UNIVERSITY OF LONDON THESIS

Degree PhD Year 2007 Name of Author ALEXIOU, Aikaterini

Understanding Multi-Agent Design as Coordination

Aikaterini Alexiou

Bartlett School of Graduate Studies

Faculty of the Built Environment

University College London

University of London

Thesis Submitted for the Degree of Doctor of Philosophy

2007

Abstract

Design decision making increasingly involves the participation of multiple agents which bring into the design process multiple, and often conflicting, needs, knowledge, and goals. To the human agents (experts from the same or different domains, clients, users, stakeholders) one should add artificial agents (computational models and tools more generally) that play an important part in the process. Design research has considered the issue of distributed decision making mainly through the concepts of cooperation and collaboration. The present thesis argues that coordination is a more apposite concept for capturing the social distributed character of design. The concept of coordination places emphasis on issues of interdependency, complexity and distribution and enables us to understand design at a systemic/organisational level, without making assumptions about agents' commitment to a common goal, or their disposition towards cooperation or conflict. Additionally, coordination is used to capture the generative, creative aspects of distributed design decision making.

The study explores and establishes the meaning of coordination through experimentation with computational models and simulations. The very process of building these models is a vehicle for exploring key hypotheses and assumptions, and developing a coherent theoretical construction.

Overall, the thesis identifies the key dimensions of coordination that are typical to the domain of design, and employs them to develop a framework (a theory) for understanding multi-agent design as a generative social process. The dimensions identified are learning, decentralised control and co-evolution. A model of coordination developed using the paradigm of distributed learning control is used as a way to establish the precise meaning of these dimensions. Based on insights from the experimentation, the concept of coordination is further refined in order to propose an organisational (complexity-informed) perspective of multi-agent design. According to this perspective, the relationship between agents, their goals, and the design variables they manipulate, is at the same time a product of the design process, but also a constraint over individual agents. Coordination is then defined as a dynamic process towards a scheme of organisation that entails the emergence of collective design solutions.

Acknowledgements

I would like to thank first of all my supervisors: Phil Steadman for his most kind support as a teacher and as a friend, and Mike Batty for taking us in and helping us out all the way through our venture. I would also like to express my gratitude to Jeff Johnson for motivating us to finish and supporting us in the best possible way.

It goes without saying that everyone in CASA (students, researchers, staff, affiliated professors and visitors) has contributed to providing a special environment to surround our efforts. But, for their friendship, I would particularly like to thank the 'people in our corner'. Here is the place to commemorate pizza club evenings, late night sprints to evacuate the building, discussions about Marx and complexity, the vision of green coconuts, solidarity hugs and hugs for no reason, boxing and other athletic games, CASA*Today* contributions, and the discussions on all that constituted the small and the big in our lives in London.

I would also like to express my special thanks to Vicky Braouzou. Without her, we might not have even been here.

The computational models and simulations reported in Chapter 5 of the thesis were developed jointly with Theodore Zamenopoulos: each of us was responsible for 50% of the design, implementation and experimentation with these models. Being my partner in crime, a collaborator, and the devil's advocate, I feel a thank you note won't do him justice. It was a great gift for us to be given the opportunity to work together and 'practice what we preach' about collaboration and distributed minds.

I also wish to acknowledge EPSRC for a doctoral training account grant, which partly supported my research from September 2002 to September 2003.

Finally, I wish to thank my brother, for his phone calls, and my parents, for their immense support - whether financial, emotional, or psychological. I hope life shows I made a good use of it. The thesis is dedicated to them.

Table of Contents

List of Figures

List of Tables

Chapter 1
INTRODUCTION

> The sceptical reader may wish to look around right now and see whether there is anything in the current environment that was not either produced or delivered to its present location by the cooperative efforts of individuals working in socially organized groups. The only thing I can find in my environment that meets this test is a striped pebble that I found at the beach and carried home to decorate my desk. (...) Every other thing I can see from my chair not only is the product of coordinated group rather than individual activity, but is necessarily the product of group rather than individual activity.
>
> (Hutchins, 1995: 176)

1.1 SETTING THE SCENE/ MOTIVATION

Design has been increasingly perceived as a social, cooperative activity. This comes from the recognition that most design projects draw upon knowledge distributed to networks of diverse participants (experts or stakeholders), which bring into the design process multiple views, needs, knowledge and goals. Moreover, design practice has become more and more connected to the use of computational constructs that mediate and support humans in their activities. The term *multi-agent* will be adopted here instead of cooperative, to reflect the fact that humans and artificial constructs are considered to be equally important agents in design. But there is a second reason to opt for the term multi-agent instead of cooperative: cooperation usually assumes (or connotes) some kind of formal or informal agreement to 'work together', whereas cooperation, in the sense that is studied here, may in fact include design processes that emerge without any deliberate conformity to work together. Finally, the term multi-agent with its connotations of distributed artificial intelligence and multi-agent systems seems to reflect more accurately the distributed and emergent nature of design phenomena.

It is also true to say that notwithstanding whether a particular design is created by a group or not, it is within a group or social environment that it becomes realised. The desires and needs that motivate design are developed within a social context, while in turn design artefacts are directed towards this environment and ultimately attain their functions within it. The term multi-agent is therefore also used here in recognition of the fact that the social context of design is an integral part of the process as it provides both the requirements and the final evaluations of the design artefacts.

The original motivation for this thesis was to develop models and tools to support multi-agent design decision making. But what are the requirements for supporting design in distributed human-computer networks? There is a considerable literature on collaboration, cooperation and participation in design, especially from the perspective of Computer Supported Collaborative Work (CSCW), Decision Support Systems (DSS), and Human-Computer Interaction (HCI). But each study makes different assumptions about the design decision making process according to the particular domain of application (for example, urban planning versus architectural design), and the objective of the application (for example, to support expert and non-expert communication, to encourage participation, to support problem solving, to support conflict resolution, or to support creativity). It gradually became clear that what was needed was a deeper, and also more generic, understanding of design as a distributed multi-agent process which is not tied to a particular application. An abstract concept was required able to capture the social distributed character of design, without making strong assumptions about the kinds of interdependencies that exist among agents and their goals, or about their dispositions towards cooperation or conflict. This thesis proposes that the most appropriate concept for this purpose is the concept of coordination.

Malone and Crowston (1990: 361) offer one of the most comprehensive definitions of coordination, as *'the act of managing interdependencies between activities performed to achieve a goal'*. The word coordination is used to include terms such as cooperation, collaboration or conflict, so long as they involve managing interdependencies among activities. Malone and Crowston (1994: 91) further classify the wide variety of interdependencies that arise in complex processes under four major categories: shared resources and task assignments (which may involve for example priority ordering); producer/consumer relationships (which may involve sequencing, standardization, or synchronization); simultaneity constraints (which involve scheduling and synchronization); and task/subtask dependencies (which usually involve goal selection and decomposition). Coordination theory is a rather new research area, developed to bring together disciplines like organisational science, management science, economics, computer science and psychology, to the study of coordination in complex systems. This interdisciplinary study of coordination is ultimately intended to work as a theoretical framework for research in computer supported cooperative work, and distributed and parallel computer systems.

Thus defined coordination is a key problem in multi-agent design deriving directly from the decentralisation and distribution of knowledge, processes and decisions. For example, the success of large development and regeneration projects across the world relies (among other things) on the effective coordination of goals and activities between the agencies concerned (ranging from local stakeholders to development firms, consultants, local authorities and so on). The Architecture-Engineering-Construction (AEC) domain is another typical example of multidisciplinary multi-agent design where coordination is crucial. In the AEC industry architects, structural engineers, mechanical engineers, and contractors, all work on the same artefact (a building) yet by utilising their own expertise, views, and models. There is in this case an imperative need for coordinating the various tasks and representations and for managing interdependencies and inconsistencies.

However, the subject of investigation in this thesis is multi-agent design as a constructive *generative* decision making process and not only as management. So there is a need for further developing the concept of coordination beyond management and specifically in relation to the question of how coordinated solutions (particularly novel or creative solutions) are generated in distributed design settings. The aim of the research project reported is therefore essentially two-fold: to elaborate the concept of coordination in a way that is appropriate for (generative) design and to develop a view of multi-agent design as coordination such that it successfully captures its distributed, social character.

To address these questions the thesis draws from various fields and disciplines like design science, distributed artificial intelligence, complex systems science, cognitive and social science, and delves into critical issues such as learning, sociality, creativity, and emergence. Computational modelling and simulation are used as a critical method for thinking through the assumptions and consequences of the theoretical ideas and for testing their coherence. Ultimately, the fundamental conviction of this thesis is that having a rigorous framework for understanding multi-agent design coordination, and the conditions that lead to the generation of coordinated solutions, will help to better exploit the creative capacity of complex distributed human and artificial networks, and inform the development of future design decision support systems.

1.2 SUMMARY OF HYPOTHESES AND OBJECTIVES

The overall objective of the thesis is to develop a coherent framework for understanding multi-agent design with an appreciation for its social and generative, creative, character. The driving hypothesis is that coordination is an appropriate concept for this purpose as it helps us to focus on the important aspects of distribution and interdependency without being tied to any particular assumption about the domain and objectives of the design decision making process.

From this general objective a set of more specific objectives can be derived:

- To introduce the concept of coordination by drawing from studies in different fields and explicate its relevance for multi-agent design
- To identify the key dimensions of multi-agent design as coordination
- To experiment with computational models and simulations in order to evaluate, and build upon, the established dimensions of coordination

1.3 METHODOLOGY

Although this research seeks to develop a framework for coordination in multi-agent design drawing from the relevant literature, the very process of building computational constructs to implement hypotheses and ideas is a significant methodological tool. Development and experimentation with computational models is intended as a vehicle to explore and develop the main hypotheses of this research and evaluate their applicability.

1.4 OUTLINE OF CHAPTERS

The current chapter offers an introduction to the field of study and clarifies the scope, purpose and main hypotheses of the research. The second chapter focuses on the key concept of coordination. It offers a review of the term from different perspectives, clarifies its relation with concepts of cooperation, conflict and collaboration, and extracts some basic conditions or dimensions that are considered critical for multi-agent design. The third chapter discusses in more detail the methodological approach adopted and the reasons behind it. The fourth and fifth chapters examine how the critical dimensions of coordination can be made more precise through a combination of conceptual and computational modelling. Together they investigate in depth the implications of the basic argument of the thesis and highlight important questions and lessons learnt. The sixth and seventh chapters continue with a more complexity-informed perspective of coordination as an abstraction of multi-agent design, and grapple with some of the questions arising from the experimentation. In particular the two chapters focus on the important subjects of sociability and emergence. The thesis concludes with a summary and discussion in chapter eight. The overall structure of the thesis is illustrated in Figure 1.

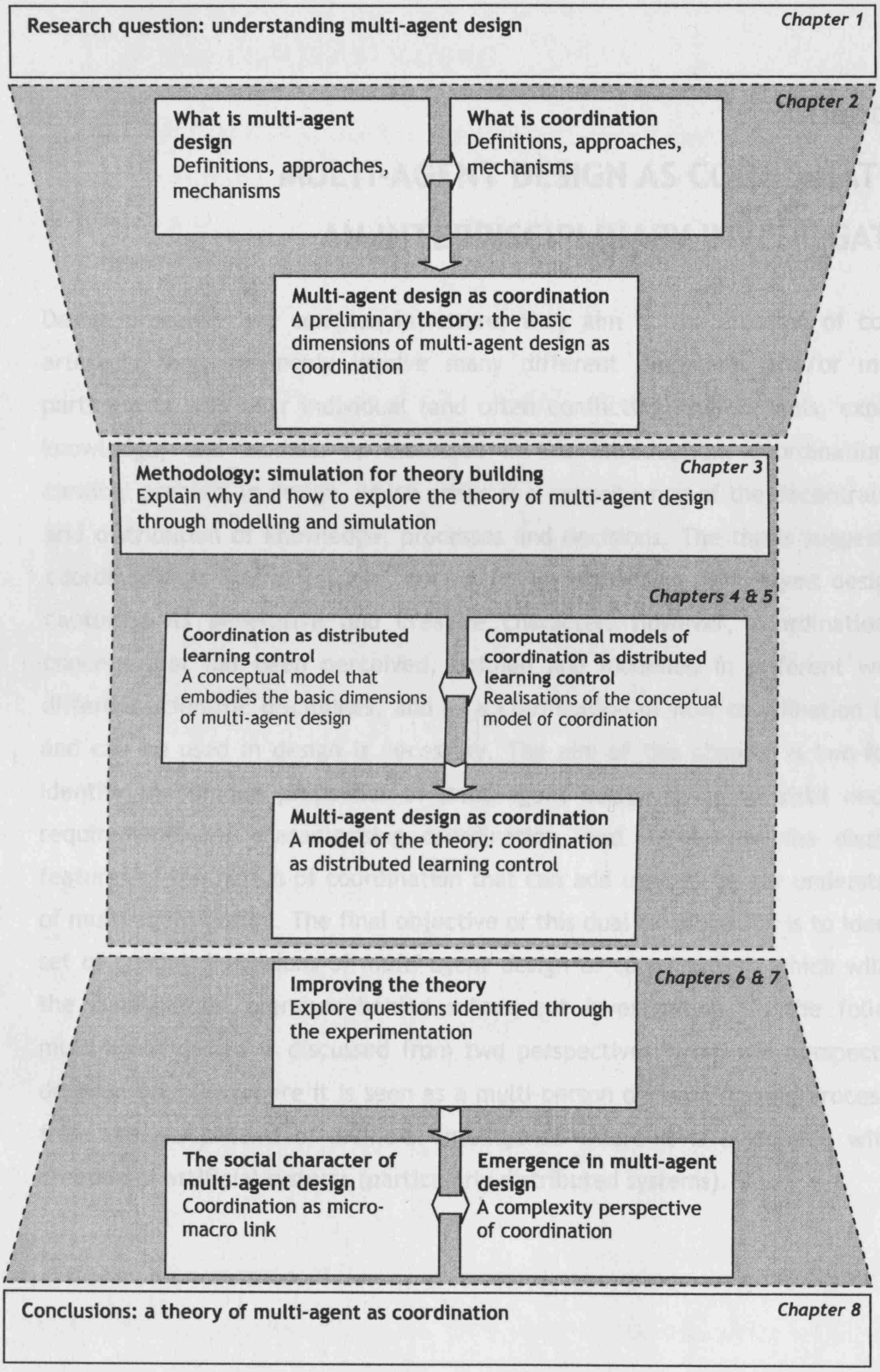


Figure 1 An outline of the thesis

Chapter 2

MULTI-AGENT DESIGN AS COORDINATION: AN INTERDISCIPLINARY INVESTIGATION

Design processes are complex in nature; they aim at the creation of complex artefacts; and commonly involve many different disciplines and/or multiple participants who bear individual (and often conflicting) views, goals, expertise, knowledge, and models. As discussed in the introduction, coordination is a cardinal problem in design, which arises as a consequence of the decentralisation and distribution of knowledge, processes and decisions. The thesis suggests that coordination is also a suitable concept for understanding multi-agent design and capturing its generative and creative character. However, coordination is a concept that has been perceived, defined and modelled in different ways by different scientific disciplines, and so a clarification of how coordination is used and can be used in design is necessary. The aim of this chapter is two-fold: to identify the unique properties of multi-agent design so as to distil necessary requirements for characterising coordination, and to identify the distinctive features of the notion of coordination that can add usefully to our understanding of multi-agent design. The final objective of this dual investigation is to identify a set of crucial dimensions of *multi-agent design as coordination*, which will form the fundamental premises behind subsequent investigation. In the following, multi-agent design is discussed from two perspectives: from the perspective of decision sciences where it is seen as a multi-person decision making process, and from the perspective of artificial intelligence where it is associated with the creation of artificial systems (particularly distributed systems).

