find T such that $C(B) > C(F)$ |
| **Cybernetics** | $(?Fc, a) \rightarrow B$<br>↓<br>$(F, b) \rightarrow A$ | Find Fc such that C(A) is reduced |
| **Artificial Intelligence** | $(?Fi, a) \rightarrow B$<br>↓<br>$(F, b) \rightarrow F$ | Find Fi such that F can be defined |

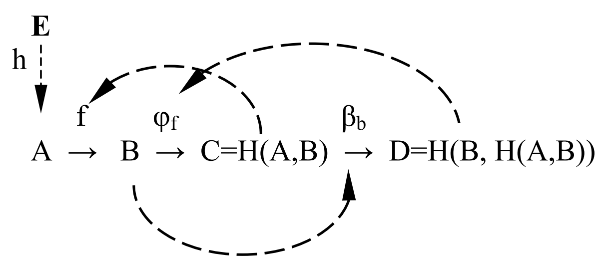
Table 1 The machine paradigm and its offspring: fundamental questions and assumptions

The realisation of the causal mapping represented in (7a) and (7b) offers a series of explanations that aim to justify why the set of effects *B* has been generated. We have previously explored Rosen's arguments with regards to this mapping. Let us have another look at these arguments with the help of the diagram (8).



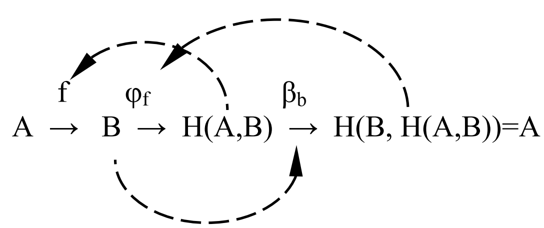

(8)

We can see more clearly that *B* can be entailed by *a* (material cause), and *f* (formal and efficient cause in this case). We can also see that there is nothing to explain *B* by its effects in this diagram, as there is nothing to ensure a function is given to *B* from within the system. Notably, the opposite holds for *f* and *a*: they can only be explained in finalistic terms (because of their function in the diagram). There is no material, formal or efficient entailment neither for *f* (the language, or program which associates observables in the world and generates possible worlds) nor for *a*, the material conditions that guide the operation of the system. This dual relationship between system and environment has some very important implications for design systems. It effectively suggests that if we are to assume a design system with the characteristics we explicated previously, then this system must be found in the environment. Because it is the environment that entails both the language of the system and the conditions of its operation, and it is the environment that assigns purpose to the system. It follows that if design worlds are to be formally represented we need to include more structure into the initial formalism: we need to include the environment into it. Note that this is not a new idea: as we saw, the inclusion of environment into the system is what abstract worlds such as cybernetic and artificial intelligence worlds have been striving to achieve. As Rosen discussed, if we wish to include the environment we need to extend the mapping (7b) by adding a new function: as seen in (3). This is a very convenient addition because $f$ can now be efficiently entailed, and a function can be given to $B$. However, it can similarly be shown that the new function $\varphi_f$ will be unentailed and the 'designer' will again have to be found somewhere in the environment. We can continue this process of adding functions infinitely… Rosen formulated the (M, R) formalism (4) so as to overcome this apparent infinite regress and close the system to efficient causation. This is an important step from the point of view of design systems. The formalism meets two of the conditions we specified at the beginning of this section: it is able to entail the definition of a language (*f*) by which to carry out the generation of possible worlds and it is able to entail its future purpose and function. Yet, there is still something missing. In diagram (9) we can see that the material cause $a$ remains unentailed: the conditions that bound the operation of the system and guide the re-definition of the language by which problems and solutions are expressed, are still to be found in the environment.




(9)

It is clear that an additional closure condition needs to be formulated. Our proposition is to define the set $D$ as a set of functions with the logical form B⇒[A⇒B] whose evaluation defines the condition $a \in A$ (10).