2.1 DECISION SCIENCES: DESIGN AND PLANNING AS MULTI-PERSON DECISION MAKING

The field of decision sciences is concerned with understanding and improving the decision making of individuals, groups and organisations (Kleindorfer, Kundreuther and Schoemaker, 1993: 3). Drawing from disciplines that range from psychology and sociology, to economics and management science, research in decision sciences encompasses both descriptive (empirical, explanatory and experiment-based) and prescriptive (formal, normative and usually technology-based) approaches to the investigation of theories, methodologies and techniques related to choice, problem solving, and problem finding. Decision making and urban planning are linked together on a fundamental level, since they both follow the same stages - which in general terms involve: defining a problem and setting goals; identifying resources and constraints; formulating alternatives; projecting the outcomes of these alternatives; and evaluating them in relation to goals and outcomes (Harris, 1972: 9; Alexander, 1992: 47). Similarly, design decision making has largely been identified with a process of creative problem solving, which involves three interdependent sub-classes of activities, namely: problem representation, solution generation and solution evaluation (Rowe, 1987: 56).

The research reported here, however, is particularly focused on multi-person decision making (group, organisational and societal), which involves dealing with complex problems and processes requiring cooperation between multiple decision makers with diverse knowledge and skills, and resolution of conflicts that may arise. Consequently, key approaches and definitions of *cooperation*; *collaboration*; *conflict* and *conflict resolution*; *collective choice*; and *consensus*, will constitute here the centre of attention for the understanding and definition of coordination.

2.1.1 Coordination in group decision making

Kleindorfer et al (1993: 213) write: 'In addition to the compelling social reasons for group cooperation and coordinated action, it is also clear that information-processing and physical limitations of individuals imply the desirability, indeed necessity, that certain problems be solved by groups'. Similarly, Domeshek et al notice: 'The lone design genius, if not mythical or completely extinct, is surely on the endangered species list. Nowadays, significant design projects require teams of designers coordinating their varied expertise to arrive at effective design solutions' (Domeshek, et al, 1994: 143). Truly, group decision making has become pervasive - and not only in the sphere of decisions that have social impact. That is because most problems (e.g. managerial, political and design and planning problems alike) call for solutions produced by individuals organised in groups, able to combine their individual expertise in order to pursue common goals or maximise some joint utility. It follows however, that group effectiveness is directly linked with cooperative behaviour[1].

Yet, there are many different approaches to studying choice and cooperation within groups. For example, as Ackoff and Emery (1972) point out, most psychologists tend to be occupied with the particularities of individual behaviour in groups; social psychologists focus on the interrelationships of people acting within a group; and sociologists are concerned with the behaviour of groups taken as a whole (collective behaviour). Similarly, scientists working on software and computer systems to support group work and decision making may focus on assisting the individuals who make up the group or, alternatively, on facilitating the overall group process.

As the intention here is not to provide a comprehensive review of group decision making, we will focus on those (descriptive and prescriptive) approaches that are more relevant to urban planning and building design or architecture. To this end, the following sections include an examination of the system-theoretic approach of group planning developed by Ackoff (1974), and a review of critical approaches to group choice and cooperation developed in game theory.

[1] Note that the benefits of group decision making do not come without certain potential liabilities, such as 'groupthink', conformity, or polarization. Kvan (2000: 410) underlines three factors that determine successful (collaborative) group performance: task interdependence (how closely group members work together); outcome interdependence (whether and how group performance is rewarded); and potency (members' belief that the group can be effective). The different theories and approaches of group decision making offer different mechanisms - such as feedback, learning or dialectical reasoning - with the aim to develop a level of trust and cooperation among group members and overcome any potential drawbacks.

2.1.1.1 Group planning in the systems age

Ackoff and Emery (1972) adopt a view of social groups as systems, which implies understanding the relationships between groups and individuals as relationships between systems and their elements. They define a social group as 'a purposeful system whose members are purposeful individuals and who are intentionally coproducers of a common objective' (ibid: 213). The intention to pursue (one or multiple) *common goals* or *objectives* seems to be an essential characteristic of a group and provides a framework for understanding group interactions and conflicts. Ackoff and Emery (ibid) further distinguish between 'goals', 'objectives' and 'ideals', and establish the concepts of 'goal-seeking', 'purposeful' and 'ideal-seeking' behaviour as a foundation for understanding human and social interaction and behaviour. A *goal* is a preferred outcome which 'can be obtained within a specified period of time', whereas an *objective* is a desired outcome that 'cannot be obtained within a specified period of time', 'but is obtainable at a later time'. Purposeful systems therefore have the ability to choose between goals and courses of action so as to achieve their objectives. Ideal seeking behaviour on the other hand refers to the high level ability of purposeful systems to pursue objectives that 'cannot be obtained in any time period but can be approached without limit': *ideals*.

Ackoff (1974) transfers this teleological approach to the planning process (strategic planning), highlighting the importance of specifying ideal future states as a reference point for planning (Kleindorfer et al, 1993: 237). He advocates 'interactive' planning as a process that involves group choice and decision making oriented to the achievement of an idealized future. Such planning practice is participative, coordinated, integrated and continuous: it facilitates the participation of a wider public which can be directly involved in the design of its future; facilitates the collective and continuous formulation and reformulation of objectives on multiple levels; and improves awareness of conflicts and constraints which can be alleviated to achieve consensus (Ackoff, 1974: 29-31). To see planning as a purposeful group choice endeavour does not mean to overlook conflict. Conflicts do arise due to the presence of contrasting individual objectives, or even due to disagreement about the courses of action that should be followed to pursue the shared objective. Cooperative behaviour is desirable in such situations as a way to improve group's effectiveness and control conflict.

Ackoff's view of planning as a social group process aims at the reconciliation of the systemic view of planning with organisational aspects of human behaviour. The proposition is that processes such as conflict and cooperation should be seen as essential aspects of planning, and should be studied in the light of the relationships developed within and between groups and organisations. Moreover, this view suggests that identification, pursuit, and coordination of collective and individual goals, is instrumental for the development of cooperation. Although systemic approaches to planning have come to pass, the idea of progress towards desirable goals (when neither the goals nor the means to achieve them are necessarily known in advance, or indeed unchangeable), and the ideas of intentionality and anticipation more generally, are still considered to be of paramount importance in planning.

This view is crucial for the present research as it is also relevant to design. Design is often defined as a purposeful human activity (Simon, 1969; Rosenman and Gero, 1998), which moves from some perceived need for change towards a new (previously unknown) state that satisfies this need. But goal-seeking and goal articulation relates to problem making, which is an integral part of design. Smithers (2002: 7) writes: '... designing must start with something that neither specifies what is required nor defines a problem to be solved, yet it must arrive at a design - a specification - for something that, when realised or implemented, will satisfy the motivating needs or desire... This apparent paradox - arriving at a kind of solution without starting with a problem - is what makes designing different from other activities. It is the characteristic feature of designing... Designing resolves this paradox by actively constructing the problem or problems whose solution or solutions can form a design or parts of design.' Smithers further argues that problem solving and problem making are iterative and tightly intertwined activities; a view that follows Rittel's position that 'problem understanding and problem resolution are concomitant to each other' (1984a: 137). This argument derives in principle from the understanding and characterisation of design problems as 'wicked' problems (Churchman, 1967; Rittel and Webber 1973;

Bazjanac, 1974; Rittel 1984a). Although Smithers attempts to distinguish planning from design on that basis, it is reasonable to argue that the kind of planning problems we are interested in here (in the domain of physical and location planning), are indeed akin to this class of 'wicked' problems, mainly due to their social and multi-participatory nature. Actually, Rittel did refer to the fields of urban planning and urban development as areas that could benefit from his approach to design methods (1984b).

But, to return to the issue of coordination, the reason why such problems require a co-evolutionary approach to problem solving and problem making, is the fact that design knowledge and decision making are '*distributed* among many people, in particular among those who are likely to become affected by the solution' (Rittel, 1984b: 320 - emphasis not in original). This fact, which Rittel called 'symmetry of ignorance', is a very important aspect of design, often overlooked in design theories that focus on the capabilities of individual designers. For the present research this phenomenon is considered to be an inherent characteristic of group decision-making, as well as an inherent characteristic of multi-agent design in general. If we take this view further, we can start understanding the iterative exploration of problem formulations and problem solutions, and the continuous pursuit and identification of goals, as a process equivalent to coordination (Alexiou and Zamenopoulos, 2002a). The subject of coordination is not only related to the effective generation and reconciliation of common goals and expectations (or ideals as Ackoff suggested), or to the combination of individual expertise, but also to the reconciliation of problems and solutions. Coordination is the equivalent of a process towards goals, as well as a process towards (re-) solutions.

It is worth mentioning here that *co-evolution* of problem and solution has been extensively discussed as a critical aspect of creative design (Maher, 1994; Smithers, 1998; Dorst and Cross, 2001). For example, the function-behaviour-structure model developed by Gero (1990; 2000), and its latest elaboration into a situated framework for designing (Gero and Kannengiesser, 2002), offers a quite comprehensive account of design as a 'goal-oriented, constrained, decision-making exploration, and learning activity' (Gero, 1990: 28). This framework models the complex formulation, synthesis, analysis, evaluation, and reformulation processes of (conceptual/creative) design that transform functional, behavioural and structural variables used to represent the problem and solution spaces, and explicitly takes into account the role of goals and expectations. Nonetheless, the focus in this work is on designing as a single agent process. Notably, although the notion of co-evolution has been increasingly perceived as a key aspect of creative design processes, it has only been understood as an individual activity or ability (even in the cases where design is taken as a process that unravels within a social context) and it has not been linked to the issue of coordination in group design and decision making.

2.1.1.2 Group choice and games

The theory of games was introduced with the seminal work of John Von Neumann and Oskar Morgenstern (1944) and can be defined as 'a theory of rational decision in conflict situations' (Rapoport, 1974: 1). It is primarily concerned with formal models of choice in groups seen from the point of view of *individual* 'players' who seek to find the appropriate 'strategies' so as to maximise their expected 'payoffs' in the 'outcomes' associated with those strategies. Game theory is therefore concerned with modelling and predicting rational decision making of individuals when the outcomes of the game depend in part on the other players, who may have different preferences and beliefs. Rationality suggests that players have consistent preferences and act so as to satisfy their own (and only their own) utilities. Moreover, rationality implies that each player expects that the same will hold for every other player (Rapoport, 1974: 1). Typically, players use knowledge about other players' preferences and beliefs in order to choose their strategies[2].

[2] One of the important aspects of games involves assumptions made about information available to players. In setting a game one must specify whether players have complete information about other players' payoffs as well as their own. Another issue related to the question of information, is captured in the distinction between simultaneous-move and sequential-move games. Simultaneous-move games are usually represented in matrices ('normal form' games), whereas sequential-move games are usually represented in a tree form ('extensive form' games). Games in extensive form show, in a graph-like diagram, the order of feasible actions/moves for each player and the information available to each player at each node/point of action (the 'information set'). The notion of beliefs mentioned here, captures what players know about the game and what they expect from each other (a probability distribution on the information set). Beliefs, goals, commitments and so on, form the basis of rational reasoning and have also been used extensively in artificial intelligence.

Cooperation or defection? The Prisoner's Dilemma Game

The assumption of rationality has significant implications for the understanding of conflict, and the emergence of cooperation and coordination. The most famous example used to illustrate those concepts is the Prisoner's Dilemma Game. The game is presented in Table 1 for a problem that involves provision of a collective good: building a highway to connect a suburb to a city (adapted from Hopkins, 2001: 87). We assume that we have two players, each representing an investor interested in land development in the suburb. Each investor has two choices: either to build (Strategy C) or not to build (Strategy D) or, in other words, either to join in building the highway ('cooperate') or to 'defect'. Let us then assign a benefit of 10 from the built highway, and a cost of 15 for building it. If any one of the investors decides to build and the other chooses not to, then the lone builder has to pay all the costs while the other will still benefit from the public highway (payoff entries 10, -5 and -5, 10). If both investors decide to build, then they enter a joint venture, which leaves them with a net benefit of 2.5 each. If they both decide not to join, there are no benefits and no costs (utility payoffs 0, 0).

	Investor 2	
	Build (C_2)	Don't build (D_2)
Investor 1		
Build (C_1)	2.5, 2.5	-5, 10
Don't build (D_1)	10, -5	0, 0

Table 1 Prisoner's dilemma game for a collective good

Individual rationality suggests that the players should choose not to build because this choice brings each of them - individually - the greater benefit. The dilemma here arises because the resulting outcome (0, 0) is obviously inferior to the (2.5, 2.5), which could be achieved if both players chose to build (therefore acting 'irrationally'). The (D_1, D_2) strategy is a unique Nash Equilibrium for the Prisoner's Dilemma Game. The Nash Equilibrium concept exemplifies a non-cooperative or individualistic solution concept (Kleindorfer et al, 1993: 246): it refers to a collective strategy such that no individual player acting on his own can improve his situation by departing from this strategy. Pareto solutions on the other hand are strategies to which at least one player can be better off, without making anyone else worse off. So, the Pareto solution concept embodies some kind of *group rationality*. For the Prisoner's Dilemma Game the Pareto solution set contains every strategy apart from the Nash Equilibrium.

The core question in the Prisoner's Dilemma Game, but also in 'social dilemma' games more generally (multi-person Prisoner's Dilemma Games), is to predict when players' choices may lead to cooperative (rational from the group perspective) or non-cooperative (rational from the individual perspective) solutions. Thus the division of games into *cooperative* and *non-cooperative*. In cooperative games, players can agree to coordinate their strategies so as to promote a joint interest (through the formation of coalitions). In measuring the degree and nature of cooperation one must take into account a wide variety of factors, which are related to the context of the game; the attributes of the players; and the interactions developed among them (Axelrod, 1984: 28). Kleindorfer et al (1993) identify communication, equity, interpersonal risk, and reputation in repeated games as the major determinants of cooperative behaviour. Concepts such as reputation and reciprocity, which grow in repeated games, are particularly important indicators for cooperation especially when communication is not a realistic alternative.

Axelrod (1984) was one of the pioneers of research on repeated games, which contributed to the establishment of evolutionary game theory, and vitally inspired research in multi-agent systems. The idea behind repeated games is that of *learning* through indirect communication. The repeated encounters allow the players to acquire knowledge about the other players' strategies by observing their behaviour, and thereafter use this knowledge to guide their own choice of strategies. The evolutionary approach to games also introduced the idea of fitness as an alternative to the idea of utility maximization, and thus offered an alternative to the predominant views of optimisation. Although several limitations have been revealed in response to this approach (especially regarding the efficiency of the tit-for-tat strategy that for Axelrod exemplified an evolutionary stable strategy), the main outcome of this discussion remains that learning is an essential aspect for achieving cooperation in conflicting situations.

Coordination games

Another class of games that is particularly relevant to this research is called *coordination games*. The term refers to games with multiple Nash Equilibria, where the players need to select the appropriate action to avoid reaching a mutually inferior outcome. A typical coordination game is shown in Table 2 for a problem that involves two partners of a building firm who try to decide the location on which to build (note: the game is an example of the 'Battle-of-the-sexes' game). The players have different preferences over the equilibrium outcomes (Partner 1 prefers Location A, and Partner 2 prefers Location B) but they can only buy one of the locations available. The two partners must therefore coordinate their choices of strategies to settle for one of the two Nash Equilibria of the game: (C_1, C_2) or (D_1, D_2).

	Partner 2	
	Location A (C_2)	Location B (D_2)
Partner 1		
Location A (C_1)	100, 50	0, 0
Location B (D_1)	0, 0	50, 100

Table 2 A 2-person coordination game

Although the concept of equilibrium is at the centre of attention in coordination games, it does not provide an answer as to how the players go about choosing between the multiple equilibria (Lucas, 1986). Theoretical and experimental research however strives to identify the parameters involved in coordination success and explain choice behaviour in coordination games. The earlier suggestion that Pareto-dominant outcomes would be selected among the multiple equilibria if indeed such an outcome was available, was proven to be false (Cooper et al, 1990), and consequently alternative criteria such as the notion of risk dominance (Harsanyi and Selten, 1988) have been proposed. As in the case of cooperation, several suggestions for the solution of such games have been made, which include (apart from social conventions and explicit rules that may apply) the existence of focal points (Schelling, 1980); adaptation and learning (Fudenberg and Kreps, 1993; Crawford, 1995; Erev and Roth, 1997; Rappoport et al, 1998; Camerer and Ho, 1999); and the facilitation of communication and reciprocity (Farell, 1987; Cooper et al, 1989). For a more detailed survey of coordination experiments and critical review of the literature refer to Ochs (1995) and Cooper (1999).

From the above discussion we can conclude that in the context of coordination games with multiple equilibria, coordination becomes a rather elusive ability, particularly because the notion of equilibrium does not offer an explanation/prediction of choice behaviour. It can be argued however, that the set of equilibria that can be identified in the games, offers a setting for observing choice behaviour. Consequently, the focus on coordination games seems to move towards the specific variables that affect coordination and more importantly towards the investigation of the role of learning and adaptation.

It is also worth noting that in trying to identify the difference between cooperative and coordination games, or the difference between cooperation and coordination for that matter, one must focus on identifying the underlying assumptions and principles of the 'game'. It seems that cooperation suggests a process that involves some kind of optimisation of joint profits based on the reconciliation of individual and group preferences. Coordination on the other hand, does not aim at optimisation per se, but it rather aims at reaching an equilibrium state. This however does not necessarily mean to undermine efficiency. Coordination simply suggests an alternative methodological/research focus; it suggests a shift from looking at how to adapt preferences, to looking at how to adapt expectations (Camerer and Knez, 1997). This is an interesting approach because it brings into play the necessity not only to investigate and support the effective exchange between players and the development of joint agreements, but also to investigate and support the processes of identification and adaptation of common goals. In any case however, a strong link between cooperation/coordination and learning can be detected.

Before we conclude with game theory, it is useful to note that since it has primarily been developed as a mathematical and economic theory, its main contribution to urban planning has been in the domains of public policy, urban economics etc., but there exist notable applications in urban design (Batty, 1977, 1996) and facility location and public goods planning (Hopkins, 1981). However, the particular value of this paradigm is that it has shaped the foundation for the development of formal approaches in group decision making. Moreover, it has brought into light an explicit preoccupation with problems of conflict, cooperation, negotiation, bargaining, and collective choice. Such influences will be recognised in the subsequent discussions of societal decision making and action in human and artificial worlds.

2.1.2 Coordination in organisational decision making

Organisational decision making is a special case of group decision making. The main distinction between groups and organisations is that in the latter there exists division of labour among the participants, or differentiation of responsibilities for different parameters of the group choice (Ackoff and Emery, 1972: 18-19). Kleindorfer et al (1993: 290) identify hierarchy, systematisation of procedures, interrelated activities, specialization of tasks, and overall complexity, as some of the distinguishing features of organisations. It follows that decision making in organisations is strongly related to effective *allocation of resources* and *management of activities*. Additionally, organisational decision making usually assumes the *existence of a shared common goal* whose achievement depends on decisions taken at various levels, and on the fulfilment of various sub-goals at different stages. However, the existence of a common goal does not imply that this goal is necessarily predefined, nor that the participants know or agree on how to achieve it (for a discussion see Camerer and Knez, 1997). Organisational decision making is therefore directly linked to the need for coordination of individual choices and decisions, as well as individual and group goals.

2.1.2.1 Mechanisms of coordination in organisations

In a discussion of coordination in terms of economic behaviour and activity in organisations, Grandori (2001) distinguishes four basic categories of mechanisms, shown in Table 3 (adapted from Grandori: 97).

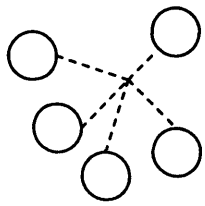
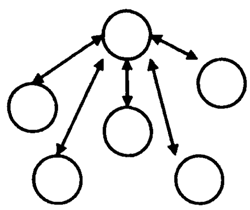
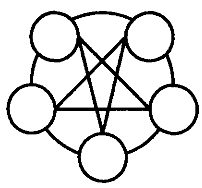
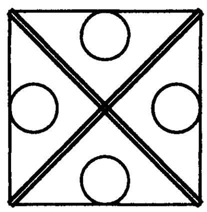
Pricing and voting	Authority and agency	Teams and negotiations	Norms and rules
			
▪ Local knowledge	▪ Bilateral communication	▪ Multilateral communication	▪ Common knowledge
▪ Common codified information	▪ Transfer of decision rights	▪ Joint decisions	▪ Unilateral non-calculative decision
▪ Unilateral decision			

Table 3 Classification of coordination mechanisms by Grandori (2001). Circles represent agents, or decision making actors, while lines represent communication between agents. Dashed lines denote no direct communication; arrows denote bilateral communication; and solid lines denote multilateral communication.

The first category includes mechanisms of *pricing and voting* where coordination can be achieved by unilateral decision making, without direct communication between parties. Such mechanisms draw on basic principles of game theory and collective choice. The second category incorporates *authority and agency* relationships where actors coordinate themselves on the basis of sharing or trading of decision rights. This requires some (bilateral) communication among actors and allocation of decision rights in an asymmetrical (hierarchical) way. The third category includes mechanisms for direct and reciprocal adjustment among actors as in *teams and negotiations*. This involves multilateral communication, knowledge sharing and integration, reciprocal control, and joint decision making. Finally, the fourth category includes mechanisms which guide behaviour without requiring ad hoc decision making, where coordination is achieved in response to conventions, *rules and norms*. These are prescribed structures or models of behaviour acceptable by the actors involved, which guide coordination in a non-calculative fashion.

In this view the *nature of knowledge* available (local or common, concentrated or distributed etc) and the *degree and mode of communication and information exchange* (tacit or explicit, unilateral or multilateral etc), are identified as major parameters for the development and subsequent evaluation of effectiveness of the various coordination mechanisms. Moreover, different configurations of these parameters also disclose different degrees of cooperative behaviour and joint decision making and highlight different structures of social interaction. This classification can easily be generalised to describe research in design, although organisational approaches to collaborative design and Computer Supported Cooperative Work (CSCW) have chiefly emphasised communication as the most fundamental aspect for the coordination of design information and tasks (Chiu, 2002; Cumming, 2002b).

2.1.2.2 Organisational learning

Another important aspect related to organisational effectiveness more generally is the idea of *organisational learning* (March and Olsen, 1975; Argyris and Schön, 1978; Kolb, 1984; Senge, 1990; Simon, 1991; Cohen and Sproull, 1996; Starkey, 1996). Learning and adaptation are highly desirable for organisations to be able to respond effectively to changes occurring both internally and externally. We have already briefly encountered learning as an important feature of cooperative behaviour in groups. Organisational learning adds an important parameter into our investigation of coordination: the need to study the relationship between learning and coordination at the global as well as the individual level. This includes investigating learning as an individual and/or an organisational ability, and particularly investigating the forms and targets of learning at various levels and their interrelationships. Moreover, organisational research investigates the link between learning and creativity in organisations, which is invaluable for design and planning research. We will explore these issues later in the thesis.

2.1.2.3 Organisations as complex systems

It is worth mentioning here that the discussion of learning and adaptation in organisations is strongly related to a view of organisations as *complex systems*. Building on the work of Simon (1957) and Newell (1990) and based on a view of organisations as complex information processing systems, several researchers have introduced a computational approach to organisations and organisational behaviour with a special emphasis on decision making, learning and adaptation (Carley and Prietula, 1994; Prietula, Carley and Gasser, 1998). 'Computational organisation theory' has brought together organisational, cognitive and social science and artificial intelligence and has grown to represent a distinctive approach within the multi-agent systems (MAS) community.

There are two key features that characterise this strand of research: the understanding of organisations as computational entities (agents), and the use of simulation for experimentation, theory building and testing. Naturally, the perception of organisations as sets of (natural and artificial) interacting agents leads to a fundamental preoccupation with coordination. On the other hand, organisations are often viewed as agents themselves with similar adaptive abilities. In a recent review of the 1998 book edited by Prietula et al, Conte (2003) suggests that this dual attribute of organisations (as collection of agents, and as entities themselves) needs to be further investigated particularly by examining 'the interrelationships between the properties that individuals have on their own account and those that they derive from the organisations', but also examining how these properties recur at the entity (global) level and how they are inherited by the individuals. This signifies a general preoccupation with the micro-macro level relationship in social systems, as we shall discuss in the multi-agent systems section below.

2.1.2.4 Organisational decision making approaches in planning

Such organisational approaches to decision making have naturally been influential on urban planning. For example, Lai (Lai and Tang, 1995; Lai, 1998; Lai 2006) uses the 'garbage can theory' proposed by Cohen, March and Olsen (1972) in order to model planning and urban development processes. The idea behind the original garbage can model is that organisations can be seen as systems in which people, problems, solutions and decision situations interact in random or highly unpredictable ways. By transferring this to planning and by adding a spatial element into the equation, Lai aims to offer a model of how actions and decisions become coordinated, and how organisation emerges out of this random interaction of elements. Another example is Xiang's coordination model. Xiang (1993) discusses coordination in distributed problem solving environments, and particularly in the context of environmental planning, with the ultimate aim of developing organisational decision support systems 'to support and augment coordinations among individual problem solvers such that they can operate independently but coherently' (ibid: 15). Drawing from Durfee et al (1990) and Suchman (1987), he argues that in distributed problem solving, apart from the need to gain shared understanding of goals and objectives, there is also a need for *plan coordination*. This is crucial for social policymaking where the consideration of multiple alternative plans and scenarios is an imperative. Xiang actually employs mathematical multiobjective programming methods usually utilised for generation of alternatives, as a mechanism to initialise the process of coordination (Xiang, 1993: 214).

Both approaches constitute invaluable contributions to the understanding of coordination in the context of urban planning and signify a need to revisit the notion of plan to reflect on the complex and multidimensional nature of decision making in this domain. However, it must also be noted that these approaches largely see coordination as a multicriteria or multiobjective optimisation process, and do not fairly take into account processes such as goal adaptation, problem reformulation or co-evolution. This is in part owing to the underlining assumption of organisational decision making that there exists a common or shared goal the individuals benignly wish to pursue. However, it can be argued that this approach is biased towards cooperation; something that in multiparticipatory planning in particular cannot be taken for granted. A multi-agent approach would require considering additional processes and mechanisms that allow the adaptation and reformulation of choices and decisions on various levels.

2.1.3 Coordination in societal decision making

Societal decision making involves the participation of individuals, groups and organisations, in choices and decisions that have consequences for society at large. Societal decision making is indeed a necessity for the resolution of problems that affect people's life and well-being; problems that have to do, for example, with the provision of public goods (provision of public facilities, or allocation and distribution of resources) or the location of hazardous facilities (such as nuclear waste disposal facilities), etc. Such problems are in many cases subject to processes of public or collective choice. Although processes of democratic or collective choice have not been regarded as particularly relevant in building design, most of the approaches developed to study issues of collaboration in design have been built around the understanding of design as a social process. Therefore, both collective choice and collaborative design will be briefly examined in this section.