(10)

We have now reached our ultimate aim. Diagram (10) seemingly encodes all the necessary conditions for the definition and explanation of design systems. We suggest that this formalism provides a framework by which to decide when a given system is a design system (any realisation of this formalism would in principle constitute a design system), and can subsequently improve our ability to observe and characterise design worlds. We mentioned at the beginning of this section that the realisation of design systems requires a specific composition of actual, possible and design worlds. Let us try to elucidate this composition in more detail. The first mapping represents a function that links certain observables to their effects in the observed world, and therefore can be considered as an expression of a problem within the actual world. The second mapping corresponds to a process of rule extraction from observations, which associates initial conditions with expected results, and therefore can be considered as a description of possible worlds. The third mapping corresponds to a process which, given the general rules generated and the actual state of the world, formulates new premises or hypotheses for the observed world. As the mapping refers to a process of both producing solutions that fulfil expected outcomes and reassessing the formulation of the initial problem that the system seeks to solve, it can be considered as a characteristic expression of design worlds. In effect, although actual and possible worlds are bounded by design worlds within a hierarchical composition of mappings, observations and effects in the actual world seem to restrict the generation of design worlds at a most fundamental level. This can be written in terms of subset relations as follows:

$$\text{Actual Worlds} \subseteq \text{Possible Worlds} \subseteq \text{Design Worlds} \subseteq \text{Actual Worlds} \tag{11}$$

Before we conclude this discussion, it is important to stress that although the formalism suggested in (10) designates when a system can be considered to be a design system, it does not define the structures that could be used to realise it. In other words, the diagram does not specify individual components of a design system, but it rather explicates the conditions that relate them or link them together in terms of function. This is the most essential feature of the organisational approach.

## 5. Discussion

The motivation for pursuing this line of research is to bring the study of design at a different level of abstraction concerned with the organisational features of design systems. An organisational view of design does not only offer an account of design phenomena and abilities irrespective of their structural components (whether these are human designers, design artefacts, or computer constructs), but can also potentially broaden our ability to explore, explain and understand complex realities. If science in general is occupied with the exploration of actual and possible worlds, we see design science as the science that complements this exploration by linking actual, possible and design worlds together.

In this sense, the study presented here can be seen as a framework for reconsidering the role, scope and requirements of decision support systems. The view adopted suggests that there is a need to move from the exclusive study of algorithms and programs to the study of high-level (abstract) functions such as learning and adaptation and the exploration of organisational concepts such as coordination. Moreover, the formalism developed puts forward a view of design systems as being fundamentally open to interaction and coupled with the environment within which they operate. This is in line with most of the studies that consider the idea of anticipation in relation to design, and essentially draw attention to the importance of re-evaluating the boundaries between system and environment. Not unlike proponents of situated or distributed cognition who place emphasis on interactions between agents and with external representations, tools and artefacts (Arias et al 2000; Hollan et al 2000; Gero 2003), this study sees design(-ing) as a process developed inevitably within a socio-technical context. Rather than focussing on how design abilities can be embodied within computational constructs, this study aspires to take a step towards developing computer systems that are open to human interaction and support the emergence of design abilities through this interaction. This view is to a certain degree also relevant to emerging research on socially intelligent agents, which seeks to investigate and establish relationships among humans and machines by putting the human in the loop (Dautenhahn et al 2002).

**References**:


Alexander, C. (1963). The Determination of Components for an Indian Village. Conference on Design Methods. J. C. Jones, Thornley, D. Oxford, Pergamon Press.

Archer, B. L. (1965). Systematic Methods for designers. Developments in design methodology. N. Cross, John Wiley&Sons**:** 57-82.

Arias, E., Eden, H., Fischer, G., Gorman, A., and Scharff, E. (2000). "Transcending the individual human mind - creating shared understanding through collaborative design." ACM Transactions on Computer-Human Interaction **7**(1): 84-113.

Arthur, W. B. (1994). "Inductive reasoning and bounded rationality." The American Economic Review **84**(2): 406-411.

Asimow, M. (1962). Introduction to Design. New Jersey, Prentice Hall.

Beer, R. D. (1995). "A dynamical systems perspective on agent-environment interaction." Artificial Intelligence **72**(1-2): 173-215.

Bentley, P. J., Ed. (1999). Evolutionary Design by Computers. San Francisco, CA., Morgan Kaufmann Publishers Inc.

Boden, M. A. (1972). Purposive Explanation in Psychology. Cambridge: Massachusetts, Harvard University Press.

Braha, D., and Maimon, O. (1997). "The design process: properties, paradigms, and structure." IEEE Transactions on Systems, Man and Cybernetics - Part A: Systems and Humans **27**(2): 146-166.