2.1.3.1 Collective choice and consensus

Collective choice is concerned with group choice situations where the main problem is to define 'fair' and 'democratic' ways of amalgamating individual choices into a single group decision. Much of this work was initiated by Arrow (1951) who interpreted the problem as a 'question of 'combining' individual preference patterns over various states of affairs to generate a single preference pattern for the society composed of these individuals' (Luce and Raiffa, 1957: 328). In other words, given a set of alternatives that for each individual are ordered according to his/her preferences, the problem of collective choice lies in *aggregating* those *preferences* to decide on a single alternative. Typically, democratic voting procedures are employed to resolve conflict and disagreement among the group members. Finding such procedures for the coordination of a common choice, essentially involves assigning an appropriate *social welfare function* to represent the 'rule which associates to each profile of preference *orderings*, a preference ordering for the society itself' (ibid: 332 - emphasis added).

The difficulty with social choice problems is that various processes and methods used to define social welfare functions do not satisfy certain democratic or socially desirable criteria. According to Arrow's Impossibility Theorem (1951), no such function exists that can satisfy four critical axioms/conditions which ideally define fair and democratic choices - given a number of more than two alternatives. Briefly, these conditions are: 1) the condition of 'unrestricted domain', which states that all possible combinations of individual orderings must be included in the social welfare function, 2) the condition of the 'weak Pareto principle', which states that if one alternative A is preferred to another alternative B by every individual, then the society must also prefer A to B, 3) the condition of 'independence of irrelevant alternatives', which states that social choice over a set of alternatives must depend only on those alternative and not on anything else and, 4) the condition of 'non-dictatorship', which states that there must be no individual whose preferences dictate the social order (note that this account follows Sen, 1979).

Several ways have been suggested to escape Arrow's Theorem, by relaxing one or more of the requirements, but a comprehensive review would go beyond the scope of this discussion (for an extensive discussion on the subject see Luce and Raiffa, 1957: 340-345). From the various approaches, one that is worth mentioning is the one that suggests relaxing the third condition. The rejection of the axiom of independence of irrelevant alternatives allows for interpersonal comparisons to be made on the basis of individual *strengths of preferences*. This line of work influenced the development of Multi-criteria Decision Making (MCDM) approaches (see Keeney and Raiffa, 1976; Zeleny, 1982), which further provided ways to formalise public choice problems with respect to multiple criteria, objectives and constraints, and gave impetus to the development of Decision Support Systems. As a general rule, such formal approaches to collective choice in planning highlight the importance of the evaluation (Voogd, 1983), and generation of alternatives (Brill et al, 1990) as a basis for the choice process. For more information on MCDM approaches in planning and in relation to public facility problems the reader is referred to Massam (1988, 1993).

From this brief examination of collective choice, we may conclude that although the term coordination is not explicit in the definition of such problems, the idea of aggregating or combining alternatives to form a single choice is an intuitively valid aspect of coordination, particularly as it involves evaluation and ordering of goals and preferences to ensure resolution of conflict and collective agreement.

Consensus planning

Consensus planning can be seen as the counterpart of collective choice where issues of participation and social interaction are emphasised. Woltjer (2000) provides a comprehensive review of the field by considering three dominant paradigms which respectively view consensus planning as: 1) a process of collaboration and learning 2) a process of bargaining and negotiation and, 3) a process of persuasion and will shaping. He concludes that the three approaches share some common general characteristics, although they differ in respect to their particular focuses (ibid: 40-41). The common ground behind these approaches is that 'reality is constructed' within the social context in which the participants meet together; there are multiple forms of understanding that the individuals bring to this process; and decisions must be taken with emphasis on the dynamics of participation rather than as a product of top-down control. Notably, these characteristics are compatible with the key dimensions of collaborative design, as we shall see below. They emphasise the importance of interaction and learning as a means for constructing common understanding of problems and goals, and for shaping the knowledge required for decision making.

It should be noted however, that in the practice of both collective choice and consensus, choices are often given in advance and therefore the activity of generating alternative solutions precedes (and for this reason sometimes stifles) the process of collective agreement. To capture the creative character of multi-agent design we need to adopt a notion of coordination that incorporates the generation of new choices and alternatives as a part of collective decision making.

The idea of distribution is a key to understanding societal decision making in these terms. It stresses the point that there is no central source of complete information and knowledge, and hence *no central control* over the process or the outcome of decision making. It further implies that collective agreement cannot be modelled in advance but should be considered to emerge through a continuous process of interaction, adaptation and learning. This observation is very important for our definition of coordination. It suggests that if coordination of knowledge and decisions is to be considered central for multi-agent design and planning, then we should also endeavour to understand it as an *emergent* property.

Coordination and emergence

This view is also endorsed by Innes and Booher (2000) who suggest that the development of community indicators can be seen as a process of coordination of individual knowledge and decisions, built from the bottom up via individual learning and adaptation. However, although the link between coordination and emergence in distributed systems is discussed in complexity science, artificial life and multi-agent systems (Theraulaz and Bonabeau, 1999; Ossowski, 1999), this issue has remained undeveloped in studies of group decision making in urban planning. Yet, it is suggested that bottom-up or emergent coordination is important for multi-agent design in general, particularly because it can serve as a catalyst for creativity (Cumming, 2002a; Zamenopoulos and Alexiou, 2002, 2003a).

2.1.3.2 Collaboration in design

The realisation that group processes are vital in most design projects has led to an increasing interest in collaborative design. Studies of collaboration in creative design domains, such as in product design, architectural design, or in the AEC industry, follow two streams of investigation, often in conjunction with each other. The first stream relates to studies of design models and methods. These are usually coupled with empirical studies of designers at work, which aim to identify the underlining characteristics and differences between individual and group design, as well as links between teamwork and creativity. The second stream of investigation relates to the study of collaboration through the prism of collaborative systems, Group Decision Support Systems (GDSS) and Computer Supported Cooperative Work (CSCW). These studies focus on questions of how technology can be used to support group decision making and facilitate collaboration. Although reference will be made to GDSS and CSCW, the focus here will be on the first stream of investigation, design studies, as they are more directly concerned with processes of decision making and in a sense precede research on computational constructs and tools for collaboration.

Paradigms for understanding collaborative activity

Kvan (1999) argues that there are three paradigms that consider (architectural) design as a collaborative activity, namely: the view of design as a cognitive act; the view of design as a situated act; and the view of design as a social act. In fact, the cognitive view is the predominant approach in design studies and it is fair to say that even the situated and the social views are descendants of this approach which takes cognitive acts and processes as the focal point for the study of design.

Let us briefly see how those approaches are linked to each other. The 'information-processing theory' introduced by Newell, Shaw and Simon (1957) established the idea of design as a rational problem-solving process which can be studied and modelled formally. This was one of the first attempts to investigate the cognitive foundations of design and formed the basis for much existing knowledge and methodologies. Another decisive view of design that grew from this paradigm was Akin's approach (1986), which emphasized psychological aspects of problem-solving behaviour. This approach promoted the value of heuristic reasoning and further developed issues related to representation of knowledge. Often cited as the opposite to the problem solving paradigm in design, the 'reflection-in-action' approach introduced by Schön (1983), offered an alternative theory aiming to reconcile professional practice and reasoning. Schön saw design as a 'reflective conversation with the situation' (ibid: 76) meaning that designers in their complex everyday practice need to act in response to the situation they are in and reformulate their understanding of the design problem in order to solve it. It can be said that one of the major contributions of this approach was the recognition that design is an interactive process that takes place within a social environment. With the rising interest in Human-Computer Interaction (HCI) and CSCW in the early '90s, designers once again turned to cognitive theories; this time embracing approaches that called attention to issues of social action and learning, such as activity theory (Vygotsky, 1978), situated action and learning (Suchman, 1987; Lave and Wenger, 1991; Clancey, 1997) and distributed cognition (Hutchins, 1991; 1995).

It must be noted that the various approaches briefly listed above influenced the study of design both within the 'design methods' community and within the 'artificial intelligence in design' community. For example, the perception of design as a social process gave rise to ethnographic studies of design teams (Bucciarelli, 1988), while the situated cognition paradigm inspired the development of formal models of designing and (computational) design agents (Gero, 1998; Gero and Fujii, 2000). Although there are valid arguments to support one or another view of collaborative design, our focus here will be instead on highlighting and establishing the key dimensions that lie behind the various approaches adopted.

Common threads

The first important dimension of group design decision making is that *knowledge is distributed* between the members of the design group. Kvan (1999: 62) argues: 'The knowledge required to produce an architectural design lies beyond the realm of one individual. Deriving a solution to an architectural design problem draws upon the collective knowledge of the group, the team, gathered for the project'. Note that this 'social network of knowledge' described by Kvan, includes human and artificial agents alike, so long as they share knowledge and join in the problem-solving tasks. Along the same lines, Fischer (1999) also emphasizes that the stakeholders who participate in design have equally limited knowledge that, when brought into a collaborative process, offers an opportunity for creativity (refer to the 'symmetry of ignorance' concept mentioned above). The notion of distribution is of paramount importance here, and must not be confused with the fact that the individual members of the group may be geographically far apart. Distribution implies that the nature of design is such that the knowledge and capabilities of the group do not derive simply from the summation of individual knowledge. Complex interactions developed between human and artificial members, adaptation, and learning, give rise to collective knowledge and 'cognitive properties that are not predictable from a knowledge of the properties of the individuals in the group' (Hutchins, 1995: xiii).

The second dimension is that group design is increasingly perceived as a *social process*. In group design, interactions take place because complex *interdependencies* evolve between individual activities and decisions, which naturally give rise to *conflicts*. This epitomizes the social character of group design: social behaviour develops from the need to resolve conflicts by taking part in a collaborative endeavour. Observations of design teams at work identified the occurrence of various forms of social interaction, such as negotiation, persuasion and compromise; and compliance with social norms, power structures, roles etc (Branki, Edmonds and Jones, 1992; Cross and Cross, 1995). However, it should again be emphasised here that, in reality, interactions evolve in - and between - human and artificial agent societies alike: computer models, tools and software agents are equally important carriers of individual knowledge and expertise. Therefore, our view of design as a social process should account for the fact that design takes place in a *socio-technical* context. Apart from the issues of conflict and social interaction, the social views of design also highlight another important issue that derives from the distributed nature of group decision making. The recognition that the individual decision makers come into the design process by contributing their expert (domain) knowledge, also suggests that they consequently maintain different ways of understanding the design problem. Developing a common understanding of the problem (problem formulation-reformulation), and resolving conflicts that arise due to conflicting viewpoints and goals, is considered to be of primary importance if some shared outcome is to be achieved (Cross and Cross, 1995; Kalay, 1998). This also implies that individual and common objectives need to be simultaneously pursued and adapted. The social aspect of design is of critical importance and will be discussed again later in this chapter, in relation to research and applications of multi-agent systems.

The third dimension of group design decision making is that *creativity is a collective attribute*. The question of creativity is the bone of contention between researchers in design studies who have been exploring questions related to the nature of creativity, the conditions that foster creative design and the features of individual and social creativity. Much of this research is reported in two influential series of conferences: the Creativity and Cognition Conference, organised by the Creativity and Cognition Research Studios of Loughborough University, UK; and the Computational Models in Creative Design Conference, organised by the Key Centre of Design Computing and Cognition of the Sydney University, AU. However, situating design within a social context means that both individual and collective creativity draws on collective knowledge, which is created by groups of people interacting with each other and with tools and artefacts (Fischer, 1999). As mentioned above, collective design becomes in many cases a process of conflict resolution or compromise. However, this is not seen as an impediment to creativity; it may lead to thorough exploration or extension of the problem and solution spaces and bring into light creative solutions that could not have been considered by an individual working in isolation.

Finally, it must be noted that effective collaboration in group decision making is often linked with *communication*. Communication is generally acknowledged as an essential aspect of collaboration because it supports the exchange of information, knowledge and ideas, and facilitates detection and understanding of problems and conflicts. There are various arguments that strive to identify the best modes of communication: synchronous or asynchronous, mediated or face-to-face etc. However, communication in the present thesis will be considered in a generic way, as a means to establish a common ground upon which social interactions are built. In this sense indirect and tacit modes of communication will be considered equally important. Such indirect forms of communication can be established for example with the use of external representations such as sketches and drawings, which do not only work as a form of external memory, but also constitute a basis for further exploration, re-evaluation and re-interpretation (Schön and Wiggins, 1992), or even invention of new concepts and requirements (Suwa, Gero and Purcell, 1999).

Note that the role of external representations and other design artefacts and tools for the development of communication and collaboration is a central issue in many studies in CSCW. These studies have their theoretical and methodological roots in theories of situated and distributed cognition as well as in ethnographic studies. Drawing on this tradition, Perry and Sanderson (1998), discuss the essential role of design artefacts for the representation and transformation of knowledge in joint design work, but also further point out their importance for the coordination of individual activities and the development of social interactions. An extended review and discussion on the role of external representations and artefacts in design can be found in Perry (1997) and Perry and Sanderson (1998).

Yet, concepts of indirect communication external to cognitive science research could also present mechanisms for the development of collaborative activity and coordination. For example, Susi and Zimke (2001) argue that the concept of *stigmergy*, which explains how the behaviour of a group of social insects can be coordinated by indirect interaction (through modification of the physical environment), has some similarities with the aforementioned cognitive theories and could offer valuable insight into the study of social interactions.

The four dimensions of group design elaborated above can help us elucidate the meaning of cooperation, coordination and collaboration, which come into view as three subtly - but crucially - different manifestations of multi-agent decision making. Let us first look at some existing comparative definitions.

Comparing cooperation, collaboration and coordination

Branki (1994) argues that groupwork involves all the three processes: cooperation generally refers to a process of working together to develop a design artefact, while collaboration involves more specifically sharing, exchanging and maintaining information to create a common pool of knowledge and build shared goals. Coordination on the other hand is a concept more related to ordering of processes to ensure harmonious work, and may have four components: goals, activities, actors and interdependencies (ibid: 37). Note that this definition is resonant with the definition offered by Malone and Crowston (1990: 360) in coordination theory and is particularly akin to engineering approaches to coordination. Similarly, Perry (1997) defines coordination as the means by which the distribution of labour is achieved (and may be emergent or arise through managerial action); while Dave (1998) relates coordinative action to a process of managing tasks through the development of plans and schedules. Additionally, Gross et al (1998) associate coordination with defining protocols for communication and agreeing on the allocation of responsibilities. Finally, Kvan (2000) - drawing largely from Mattessich and Monsey (1992) - offers a more comprehensive account of the three concepts. In his view, cooperation is a more low-level concept used to describe informal relationships between the members of a group and does not imply any commonly defined mission or goal. Collaboration on the other hand is seen as joint problem solving and therefore requires a higher sense of working together and greater commitment to a common goal. In between, coordination is characterized by some formal relationships and understanding of compatible missions, but does not imply full commitment to common goals (Kvan, 2000: 410-411). It follows that group design, although it can be defined as a social act, it is not necessarily collaborative.

The position adopted in the present study is that in contrast to cooperation, coordination is an abstraction that matches well with the view of design as a social process, but unlike collaboration, is generic enough to capture a large class of design problems. The main issue captured by the concept of coordination is that decisions and actions of individuals that participate in a group (human or artificial agents) are interdependent. This necessarily leads to the development of some (weak) form of social structure that sets the scene for the effective management of interdependencies, resolution of conflicts, and construction of collective knowledge. Furthermore, the concept of coordination allows us to include, in our understanding of group design, processes where there are no shared goals, and collaboration is not a prerequisite for the formation of design solutions (see the concept of unstructured collaboration discussed by Craig and Zimring, 2000).

2.1.4 Summary

Multi-person decision making is an important aspect of design and planning and constitutes a significant area of investigation for the development of computational models and tools to mediate and support design and planning processes. Group, organisational and societal approaches to decision making were briefly reviewed in order to identify the role of coordination and its relation to issues such as cooperation, collaboration, conflict resolution, collective choice, and consensus. In turn, various aspects were uncovered viewing coordination as coordination of goals, choices, preferences, and expectations, or aggregation of choices, but also as coordination of actions, distribution and allocation of resources etc.

2.1.4.1 Coordination in multi-person decision making

We first examined Ackoff and Emery's view of planning as purposeful group choice, where processes of interactive identification, pursuit and coordination of individual and collective goals were assumed to achieve cooperation. We also saw that the view of planning as purposeful decision making shared common characteristics with design, where pursuit and identification of goals is tied with a need to take a co-evolutionary approach to problem making and problem solving.

On the other hand, we saw that game theoretic perspectives viewed cooperation as joint utility maximization, highlighting instead the need for coordination of individual choices and preferences. Game theory introduced into planning literature the fundamentals of social dilemmas associated with location of public goods and resources and offered formal tools for studying processes of negotiation, bargaining, conflict resolution etc. A brief discussion of the relationship between cooperation and coordination revealed both similarities and differences in terms of underlying assumptions, attributes and mechanisms. For example, although there could be identified a common concern with reviewing rationality assumptions and detecting the factors that influence optimum behaviour (e.g. communication, the use of norms and rules, etc), coordination games were distinctly focused on the investigation of equilibrium as a potential quantitative measure or predictor of coordinated behaviour and the subsequent proposition of alternative mechanisms for adaptation (coordination of expectations).

Additionally, we saw that similarly with Ackoff and Emery and the approach of cooperative games, organisational decision making approaches adopted the view of cooperation as a process towards fulfilling a common purpose. Here, decision making was more directly linked with the organisation of actions and tasks, and hence coordination was seen also in relation to effective allocation of resources and management of activities. However, we also saw that although organisational decision making introduced explicitly the notion of coordination in urban planning, relevant applications over-focused on issues of defining options for planning, coordinating actions, or even generating alternatives, therefore leaving out issues of goal adaptation and co-evolution.

In the discussion of societal decision making, the focus for the investigation of coordination shifted to issues of collective choice and consensus. In collective choice, coordination was largely identified with the development of a social welfare function that defines an ordering of choice preferences for the society by aggregating individual preferences. Although this approach does not capture the totality of processes involved in multi-participatory design and planning, it adds to our understanding of coordination the crucial facet of evaluation of alternatives. Issues of participation and social interaction are instead more directly linked with the ideas of consensus planning. We saw that consensus planning emphasised the need for constructing a common understanding of problems and goals, and for shaping the knowledge required for decision making from the bottom up. This additionally introduced the idea of emergence as a crucial attribute of coordination. Table 4 provides a summary of main points relating multi-person decision making and coordination that were exposed through this discussion.

Group decision making and game theory	Goal oriented activity Coordination is seen as a concept between cooperation and conflict that reconciles individual and group rationality Focus on individual learning Importance of adaptation of preferences and expectations Notion of equilibrium
Organisational decision making	Assumes shared goal Coordination also includes goal generation and adaptation Focus on individual as well as collective learning Notions of management and allocation of resources, division of labour, and authority
Societal decision making and collective choice	Assumes multiple objectives Coordination related to concept of emergence Importance of generation and evaluation of alternatives Notion of amalgamating preferences, notion of consensus

Table 4 A summary of the main aspects of coordination in group, organisational and societal decision making

2.1.4.2 Coordination from the perspective of design and planning

Despite differences in the aforementioned approaches, certain common features were identified which become crucial not only for the understanding of coordination itself but also for the understanding of how coordination is linked to design and planning processes. The first feature identified was that multi-person decision making in design and planning does indicate the existence of distributed or *shared resources*, tasks and responsibilities, but also the existence of *distributed knowledge*. This feature establishes a view of coordination as a product of decentralised control over resources, processes and activities, but it also explicitly links coordination with a constructive process of collective knowledge.

The second feature of multi-person design revealed from this review puts forward *adaptation* and *learning* as key mechanisms for the construction of this knowledge. Although there are various approaches to how such abilities can be achieved (for example with or without communication), they all agree on the dependence of adaptation and learning on the basis of interaction and feedback from the environment.

The third commonly identified feature among the various approaches reviewed, was that design and planning processes are *goal-oriented*; coordination can therefore be identified with a process of individual and collective goal identification, pursuit and adaptation. As goal finding and adaptation is directly linked with problem finding and reformulation, the notion of *co-evolution* of problem and solution was also strongly emphasised.

Finally, multi-agent design was related to the notion of *creativity* as this is developed within a social setting. Coordination in this sense was seen as a product of collective knowledge created through processes of social interaction.

All these features should be taken into account in the development of a valuable definition and model of coordination in the domains of design and planning. Although some work was presented which seeks to define coordination particularly in relation/contrast to cooperation and collaboration, no work exists which effectively takes into account all the aforementioned features. However, the review of these approaches helped us identify the distinguishing characteristics of coordination, where social interaction is emphasised without compliance to (benevolent) collaborative behaviour.

Before we continue with the review of coordination in other fields, it is worth noting that the above review was slightly biased towards approaches that formalise coordination as a uniformly distributed process. For example, approaches such as the principal agent approach or the unitary actor model found in group and organisational studies have not been reviewed. Although these approaches may be more appropriate in certain conditions, the idea of homogeneous (non-hierarchical) control captures more effectively the idea of distribution as it was discussed above, and provides a more generic framework for the study of coordination.

It should also be mentioned that the review was focussed on (descriptive and prescriptive) theories of multi-person decision making rather than specific methods and practices. Consequently, the thesis only mentions briefly a few tools and methods for collaborative decision making (e.g. multi-criteria optimisation) and does not discuss extensively the different existing techniques for facilitation and group decision support, like for example, brainstorming (Osborn, 1963), the 'Six Thinking Hats' (de Bono, 1985), or the Delphi method (Turoff and Linstone, 1975). The thesis also does not review processes and methods for community participation in design and planning. For an overview on the subject see for instance (Sanoff, 2000, 2007; Wates, 2000).

2.2 ARTIFICIAL INTELLIGENCE: DESIGN AS THE SCIENCE OF THE ARTIFICIAL

Artificial Intelligence (AI) is defined by John McCarthy (2002), who actually coined the term in 1955, as 'the science and engineering of making intelligent machines, especially intelligent computer programmes'. The history of AI is quite long (some people trace its origins back to Aristotle and the ancient Greek automata), and so is the list of definitions, interpretations and research paradigms within the field. However, research in AI has generally aimed to fulfil two purposes: first, to construct intelligent computer models in order to understand animal (and particularly human) behaviour, and second, to emulate intelligence on computing machines, so as to complement or augment human capabilities in solving problems and performing complex tasks. Naturally, artificial intelligence is a multidisciplinary research field and draws from studies in diverse areas such as philosophy, cognitive sciences, computer science, organisational sciences etc. For comprehensive accounts of AI the reader is directed to Boden (1996) and Russell and Norvig (1995).

The reason why this field is relevant to the research reported here is that, in accordance with Simon's introduction to The Sciences of the Artificial (1969), artificial intelligence has been increasingly perceived as a science of design.

Simon not only established the view of design as a process that can be modelled and performed computationally, but more importantly he also established the view of design as the core of the sciences of the artificial. It is in this sense that devising intelligent machines becomes a tool for both the understanding and the devising of our world. In a similar vein, Winograd and Flores (1986: 4) perceive the question of design, as 'the interaction between understanding and creation': a discourse between understanding our environment (and ourselves) and creating tools to re-interpret and change this environment.

Although artificial intelligence forms a general background for the present research, the study here will focus on a sub-field of AI which engages in the study of groups of intelligent agents, namely Distributed Artificial Intelligence (DAI). This opens up the present review to consider more formally aspects of coordination in artificial societies ('societies' of computational agents). It is useful to note that although coordination as *'the act of managing interdependencies between activities performed to achieve a goal'*(Malone and Crowston, 1990: 361) is generally considered crucial for the understanding and development of artificial societies, the perspectives of coordination and the techniques used to achieve it vary substantially. This is in part attributable to the fact that DAI grew out of two different traditions in artificial intelligence: the connectionist tradition that emphasises issues of learning and parallel distributed processing, and the rationalistic tradition that emphasises issues of symbolic/logical knowledge representation and modularity. In some cases a third influence can be detected, that has its routes in artificial life (AL). This 'reactive' view influenced the understanding and development of autonomous agents by alternatively emphasising issues of agent-environment interaction and adaptation. Next, we will briefly discuss some characteristics of systems composed of multiple agents and their relevance to design and planning tasks, and consequently review fundamental perspectives and mechanisms for coordination.

2.2.1 Distributed artificial intelligence and coordination

The general standpoint in DAI (built on the concept of bounded rationality proposed by Simon (1957)) is that individual agents have limited knowledge and resources, information processing abilities, and viewpoints, and therefore need to engage in some kind of collective activity or interaction in order to improve their problem solving and goal attainment capabilities. It goes without saying that the term distribution is the quintessence of artificial societies: it suggests the existence of incomplete and dispersed information, interdependency of actions and decisions, and lack of global control mechanisms or rules to dictate global behaviour.

2.2.1.1 An overview of DAI approaches

DAI scientists typically distinguish two areas of research (Bond and Gasser, 1988; Moulin and Chaib-Draa, 1996): distributed problem solving and multi-agent systems. For more thorough reviews of the two areas see Durfee et al (1989) and Sycara (1998) respectively. In brief, *Distributed Problem Solving* (DPS) studies how a group of intelligent agents (or nodes) can share their resources and harmonise their activities, so that they can collectively solve a particular problem. According to Ossowski (1999: 39-40), DPS focuses on aspects of *cooperation* within a dynamic problem-solving environment, and typically conforms to three assumptions: the *benevolence* assumption, the *common goal* assumption, and the *homogeneous* agents assumption. This means that agents are in principle predisposed to cooperation and aware that they pursue a common goal, therefore conflicts only arise due to limited viewpoints and not to conflicting goals or interests. On the other hand, research in *Multi-Agent Systems* (MAS) is concerned with a loosely coupled network of agents who may work together to solve a problem. The focus therefore is on the *interactions* between agents, and the main assumptions are: *non-benevolence, multiple goals, autonomy* and *heterogeneity* (Ossowski, 1999: 43-44). In other words individual agents are assumed to act based on individual rationality (see the discussion of game theory), which suggests that they maintain multiple, partially conflicting goals and interests. Despite their 'self-interested' nature, agents may wish to interact with each other so as to perform activities that assist them in the achievement of their own objectives (Lesser, 1999). Heterogeneity suggests that other kinds of conflicts may also arise due to incompatible communication protocols, or agent architectures.

The MAS view offers a more generic perspective on distributed artificial intelligence and in recent years has been used in an inclusive sense to refer to the study, modelling and construction of artificial societies (for more recent overviews of MAS see Ferber, 1999; Weiss, 1999; and Wooldridge, 2002). The distinction between DPS and MAS however, offers a useful view of the processes found in human design and planning groups (Alexiou and Zamenopoulos, 2001). Distributed problem-solving is inherent in situations where different experts work collaboratively on a project, and therefore need to combine their expertise, languages and viewpoints, so as to reach an optimum solution. The problem/goal of the design or planning process is defined in common, but conflicts arise due to conflicting partial requirements, distributed or shared resources, and shared routines. This situation is typical in design firms (Coyne et al, 1996) or in emergency planning organisations (Dickey and Vernon, 1998). In such cases coordination typically reflects the organisational and problem-solving ability of the group. On the other hand, the view of multi-agent systems mirrors situations where different stakeholders take part in a design or planning process so as to promote their individual goals. To this end, interactions are developed as a means to resolve conflicts and establish conditions for mutual benefit. This situation is typical in participatory planning (Innes, 1996; Forester, 1998; Innes and Boocher, 2000) and in unstructured collaborative processes (Craig and Zimring, 2000; Findlay and Haugen, 2000; Van Loon, 2000). In such cases coordination is to a great extent a process equivalent to discovering and promoting common goals. Despite the differences identified in DPS and MAS, it is true in both cases that coordination is an important measurement of coherence and group performance.

2.2.1.2 Coordination techniques

Researchers in the DAI community have formulated various coordination techniques. For some existing classifications and descriptions the reader is referred to Jennings (1996), Sycara (1998), and Ossowski (1999). In his account of coordination Ferber (1999) has identified four fundamental coordination mechanisms: synchronization, planning, reactivity and regulation. *Synchronization* refers roughly speaking to a process of chaining actions so that they can be performed simultaneously at a certain time; *planning* refers to determining a set of actions to be carried out so as to achieve an overall goal and involves generation, coordination, choice and re-evaluation of plans; *reactivity* refers to a coordination process built by agents reacting and adapting to changes that occur in their environment and in relation to actions taken by other agents; and *regulation* refers to the use of social laws or conventions to ensure coordination and conflict avoidance. This account offers quite a wide perspective on coordination approaches and mechanisms particularly because it takes into account both the rational/deliberative and the reactive/behaviour-based traditions in DAI. The review here will focus more on identifying current trends in multi-agent coordination and investigating research issues and questions relevant to the study of design in the context of architecture and urban planning.