Braha, D., and Maimon, O. (1998). A Mathematical Theory of Design: Foundations, Algorithms and Applications. Dordrecht, The Netherlands, Kluwer Academic Publishers.

Brazier, F. M. T., Moshkina, L.V., and Wijngaards, N.J.E. (2001). "Knowledge level model of an individual designer as an agent in collaborative distributed design." Artificial Intelligence in Engineering **15**(2): 137-152.

Brown, D. C. (1998). Functional emergence: a position paper. Emergence in Design Workshop , AID '98 Conference, Cambridge, UK.

Butz, M. V., Goldberg, D. E. and Stolzmann, C. W. (2002). "The anticipatory classifier system and genetic generalization." Natural Computing **1**(4): 427-467.

Butz, M. V., Sigaud, O., and Gérard, P. (2003). Internal models and anticipations in adaptive learning systems. Anticipatory Behavior in Adaptive Learning Systems: Foundations, Theories and Systems (LNCS 2684). M. V. Butz, Sigaud, O., Gérard, P. (eds). Berlin, London, Springer-Verlag**:** 87-110.

Casti, J. L. (1992). Reality Rules: II Picturing the world in mathematics - The frontier. New York, John Wiley & Sons.

Casti, J. L. (2002). "'Biologizing' control theory: How to make a control system come alive." Complexity **7**(4): 10-12.

Chandrasekaran, B. (1990). "Design problem solving: a task analysis." AI magazine **11**(4): 59-71.

Christensen, W. D., and Hooker, C. A. (2000). "Anticipation in autonomous systems: foundations for a theory of embodied agents." International Journal of Computing Anticipatory Systems **5**: 135-154.

Coyne, R. D., Rosenman, M. A., Radford, A. D., Balachandran, M., and Gero, J. S. (1990). Knowledge Based Design Systems. Reading, Mass., Addison-Wesley.

Dautenhahn, K., Bond, A.H., Cañamero, L., and Edmonds, B., Ed. (2002). Socially Intelligent Agents: Creating relationships with computers and robots. Dordrecht, The Netherlands, Kluwer Academic Publishers.

Dautenhahn, K., Ogden, B., and Quick, T. (2002). "From embodied to socially embedded agents - Implications for interaction-aware robots." Cognitive Systems Research **3**(3): 397-428.

Davidsson, P. (1997). Learning by linear anticipation in multi-agent systems. Distributed Artificial Intelligence Meets Machine Learning (LNAI 1221). Berlin, Heidelberg, Springer-Verlag**:** 62-72.

Dorst, K., and Cross, N. (2001). "Creativity in the design process: Co-evolution of problem-solution." Design Studies **22**(5): 425-437.

Dubois, D. M. (1997). "Generation of fractals from incursive automata, digital diffusion and wave equation systems." BioSystems **43**(2): 97-114.

Dubois, D. M. (1998). Computing anticipatory systems with incursion and hyperincursion. In Proceedings of the First International Conference on Computing Anticipatory Systems (CASYS 1997). D. M. Dubois. Melville, New York, American Institute of Physics, AIP Conference Proceedings 437**:** 3-29.

Dubois, D. M. (2000). Review of incursive, hyperincursive and anticipatory systems - Foundation of anticipation in electromagnetism. In Proceedings of the Third International Conference on Computing Anticipatory Systems (CASYS 1999). D. M. Dubois. Melville, New York, American Institute of Physics, AIP Conference Proceedings 517**:** 3-30.

Dubois, D. M., and Resconi, G. (1992). Hyperincursivity: a new mathematical theory, Presses Universitaires de Liège.

Edmonds, B. (1999). "Capturing social embeddedness: a constructivist approach." Adaptive Behavior **7**(3-4): 323-347.

Ekdahl, B. (2000). Agents as anticipatory systems. In 4th World Multiconference on Systemics, Cybernetics and Informatics (SCI 2000) and 6th International Conference on Information Systems Analysis and Synthesis (ISAS 2000). Orlando, Florida. **Vol. II:** 133-137.

Ekdahl, B., Astor, E. and Davidsson, P. (1995). Towards anticipatory agents. Intelligent Agents: theories, architectures and languages (LNAI 890). M. a. J. N. R. e. Wooldridge. Berlin, London, Springer-Verlag**:** 191-202.