2.2.2 Artificial societies and social agents

A first issue that stands out in current research is the *social aspect* of MAS. This not only includes defining or achieving social behaviour within groups of artificial agents, but also involves investigating the relationships and interactions of agents with their environment in general including potentially with humans. Social reasoning and action in MAS is a generally desirable field of investigation for this thesis, because it contributes to the definition and modelling of design as a social activity. Naturally, there are various approaches to 'sociability' in MAS, which are coupled with different mechanisms for coordination.

2.2.2.1 'Rationalistic' approaches to sociability and coordination

One prominent approach to sociability draws on the tradition of game theory and rational decision making in general. Roughly, game theoretic perspectives consider a social setting where individually rational agents interact with each other in order to define a coordinated outcome on the basis of a known payoff matrix. Wooldridge (2002: 210) likens this to a process of *'coordination by mutual modelling'* whereby agents build models of each other (based on their beliefs, utility functions and so on) to serve as guidelines for coordination. An interesting aspect of this view is that coordination and sociability do not require communication - sociality is expressed on the basis of the interdependencies represented in the payoff matrix. Other approaches have focused on mechanisms of auctions, negotiation, bargaining, coalition formation, or voting (social choice) thereby reinstating concepts of Pareto efficiency, equilibrium, social welfare and so on (see Sandholm, 1999 for a brief summary of 'distributed rational decision making'). A distinctive approach that came out from this line of research has introduced the notion of norms and *social laws* as a means for coordination. This approach emphasises the design of protocols or 'rules of encounter' to determine agent behaviour (Rosenschein and Zlotkin, 1994; Shoham and Tennenholtz, 1992). Such social laws can be designed offline and subsequently used as a set of *constraints* to bound the negotiation process; or can emerge from within the system as a process of *strategy choice* and change (Wooldridge, 2002).

Other researchers have emphasised *intentional cooperation* and joint commitments as a basis for social action. Levesque et al (1990) adopt the view that mental states such as beliefs, desires, and particularly intentions, offer valuable insights into the study of teamwork. Specifically, they discuss the notion of a *'joint persistent goal'* as the glue of social action: being part of a team implies the existence of a common goal and the shared intention (or responsibility) to work cooperatively towards the fulfilment of this goal. Jennings (1993, 1996) has taken this view further to formalise coordination as a *'distributed goal search with multiple loci of control'* where *commitment and convention* are utilized as the key agent structures: 'Commitments are viewed as pledges to undertake a specified course of action, while conventions provide a means of monitoring commitments in changing circumstances'. Put together, commitments and conventions provide a certain degree of predictability for future activity, and a flexibility to cope with dynamically changing environments. In fact, Jennings claims that *'all coordination mechanisms can ultimately be reduced to commitments and their associated (social) conventions'*.

Current research however, has identified limitations in both the 'individualistic' and the 'cooperative' approach to sociability described above. On the one hand, game theory has been criticised as an inappropriate paradigm for the definition of social action in MAS, not only due to its inherent weaknesses (e.g. inability to predict selection of equilibria; failure to account for inconsistent preferences; strong assumptions about availability of resources) but also because it fails to consider *goals* and *motives* for joint activities, which offer content and context for the decision making (Castelfranchi and Conte, 1998). On the other hand, the view of collective action based upon joint persistent goals and intentional cooperation, leaves out the explicit consideration of *conflict* as a form of social interaction (Castelfranchi and Conte, 1996). Starting from such observations, some recent work has stressed the need to devise alternative mechanisms for coordination based on notions of rationality that capture the interdependencies between individual decision making and social action, that is, notions of *social intelligence and rationality* (Jennings and Campos, 1997; Hogg and Jennings, 1997). Nonetheless, a general concern with deliberative approaches to social action is their weakness when considering the *emergence* of social structures or how these *feed back* into the individual agents (Castelfranchi, 1998). Although the above criticisms come from inside the cognitive agents camp, others consider weaknesses of the deliberative approach as a whole and particularly the inefficiency of agents in taking into account dynamically changing environments and responding to changes in a timely way.

2.2.2.2 Reactive approaches to sociability and coordination

The alternative approach to ('rational') sociability in MAS draws on artificial life and the tradition of reactive agents. The reactive agents approach is particularly popular in agent based modelling and simulation circles and generally considers subcognitive agents who do not have an internal representation of their environment but act based on a stimulus-response type of behaviour (Ferber, 1994: 8). The reactive paradigm has strong links with biology and complexity science and focuses on the importance of the environment (particularly the *space*) within which human, animal and artificial societies are situated. In principle, coordination is not specified a priori by way of explicit ontologies or strategies; instead coherent group behaviour and social structures *emerge* from agent-environment and agent-agent interactions developed on the basis of simple patterns of behaviour or rules that pose limited bounds to the individual agents (Epstein and Axtell, 1996). This type of implicit or emergent coordination is normally obtained through processes of *adaptation* and *learning*. Various mechanisms have been used to this end, which range from simple rules to cellular automata (Epstein and Axtell, 1996; Resnick 1994), various evolutionary techniques (Axelrod, 1997; Kennedy, 1997), competing tasks (Drogoul and Ferber, 1992; Ferber, 1994), neural networks (Hutchins and Huzlehurst, 1995; Parisi, Cecconi and Cerini, 1995), stigmergic algorithms (Deneubourg et al, 1991; Bonabeau et al, 1994; Beckers, Holland and Deneubourg, 1994), and so on. Well-placed within the reactive agents tradition is also the 'behaviour-based' robotics route established by Brooks and his colleagues (Brooks, 1986; Maes, 1993; Steels and Brooks, 1993), which has greatly influenced research in collective robotics. This approach additionally emphasises issues of *embodiment* and *situatedness*, arguing for an understanding of intelligent behaviour as a socially and physically embedded activity - vitally linked to the development of interactions between agents and their environment. The roots of this approach can be found in theories such as activity theory, situated action and distributed cognition.

In general, the approach of reactive agents has in its turn raised criticisms, which mainly focus on the observation that the strict dependence of agents on local information may lead to unpredictable and unstable system behaviour (Sycara, 1998: 82). More specifically however, in terms of sociability, there are two arguments that deserve particular emphasis. The first argument is that the reactive approach misses an important aspect of sociality and collective action, that is, the social character of individual action, or otherwise the *individual social action and mind* (Castelfranchi, 1998). Castelfranchi argues that not every (inter-) action is social; instead there needs to be at least some form of cognitive intelligence (understanding of beliefs and goals of other agents) to be able to define and model social action. The second argument is that the *concept of emergence*, although of central importance for the understanding of the relationships between local behaviour, environment, and global behaviour, is neither clearly understood nor formally defined (Wooldridge, 2002: 97). In turn the concept of emergence cannot serve as an indicator of cohesive (coordinated) behaviour or a measurement of system performance in general. Similarly, Epstein (1999) also argues that the lack of a consistent definition of emergence is crucial for the agent-based modelling research because it poses limits to its predictive and explanatory value.

2.2.2.3 Current questions and directions

Looking at the critiques mentioned above, one can discern a general preoccupation with the problem of the micro-macro link in MAS, namely the link between the macro-social system and the individuals (Conte and Castelfranchi, 1995). Recent research has favoured cross-fertilisation between various approaches in MAS and the reconciliation of intentionality and emergence to explore the formation of social structures. Consequently, we can also discern three main research questions: to (re-) consider coordination and sociability in relation to both self-interested and altruistic behaviour by taking equally into account structures of *conflict and cooperation* (Tessier et al, 2001); to investigate the definition and development of *social agents* as autonomous individuals and as members of *social communities* (Dautenhahn, 2000b; Dautenhahn et al, 2002); and to review coordination and emergence in formal ways, particularly by looking into the effects of individual action on the macro-level but also the feedback of social structures to the individuals (Steels, 1993; Gilbert, 1995; Castelfranchi, 1998).

2.2.3 MAS in urban planning

In urban studies multi-agent systems have been used mainly as a simulation tool. Multi-agent modelling in urban development and planning is a relatively new research area that has grown from the tradition of using artificial intelligence and artificial life techniques (mainly cellular automata) to investigate and model spatial dynamics such as land use change, urban growth, population dynamics, traffic etc (for overviews see Besussi and Cecchini, 1996; White and Engelen, 2000). Adopters of the MAS paradigm have used agent technology to capture and represent the dynamics of individual behaviour in cities (Batty et al, 1998; Benenson, 1998; Schelhorn et al, 1999; Bishop and Gimblett, 2000; Torrens, 2001; Batty et al, 2002; Dijkstra et al, 2002). These models have focused on the reproduction, visualisation, and analysis of dynamic behaviour and have served mainly exploratory, descriptive and to a certain extent predictive purposes. So far, the majority of such models have not directly incorporated multi-person problem solving and decision making processes, therefore leaving out issues such as conflict resolution and goal adaptation; but also more importantly for this research, they have not considered processes of collective plan generation. Little work however, has been reported that attempts to deal with some of these issues.

Ligtenberg et al (2001) describe a spatial planning model that simulates land use change as a result of actor (user) decision making, based on a technique that combines MAS and cellular automata. This model considers a process whereby a number of intentional actors interact with the environment and with each other (with simple message passing) in order to define individual preferences for land use allocation. These preferences are then translated into a new configuration through voting and subsequent weighting. Although the configuration of alternative land use plans is a distributed process, coordination is equated to a simple voting rule, which inevitably brings about problems inherent in collective choice. This is detectable in one of the simulations where equal voting power was assigned to the participating actors. In this case the system resulted in an impasse due to unresolved conflicts, which signifies that the process of coordination had not been fully worked out. Nevertheless, the overall approach is noteworthy because it effectively takes into account decision making of multiple actors, by combining spatial processes with agent interactions. Yet, as the authors acknowledge, further work to implement techniques for knowledge construction and learning would greatly enhance the model.

Arentze and Timmermans (2002) propose a multi-agent model for land use (retail) allocation that focuses on negotiation protocols and strategies. Taking a more rationalistic standpoint, this work is also an interesting approach that fruitfully incorporates user (cooperative and competitive) behaviour and decision making for the investigation of macro-level dynamics. Here, coordination is equated to a process of negotiation but issues of learning and feedback would still deserve more attention if the intention to deal with complex behaviours and decisions was developed within such a set of interconnected organisations and authorities.

Finally, Saarlos et al (2001) discuss the role of MAS in the context of planning support systems in general, drawing attention to the current lack of systems that incorporate multidisciplinary expertise essential for the planning process. They further point out several requirements for land use planning systems which in brief include: support for the regulation of *physical* aspects of the built environment; support for the simulation and visualisation of decision making processes as they *evolve in time*; support for the *generation* and exploration of alternatives; provision of *'process-related tools'*; and provision for access to *multidisciplinary knowledge*. The authors see MAS as an essential component for meeting the requirements of process-related tools and multidisciplinary expertise and present a conceptual model developed toward this direction. However, as we have already seen in the discussion above, one can exploit all the powers related to the distributed nature of MAS from their problem solving capability, their adaptability and modularity, to their ability to capture and generate dynamic and emergent phenomena. Moreover, urban planning studies have a lot to benefit from the social views of MAS which focus on the links between the micro-level of agents and the global-level of the agent society, and the investigation of interactions between agents and environment. This would be achieved with a consistent formal view of coordination in multi-agent planning models. Such views should take into account the particular requirements of planning that emphasise not only aspects of consistent collective choice of plans and strategies, but also aspects of adaptation and evolution of goals and objectives, co-evolution of problem and decision spaces, and learning. It is also worth noting that all three paradigms of MAS applications in planning we have seen here, have been in land use planning.

2.2.4 MAS in design

Research on MAS has been pursued within the 'AI in design' community to study design formally, and support it computationally, as a complex multidisciplinary process that involves group decision making and collaborative working. Theoretical and technological advances in this area have gone hand in hand with advances in concurrent engineering[3], CSCW and collaborative CAD. Some

[3] Concurrent engineering is an organisational strategy, a multi-disciplinary design methodology, which aims at the integration of the design, development and manufacturing stages of a product. Although the idea of coordination is central to research in concurrent engineering, delving more into the field is outside the scope of this thesis. Nonetheless, since engineering is for many the

groundwork research is reported in Gero and Maher (1993). Research on MAS in design can be roughly grouped under two main issues or directions of investigation: the development of models of (collaborative) design, and the development of design support systems.

2.2.4.1 MAS for modelling design processes

Models of design may focus on the development of theories of collaborative design, the formalisation of design process, but also the definition and conceptualisation of design agents and their properties.

For example, Campbell et al (1998, 1999) present a theory, called A-Design, which seeks to capture conceptual engineering design as an iterative, evolutionary process where solutions evolve together with goals, constraints and user preferences. A-Design maintains and evolves both populations of designs and populations of agents who create them, and incorporates agent collaboration with an adaptive selection of designs. More specifically, A-Design consists of four basic subsystems: a hierarchical architecture of goal-oriented agents responsible for creating and improving design alternatives; a representation of the conceptual design problem that makes it possible to describe the behaviour of components and instantiate behaviour with real-world components; a scheme for multi-objective optimisation which divides alternative designs based on Pareto optimality to add adaptability and diversity to design states; and, an evaluation-based iterative algorithm for improving basic design concepts towards successful solutions (Campbell et al, 1999:172-173).

In this theory, agents contain knowledge of how to design based on their individual strategies and preferences, essentially implemented as deterministic algorithms, and are divided in three classes: maker-agents, modification-agents and manager-agents. The maker-agents create design alternatives based on an

---

'counterpart' of design, some engineering approaches to MAS will be reported which deal with issues of concurrent or cooperative design. An overview of coordination approaches and systems in MAS from an engineering perspective is offered by Whitfield and et al (2000) and Coates et al (2000) in a two-part paper. Also, for a general overview of MAS in concurrent engineering see Shen at al (2001).

initial problem specification, and are responsible for re-building designs returned by modification-agents; the modification-agents take designs from the evaluation stage and modify them to improve design states; and finally, the manager-agents provide feedback about how to improve the design process according to user preferences, mainly by controlling the population of particular agent types contributing to better designs. Although coordination/collaboration of agents is based on a hierarchical architecture whereby manager-agents (in practice one agent) invoke the operation of the other classes of agents and provide basic criteria for design selection and process convergence, the overall model relies on the concurrent working of agent populations operating in a dynamic and adaptive environment. This view is in agreement with reactive approaches to MAS, which concentrate on complex behaviour emerging from populations of relatively simple agents interacting with each other.

The majority of researchers in MAS interested in models of collaborative design currently endorse the 'weak notion' of agency proposed by Wooldridge and Jennings (1995), which denotes a view of agents as computer systems that exhibit the properties of autonomy, social ability, reactivity, and pro-activeness, and does not include stronger mentalistic notions such as beliefs and intentions.

Brazier et al (2000), suggest that reactivity and social ability are related to agents' ability to interact (communicate) with each other and the material world and often rely on successful acquisition and maintenance of knowledge about other agents and their environment. To this purpose, Brazier and her colleagues focus on knowledge level formalisations in order to model agents and design tasks. These formalisations include components for management of an agent's own processes, interaction with other agents, interaction with the external (material) world, performing specific tasks, but also coordination of the design process and manipulation of requirement qualifications and design object descriptions (Brazier et al, 1996; 2000).

Drawing on these preceding knowledge-level models, Brazier et al (2001) propose an extended agent architecture that focuses on design as a collaborative distributed process. The underlining principle for this model is that, in a distributed design setting, decisions are not only influenced by the specific knowledge available to an agent, but also by the agent's interpretation of his/her own situation. The extended model explicitly accounts for elements that define a situation in a collaborative distributed design setting by including information on design interactions and results from design interactions, but also reflective reasoning at a strategic level and reflective reasoning from the viewpoints of other agents.

While the work of Campbell et al (1999) described above focuses on devising a theory of engineering design, the work of Brazier and colleagues is focused on the development of a generic model for weak agenthood. This model is specifically targeted at questions of agent design, which, as any other process of (creative) design, involves taking into consideration issues of iterative formulation of requirements, goals and object descriptions; interactivity; cooperation, reflective reasoning; and development of shared understanding between multiple decision makers. Coordination in this model refers to processes rather than decisions; it refers to the organisation, management and orderly performance of tasks, and is carried out by individual agents organised in a hierarchy. At any level, coordination relies on communication between agents and management of agent's own process over time.

A final point worth noting is that accordance with the weak notion of agency in Brazier's et al (2001) work suggests also accordance with a situated view of design agents. In artificial intelligence research, notions of situated action have mainly been used to guide the formalisation and synthesis of individual agents anchored in an ongoing interaction with dynamically changing environments (Agre and Chapman, 1987; Agre and Rosenchein, 1996). Similarly, situated views of cognition and intelligence in design have been concerned with the understanding and modelling of design processes and activities by focusing on the relation between agents and their environment. In these views designing has generally been considered as an interactive process that takes place within a cultural and social context.

Gero's model of designing (design process) is developed in order to serve as a framework for the creation of situated design agents (Gero, 1998; 1999; Gero and Kannengieser, 2002). This model is inspired by Clancey's (1997) view of situatedness and considers decisions as 'a function of both the situation [an agent is in], and the way the situation is constructed and interpreted' (Gero, 1998: 52). In this sense, a situation is a projection of a state of the environment within an agent. A determining characteristic of this approach is the notion of 'constructive memory', which suggests that memory it is not only constructed by experience, but is also re-interpreted and re-constructed in the light of the present situation. Designing can in turn be considered as a series of constructive memory acts.

This approach is elaborated by Gero and his colleagues to explore the idea of situatedness in relation to crucial aspects of design, such as creativity, concept formation (Gero and Fujii, 2000), and learning (Reffat and Gero, 1999; Liew and Gero, 2002). Although this model proclaims a view of design agents as being situated within a social context, it too is primarily concerned with the formalisation of single design agents. It can be argued that the model is mostly focused on issues of individual cognition, namely issues of how individual agents interpret and act on the external world by constructing an internal representation of this world, based on memories, experiences, as well as expectations (hypothetical reasoning).

Consideration of the prospect that the environment of an agent may also consist of other agents receives more appropriate attention in Kannengiesser and Gero (2002), where the authors focus on communication as a means of interaction among agents as well as a means of achieving the coordination of activities. In particular they investigate the notion of common ground as an abstraction that can support the view of communication as a situated act. The aim in this case is to develop a framework for the construction of common knowledge and understanding among heterogeneous agents that does not rely on pre-defined static ontologies. Other multi-agent approaches developed within Gero's group adopt more artificial life-like methodologies, particularly to study and formally model aspects of design such as creativity and emergence (Saunders and Gero, 2001; Gero and Sosa, 2002; Saunders and Gero, 2002a; Sosa and Gero, 2002).

Researchers interested in modelling design processes and agents have also focused on issues of conflict management and resolution. Conflict resolution is closely related to coordination and has usually been studied and modelled as a negotiation process. For an overview of negotiation and conflict resolution in concurrent engineering see Shen et al (2001). Klein (Klein and Lu, 1989; Klein, 1991) discusses a model for knowledge-intensive run-time conflict resolution based on a taxonomy of conflicts and a proposed representation of conflict resolution knowledge alongside domain and control knowledge.

Another approach to conflict resolution in concurrent engineering design is suggested by the development of Single Function Agents (SiFAs). SiFAs, introduced by Brown and his colleagues (Brown et al, 1994; Dunskus et al, 1995; Berker and Brown, 1996; Grecu and Brown, 1996), are simple cooperating knowledge-based (expert) systems able to perform 'a single function, on a single target, and from a single point of view'. These three parameters (function, target and point of view) are used to specify individual agents: the function defines the types of operations an agent can perform; the target represents a parameter that an agent has an immediate effect on; and the point of view specifies a perspective that an agent takes as it performs a function, usually expressed as a goal that the agent tries to optimise. There is a limited set of functions that define problem solving activities in design, such as advice, analysis, criticism, estimation, evaluation, planning, selection or suggestion. SiFAs have mainly been used to address non-routine parametric design problems, which involve consideration of multiple constraints and requirements as well as multiple dependencies between parameters.

The approach of using single function agents with relatively simple negotiation strategies is seen as a vehicle for the investigation of the 'primitives' of knowledge and reasoning for design, redesign and conflict resolution. SiFAs are therefore seen primarily as a framework for the investigation, definition and evaluation of elementary forms of conflicts as well as types of knowledge needed for problem solving and conflict resolution. An additional characteristic of the

SiFAs approach is the emphasis on equipping agents with the ability to learn. This enhances the ability of agents to predict future interaction patterns, which can result in a reduction of the number of conflicts and/or number of messages exchanged. The subject of learning in SiFAs deserves particular attention and will be discussed in more detail in the next chapter.

2.2.4.2 MAS for design and planning decision support systems

The second direction of investigation, corresponding to research on MAS in design, is the development of design decision support systems. This research is critically linked to the development of CSCW technologies, intelligent CAD systems, as well as models and simulations that can provide input into the decision making process. Multi-agent technologies may be used for a range of purposes (and kinds of support) including the development of knowledge-rich systems, the design of (multi-) user interface(s), the integration of databases and tools, the development of languages for representation and communication, the creation of intelligent drafting and graphic recognition systems and the development of systems for prediction and evaluation.

For example, Liu et al (2002) see multi-agent technology mainly as a tool for combining diverse sources and types of information and reasoning, in order to support both human and artificial agents that take part in complex problem solving processes in design. This work includes the development of a model for a design agent based on an analysis of the design activity and its relation to learning (Liu et al, 2001). Lees et al (2001) also discuss multi-agent technology for design support in the context of concurrent engineering. They suggest the use of a set of interacting agents with different roles and problem solving capabilities (user interface agents, design critics, service agents and an agent communications server) to provide process support, particularly in the early conceptual phases of design. Notably, learning capabilities are yet again considered to be a valuable characteristic of agents. Edmonds et al (1994) consider the use of multi-agent technology to provide for more creative aspects of design. They emphasise the importance of developing agents that can potentially perceive emergent forms in design (object) representations. Multi-agent recognition of graphic units, shapes or sketches is also used for the development of intelligent drafting tools (Achten, 2002; Leclercq, 2002). Finally, some researchers have experimented with multi-agent models, such as swarms, to explore and/or generate architectural spaces and forms (Küppers et al, 2000; Coates and Thum, 2000).

Another class of applications, developed for decision support, employ multi-agent simulations to explore and visualise individual behaviour as a means of evaluating designs. For example, Dijkstra and Timmermans (2002) describe a multi-agent model, developed by combining multi-agent technology and cellular automata (CA), for the simulation of individuals moving within buildings such as shopping malls or public places. Multi-agent simulation in this case is used to predict and analyse pedestrian traffic, in order to assess the outcome of design choices and decisions. Saunders and Gero (2002) describe another model developed specifically to support the evaluation of environments that are designed to stimulate exploration, such as galleries and museums. The social force model proposed by Helbing and Molnár (1995) to simulate crowd behaviour, is extended with the realisation of an architecture for 'curious agents' that have the ability to learn from experience and provide evaluations of the interestingness of designs based on a variety of preferences for novelty. Note that in both examples coordination is achieved essentially through reactive-like rules (from velocity dependent rules to separation and collision avoidance for example).

2.2.5 Summary

An important part of current design research is devoted to the development and application of artificial intelligence theories and techniques. A sub-field of AI, distributed artificial intelligence, received particular attention in this section as it focuses on groups of intelligent distributed agents which need to coordinate their activities and decision making in order to achieve individual and global goals and improve their problem solving abilities. In more detail, the relevance of DAI and MAS with design and planning research was briefly discussed and a review of current approaches was offered, so as to highlight crucial aspects of coordination (in the design and planning domains) and identify directions for further research.

From the various approaches developed in MAS related to the question of coordination, two emerging topics of research were isolated for further investigation: sociability and learning. Both issues have previously been identified as important aspects for the understanding, modelling and support of coordination in multi-agent design.

In a brief discussion of deliberative and reactive approaches to social reasoning and action we recognised three crucial research questions: the understanding of coordination and sociability in relation to both self-interested and altruistic behaviour, by taking into account structures of conflict as well as cooperation; the definition and development of social agents as autonomous individuals, as well as members of social communities; and the examination of coordination and emergence in formal terms, particularly by looking into the effects of individual action on macro-structures, as well as the feedback of social structures to the individuals. But how were the above issues treated in the context of design and planning?

We saw that multi-agent systems in the planning literature were mainly adopted as simulation tools, to capture and represent the dynamics of individual behaviour in cities. Despite the indubitable value of such models as 'tools to think with', or as vehicles for capturing the manifestation of social phenomena in space, the exclusion of multi-person problem solving and decision making processes from these models presented us with a different challenge. That is the challenge to study and use MAS to investigate coordination in group decision making on the basis of cooperation, conflict resolution, and goal adaptation; as well as to explore processes of collective plan generation.

In the design literature on the other hand we saw that there is a great interest in modelling design agents with social abilities, although this is still a largely unexplored research area and predominantly influenced by the 'engineering' point of view. Yet, unlike most applications of MAS in other domains, applications of MAS in design emphasise the fact that coordination needs to be thought of, and modelled as, a goal-oriented and highly exploratory or co-evolutionary process.

Another aspect of coordination in design literature was that coordination has been thought both in relation to conflict resolution and management and in relation to cooperation. Finally, it was revealed that the association of coordination with generative abilities is a pertinent feature for the understanding and modelling of coordination in design domains. Notably, this feature could be very useful for the enrichment of urban planning applications.

We also saw that the various researchers working on MAS in design have focused on equipping their systems with learning and adaptation abilities, as well as the ability to construct some form of common knowledge and understanding. It should be noted that the main approach to sociability, however, has been established through the notion of situatedness, and translated into a problem of communication or learning. These issues are indeed vital for investigating sociability in design domains, but the study reported here argues that the link between coordination and sociability has remained unexplored. This link can be established by focusing on the investigation of the micro-macro relation between agents and systems, as well as on issues of emergence. It should be mentioned that some work under way investigating aspects of 'hypothetical' or reflective reasoning could be considered as a step in this direction. Nonetheless, studies discussing issues of emergence, as regards to the notion of creativity, have not attempted to establish a linkage between coordination and creativity. This is another aim of the study reported here.

2.3 THE BASIC DIMENSIONS OF MULTI-AGENT DESIGN AS COORDINATION

The thesis presented here starts with the recognition that coordination is inherent in design and planning processes evolving in complex, dynamically changing, multi-agent environments. The recognition of coordination as a crucial problem in design and planning is also a recognition of the multi-agent nature of these processes. In order to gain a comprehensive understanding of coordination, two relevant fields of research were chosen for a closer investigation: decision sciences and artificial intelligence. In the field of decision sciences, the investigation was focused on fundamental aspects of coordination found in multi-person decision making approaches (group, organisational and societal decision making) and their relevance to design and planning. Similarly, in the field of artificial intelligence, the focus was on identifying current trends in multi-agent coordination, and in parallel investigating research issues and questions relevant to the study of design in the context of architecture and urban planning.

From this review some critical aspects of multi-agent design related to coordination were identified. At the most fundamental level, coordination is associated with distribution. The distribution of decision making in multi-agent settings suggests that there is no central authority capable of controlling the process and outcome of design. In this sense, coordination is related to distributed control. An important corollary of this observation is that knowledge is also distributed among individual agents (human or artificial) that take part in the process. In this sense, sharing of information and learning are very desirable for improving the adaptability of agents in their ever-changing environment, and for ensuring consistent problem solving behaviour. As these conditions can be associated with multi-agent coordination in general, it is also important to highlight how these conditions are understood and further adapted in the context of design and planning.

From Ackoff's systemic view of planning to Gero's situated view of design cognition, design is defined and understood as a purposeful decision making process. However, in multi-agent design, the articulation and pursuit of goals is not only a collective process, but also a process that binds together problem finding and problem solving. In this sense, adaptation of goals and expectations, generation and evaluation of alternatives, and co-evolution of problem and solution, all are essential characteristics of coordination in design domains. The recognition of these characteristics is important for understanding coordination not only in relation to the management of activities (i.e. ordering, scheduling and synchronization), which is common knowledge in engineering domains, but also in relation to the generation and adaptation of knowledge and decisions.