El hadouaj, S., Drogoul, A., and Espié, S. (2000). How to combine reactivity and anticipation: the case of conflicts resolution in a simulated road traffic. Multi-Agent Systems and Agent-Based Simulation 2000 (LNAI 1979). S. a. D. Moss, P. (eds). Berlin, Heidelberg, Springer-Verlag**:** 82-96.

Frazer, J. (1995). An Evolutionary Architecture. London, Architecture Association.

George, F., and Johnson, L., Ed. (1985). Purposive Behaviour and Teleological Explanations. New York ; London, Gordon and Breach Science Publishers.

Gérard, P., Stolzmann, W. and Sigaud, O. (2002). "YACS: a new learning classifier system with anticipation." Soft Computing **6**(3-4): 216-228.

Gero, J. S. (1990). "Design prototypes: a knowledge representation schema for design." AI magazine **11**(4): 26-36.

Gero, J. S. (1998). Towards a model of designing which includes its situatedness. Universal Design Theory. H. Grabowski, Rude, S. and Grein, G. (eds). Aachen, Shaker Verlag**:** 47-56.

Gero, J. S. (2000). "Computational models of innovative and creative design processes." Technological Forecasting and Social Change **64**(2-3): 183-196.

Gero, J. S. (2003). Situated computing: a new paradigm for design computing. In Proceedings of the 8th CAADRIA 2003 Conference. A. Choutgrajank, Charoenslip, E., Keatruangkamala, K., and Nakapan, W. (eds). Thailand, Rangsit University**:** 579-587.

Gero, J. S., and Kannengiesser, U. (2002). The situated function-behaviour-structure framework. Artificial Intelligence in Design '02. J. S. Gero. Dordrecht, The Netherlands, Kluwer Academic Publishers**:** 89-104.

Glasersfeld, E. v. (1998). Anticipation in the constructivist theory of cognition. In Proceedings of the First International Conference on Computing Anticipatory Systems (CASYS 1997). D. M. Dubois. Melville, New York, American Institute of Physics, AIP Conference Proceedings 437**:** 38-47.

Goel, V., and Pirolli, P. (1989). "Motivating the notion of generic design within information processing theory: The design problem space." AI magazine **10**(1): 19-36.

Grecu, D. L., and Brown, D.C. (2000). Expectation formation in multi-agent design systems. Artificial Intelligence in Design '00. J. S. Gero. Dordrecht, The Netherlands, Kluwer Academic Publishers**:** 651-671.

Hoffmann, J. (1993). Vorhersage und Erkenntnis [Anticipation and Cognition]. Goettingen, Germany, Hogrefe.

Hollan, J., Hutchins, E., Kirsh, D. (2000). "Distributed cognition: toward a new foundation for human-computer interaction research." ACM Transactions on Computer-Human Interaction **7**(2): 174-196.

James, W. (1890). The Principles of Psychology. London, Macmillan.

Kelly, G. A. (1955). The Psychology of Personal Constructs. New York, Norton.

Knight, T. (2003). "Computing with emergence." Environment and Planning B: Planning and Design **30**(1): 125-155.

Maher, M. L. (2000). "A model of co-evolutionary design." Engineering with Computers **16**(3-4): 195-208.

Maturana, H. R., and Varela, F. J. (1980). Autopoiesis and Cognition. The Realization of Living. Dordrecht, Holland, D. Reidel Publishing Company.

Neisser, U. (1975). Cognition and Reality. San Franciso, Freeman.

Oxman, R. (2002). Anticipated emergence. Artificial Intelligence in Design '02 Accepted Posters, Cambridge, UK.

Pfeifer, R. (1995). "Cognition - Perspectives from autonomous agents." Robotics and Autonomous Systems **15**(1-2): 47-70.

Piaget, J. (1954). The Construction of Reality in the Child. New York, Ballentine.

Prem, E. (1997). Epistemic autonomy in models of living systems. In Proceedings of the 4th European Conference on Artificial Life. P. Husbands, and Harvey, I. (eds). Cambridge, Massachusetts, MIT Press**:** 2-9.

Quick, T., Dautenhahn, K., Nehaniv, C. L. and Roberts, G. (2000). The essence of embodiment: a framework for understanding and exploiting structural coupling between system and environment. In Proceedings of the Third International Conference on Computing Anticipatory Systems (CASYS 1999). D. M. Dubois. Melville, New York, American Institute of Physics, AIP Conference Proceedings 678**:** 649-660.

Rashevsky, N. (1954). "Topology and life: in search of general mathematical principles in biology and sociology." Bulleting of Mathematical Biophysics **16**: 317-348.

Riegler, A. (2001). The role of anticipation in cognition. In Proceedings of the Fourth International Conference on Computing Anticipatory Systems (CASYS 2000). D. M. Dubois. Melville, New York, American Institute of Physics, AIP Conference Proceedings 573**:** 534-541.

Riegler, A. (2003). Whose anticipations? Anticipatory Behavior in Adaptive Learning Systems: Foundations, Theories and Systems (LNCS 2684). M. V. Butz, Sigaud, O., and Gérard, P. (eds). Berlin, London, Springer-Verlag**:** 11-22.

Rosen, R. (1985). Anticipatory Systems. Philosophical, mathematical and methodological foundations. Oxford, Pergamon Press Ltd.

Rosen, R. (1991). Life Itself. A comprenhensive inquiry into the nature, origin, and fabrication of life. New York, Columbia University Press.

Rosen, R. (2000). Essays in Life Itself. New York, Columbia University Press.

Rosenman, M. A. (1996). The generation of form using an evolutionary approach. Artificial Intelligence in Design '96. J. S. a. S. Gero, F. (eds). Dordrecht, Kluwer**:** 643-622.

Schön, D. A. (1983). The Reflective Practitioner: How professionals think in action. Aldershot, Arena.

Simon, H. A. (1957). Models of Man: Social and Rational - Mathematical essays on rational human behaviour in a social setting. New York, Wiley.

Skinner, B. F. (1938). The Behavior of Organisms. New York, Appleton-Century Crofts.

Smithers, T. (1995). "On quantitative performance measures of robot behaviour." Robotics and Autonomous Systems **15**(1-2): 107-133.

Smithers, T. (1998). Towards a knowledge level theory of design process. Artificial Intelligence in Design '98. J. S. Gero, and Sudweeds, F. (eds). Dordrecht, Nertherlands, Kluwer**:** 3-22.

Smithers, T. (2002). Synthesis in design. Artificial Intelligence in Design. J. S. Gero, (ed). Dordrecht, The Netherlands, Kluwer**:** 3-24.

Stiny, G. (1991). "The algebras of design." Research in Engineering Design **2**: 171-181.

Stiny, G., and March, L. (1981). "Design machines." Environment and Planning B **8**(3): 245-255.

Stout, R. (1996). Things That Happen Because They Should: A teleological approach to action. New York, Oxford University Press.

Suh, N. P. (2001). Axiomatic Design. New York, Oxford University Press.

Thòm, R. (1972). Structural Stability and Morphogenesis. New York, Benjamin.

Tolman, E. C. (1932). Purposive Behavior in Animals and Men. New York, Appleton.

Tsoukalas, L. H. (1998). "Neurofuzzy approaches to anticipation: a new paradigm for intelligent systems." IEEE Transactions on Systems, Man and Cybernetics - Part B: Cybernetics **28**(4): 573-582.

Veloso, M., Stone, P., and Bowling, M. (1999). Anticipation as a key for collaboration in a team of agents: a case study in robotic soccer. In Proceedings of SPIE Sensor Fusion and Decentralized Control in Robotic Systems II. Boston. **Vol. 3839**.

Von Neumann, J., and Morgenstern, O. (1944). Theory of Games and Economic Behavior. Princeton, Princeton University Press.

Xinqing, L., Tsoukalas, L. H., and Uhrig, R. E. (1996). A neurofuzzy approach for the anticipatory control of complex systems. In Proceedings of the Fifth IEEE International Conference on Fuzzy Systems, IEEE. **1:** 587-593.

Zamenopoulos, T., and Alexiou, K. (2003). Computer-aided creativity and learning in distributed cooperative human-machine networks. In Digital Design: Research and Practice. Proceedings of the 10th International Conference on CAAD Futures 2003. M.-L. J.-Y. Chiu, T., Kvan, T., Morozumi, M., and Jeng, T. (eds). Dordrecht, The Netherlands, Kluwer Academic Publishers**:** 191-201.