Moreover, the above review showed that even though social aspects are pointed out in every field where some kind of group activity or decision making is required (from game theory to multi-agent system applications), sociability is an attribute related to design and planning at a fundamental level. This has a series of implications. The first is that, most obviously, multi-agent design and planning involve dealing with conflict that arises due to incompatible viewpoints and goals. This is connected with a second implication; the need to construct some form of common understanding (of the problem) through adaptation of goals and expectations. Actually, we saw that various studies in design and planning emphasised learning as a vehicle for the creation of a level of collective knowledge and common understanding. Finally, the view of design and planning as social processes implies that distinctive abilities associated with design, such as creativity, are (and should be) perceived as collective properties. To understand and potentially model coordination in multi-agent design domains necessitates therefore understanding and formalising creativity as a collective process.

This thesis seeks to develop a framework for understanding and modelling coordination that takes into account the aforementioned features. Although the various approaches reviewed have emphasised one aspect or another, currently there exists no study that effectively links the various aspects together. For example, despite the fact that co-evolution has been linked to creativity, it has not been considered within a social setting, nor it has been linked to the question of coordination. One of the positions of this thesis is that coordination can be used as an abstraction for the definition of design as a social process, wherein interdependency, conflict and construction of collective knowledge become the focal points. Additionally, in this setting, coordination can be linked to creativity through the concept of emergence, which involves considering the relation and interaction between the local level of individual agents, and the global level of the collective entity (social group).

We can now summarise the crucial dimensions of multi-agent design using the abstraction of coordination as follows:

- Firstly, distributed design tasks require knowledge that is spread among local agents. Coordination thus involves synthesising or constructing the knowledge necessary for the collective task. In this sense, learning is seen as an important instrument not only for enhancing the individual ability of agents to derive design solutions, but also for creating shared knowledge about the design task and its constraints.
- Secondly, in multi-agent design, decisions are driven by individual goals and requirements, and are taken at a local level without any external centralised source of control. Coordination is thus seen as a distributed control process which leads to the emergence of design solutions.
- Finally, in distributed design decision-making, the definition of problem and solution spaces also becomes a collective undertaking. Coordination signifies the need for a parallel exploration, generation, and reformulation of problem and solution spaces.

In the following chapters these dimensions will be explored in detail, in order to better understand their significance and their implications.

Chapter 3

INTERFACING DESIGN WITH COMPLEXITY: METHODOLOGICAL PREMISES OF THE THESIS

This chapter presents the epistemological assumptions and methodology followed in the thesis. It starts by discussing the role of science in studying and developing theories of design and by explicating the reasons why complex systems science is adopted here as an appropriate framework for this purpose. It is argued that complex systems science is suitable because it can help understand multi-agent design at an appropriate (systemic) level by offering a methodology that embraces emergence and creative construction. The following section further explains how computational modelling and simulation are used as a way to develop a theory and model of multi-agent design as coordination. This sets the backdrop against which the exploration in the coming chapters should be understood.

3.1 DESIGN, SCIENCE, AND COMPLEXITY

Design is a natural human activity and a fundamental source of change in society. It is something people do, either professionally (e.g. designing engines, buildings, cities, policies or organisations), or as part of everyday life (e.g. designing a home, planning a holiday, creating a menu or setting up a table). Despite the omnipresence of design, design research as a field was only born officially in the '60s when the first works aiming to develop a methodology for design appeared (e.g. Alexander, 1964; Archer, 1965; Jones and Thornley, 1963). Although this movement was rejected soon after, new forms of research and knowledge emerged to sustain the understanding of design as a distinct discipline. For an overview of the history and development of design research refer to Cross (1984 and 1993).

It is interesting however to see that despite this history, there is to date no general agreement about what design is, or how we should approach it. Here is a noteworthy example. In 2005 a five year research initiative was shaped in UK in order to bring together the diverse design community and establish new directions for design research and practice in the future. The initiative is called Designing for the 21st Century (www.design21.dundee.ac.uk) and is jointly funded by the Engineering and Physical Science Research Council (EPSRC) and the Arts and Humanities Research Council (AHRC). In its first round of funding the initiative offered resources to twenty one diverse multidisciplinary groups ('research clusters') so as to carry out the task of exploring the future challenges of design. In the Research Clusters Reflection & Projection Workshop held on 8-10 November 2005 in Glasgow, the different groups were invited to give an overview of their activities and their outcomes. The majority of the presentations there revealed that a great part of the efforts and discussions of the groups revolved around defining design and establishing a common understanding of concepts, terms and practices used across the design sub-disciplines. Although we all intuitively understand and agree on the idea of design, we do not have a commonly agreed theory of it!

An important reason for this is that design is a very diverse field containing different domains with different objects of investigation. Even more importantly, design research itself is varied and multidimensional. It includes studies that aim at understanding design products, processes and knowledge, through theoretical, formal and empirical examination, but also studies that seek to support designing through the development of appropriate methodologies and tools. Additionally, we can generally discern different foci of investigation in design research: a focus on *design products and artefacts*, a focus on *design processes and methodologies*, a focus on *design reasoning and knowledge*, and a focus on the *environment and practice of design*. This diversity of purposes and objects of investigation in design does not however make the task of developing a theory of design futile. It can be argued that the lack of a theory that matches people's intuition about design and design processes is what makes its development an important and even necessary task.

The real question of interest here is not whether a deeper understanding of design is necessary, but how to study design and why. The present enquiry adopts the view that design in general, and the question of coordination in particular, can and should be studied in a scientific way. It is instructive here to reflect on the relationship between design and science.

In his review of design methodology and research Cross (1993) discusses the various approaches to the development of design as a discipline, where the question of the relationship between design and science is of central concern. Many distinguish design from science using the argument that design is an essentially constructive, inventive and future-oriented enterprise, while science is analytic and concerned with existing realities and their substance (e.g. Simon, 1969; March, 1976; Steadman, 1979; Batty, 1980). Moreover, as we saw in Chapter 2 (section 2.1.1.1) design is characterised by a kind of creative reasoning which involves anticipatory thinking and reasoning from effect to cause; while science traditionally involves deductive inference, and reasoning based on observed properties. On the other hand, similarities recognisably do exist between the two, not least in terms of methodology. They both for example entail to some degree the processes of analysis, synthesis and evaluation. Understandably then, the relation between science and design has received various interpretations. Cross (1993) suggests three notions to express the different interpretations of this relationship: the notions of scientific design, design science, and science of design. *Scientific design* is used to characterise modern (industrialised) design, which uses scientific knowledge but generally may employ a variety of methods (intuitive and non-intuitive). *Design science* refers to approaching design in a comprehensively systematic way, including using technical knowledge about the product and process of design, as well as scientific methodologies and theories. This approach essentially understands design as part of science. Finally, the term *science of design* refers to the endeavour of studying design per se using scientific methods.

It is this last notion of a science of design that underlies the present thesis. For the thesis, design is regarded as a distinct 'problem' involving its own category of phenomena, processes and knowledge which can be studied using the instruments of science. However, the thesis also adopts a particular point of view about the science which is most useful and appropriate for studying design processes and problems such as coordination. That is the science of complex systems.

There are many points of intersection between complexity and design. Simon (1996) for example considered complexity as an intrinsic characteristic of design problems. Design problems are customarily thought to be complex because requirements, needs and goals are ill-defined, conflicting, and likely to change during the design process. However, complexity can also be associated with design objects or products, their special characteristics and the way we perceive and represent them (refs Gero and Kazakov, 2003; Johnson, 1995; Koutamanis, 2001; Maimon and Braha, 1996), as well as with the design process and issues such as division of labour, communication, ordering, scheduling, synchronisation, or management (e.g. Austin et al, 2002; Calinescu et al, 1998; Earl et al, 2005; Eckert et al, 2004; Qin and Johnson, 2005; Suh, 1999). This includes the study of design teams themselves as complex systems with characteristic network structures and rules of interaction (Klein et al, 2002; Sosa and Gero, 2005). Moreover, understanding design products and processes is increasingly thought to involve taking into account the complex context within which they are developed, including historical, cultural, social or technological circumstances (e.g. Ball and Ormerod, 2000; Bucciarelli, 1988; Eckert and Stacey, 2000; Fischer, 2000; Gero, 1998; Hughes et al, 1997; Liu, 1998; Lloyd, 2003). For a detailed exposition of the relationship between design and complexity and review of relevant literature see also Zamenopoulos and Alexiou (2005).

But why is complex systems science useful for studying design? Let us now briefly look at the premises of complex systems science in order to understand what it has to offer.

Complex systems science grew out of systems theory, cybernetics and information theory and the later developments in physics, biology, computing and mathematics. It encompasses theoretical and technical advances in thermodynamics, quantum mechanics, evolutionary theory, distributed computing, artificial life, and artificial intelligence and can also be thought as an intellectual offspring of holistic and constructivist approaches in philosophy, cognitive science, and sociology. A common rendering of complex systems evokes the existence of multiple components linked to one another that interact dynamically to produce some global property, structure, or behaviour (for example Holland, 1995; Kauffman, 1995). Typical examples of complex systems include the human body, chemical networks, ecosystems, social groups, cities, technological systems and so on. Complex systems scientists see complexity as a characteristic quality that explains high level 'abilities' (such as life, intelligence, evolvability or sociability), and have been developing methodologies and tools for modelling, prediction, hypothesis generation, decision making and management. The advancements in the field are also coupled with the development of sophisticated engineering products, from computer chips to robots. Like in design however, there is no generally agreed definition or theory associated with complexity. For example, while complexity usually suggests a system of interacting agents, not all conceptions of complexity reside on a notion of (multi-) agency; as a matter of fact, notions such as distribution, variety, randomness, criticality, non-linearity, incomputability, unpredictability, or irreducibility are equally central in the science of complex systems (e.g. Bak, 1996; Bennett, 1985; Chaitin, 1974; Kolmogorov, 1965; Nicolis and Prigogine, 1977). For a review of definitions of complexity see Edmonds (1999). The ambiguity of the term complexity, again like design, is due to the reality that complex systems science as a research field and domain of practice is intrinsically diverse and multidisciplinary.

Despite the diversity, there is indeed a common thread of enquiry connecting complex systems scientists, and what happens to be common among them is not a single theory, but a shared attitude towards science. Complex systems scientists embrace the notion of emergence in place of pre-determined order, and tend to look at systems by focussing on relationships and interactions. There is also a shared commitment in using generative methods and simulation in place of traditional analysis or decomposition methods. This general stance can be identified with an effort to study systems by conserving their organisation; their structural and functional characteristics, as well as the causal relationships developed among them and the environments they reside in. The idea is that we cannot truly know things if we break them apart or take them out of their context. This synthetic and holistic approach to phenomena is a central reason why complex systems science is chosen to study design in this thesis. Complexity is apposite for that purpose not only because design knowledge, processes and products are complex, but also because it is close to design as a way to approach phenomena.

Another important point to note is that at the core of complex systems science also lies the shared assumption that complexity can be used to express universal phenomena. That is, complexity can be used to make distinctions among different classes of phenomena in terms of characteristic features and structures (i.e. in terms of their organisation). To put it differently, complexity and organisation can be thought as a kind of signature for distinguishing and grouping systems of different types. For example, the notion of autopoiesis, which is important in complex systems science, has been used to typify biological (e.g. Varela, 1979), cognitive (e.g. Maturana and Varela, 1980) or social systems (e.g. Luhmann, 1986). The notion offers both a kind of distinction (for instance mechanical systems are not autopoietic systems), and a kind of characterisation (autopoiesis refers to a particular kind of structure or ability for self-construction and maintenance). If we take the view that design is indeed a distinct type of problem, capacity, or phenomenon, then taking a complex systems science perspective seems a useful thing to do.

So the main supposition behind choosing complex systems science in order to study design is that it can help us capture its characteristic features by focussing on causal processes and relationships. There is presumably also an added value in that complex systems science facilitates the systematic study of design but with an appreciation for emergence, synthetic processes, and contextual qualities. The case for complex systems science becomes even stronger when we consider the question of multi-agent which is the focal point of the thesis. We have discussed in Chapter 2 that the complexity of multi-agent design comes directly as a consequence of the decentralisation and distribution of knowledge, processes, goals and decisions. It is argued that the notion of coordination provides a way to capture the generative and creative character of multi-agent design, and account for its social nature. The deliberate choice of coordination as an abstraction for multi-agent design is exactly so that we can focus on systemic properties rather than individual (cognitive or decision making) abilities and investigate relational or organisational characteristics. When talking about organisational properties we will refer to relationships and influences between agents and the environment within which they are situated, but also to structures that characterise the design space (relationships between goals, requirements and knowledge that shape the problem and solution spaces at the individual and collective level).

These assumptions drive the specific methodological choices of the thesis. The experimentation with computational models and simulations in Chapter 5 will form the core of the overall research. Constructing models of multi-agent design is thought here to provide an appropriate means for investigating the causal processes involved. As discussed, this is a method of understanding through construction which can be contrasted to more traditional methods of investigation based on analysis of given existing systems. The role of simulation in the thesis however is somewhat peculiar and thus will be discussed in more detail in the next section. The chapters following Chapter 5 further build upon the crucial aspects uncovered via the experimentation in order to establish the unique organisational properties and characteristics of multi-agent design. Insights from complex systems research (theories and formal representations) are used in order to reflect on the role and fundamental nature of emergence in multi-agent design and formulate a view of coordination that accounts for social aspects.

Finally, it is also important to note that another critical methodological aspect of this thesis is that it strives to develop an interdisciplinary understanding of multi-agent coordination. Drawing on knowledge developed in different research domains (e.g. computer science, decision science, artificial intelligence, mathematics, cognitive science, sociology, and design research itself) is thought to enable gaining a broader perspective of the problem in hand, the available solutions, as well as the potential tools for dealing with it. Taking up interdisciplinary investigation is typical - and even necessary - both in design and in complexity research.

3.2 THE ROLE OF MODELLING AND SIMULATION IN THE THESIS

Generally stated, (computer) simulation consists in implementing a model of some real or imaginary system as a computational process. There are many ways in which simulation can be used, including applications in education, training, or entertainment. Here we are interested in the use of simulation for scientific purposes. Scientific investigation, as we discussed, aims at understanding the workings of a given reality. A fundamental assumption of science is that systems found in the world (natural, physical, social etc) operate on the basis of a set of laws or, to put it in a less strict sense, follow and entail certain causal processes. A typical way to go about understanding reality is through modelling. Modelling consists in devising another system (expressed in a different 'language', mathematical, computational, or other) in order to reproduce or represent the system in hand. A model embodies a set of rules or axioms that are supposed to represent the real system. By running a simulation of such a model one expects to match the behaviour or outcomes of the real system. The idea in this case is that if the simulation replicates this behaviour or outcomes faithfully, then it can be used to make *predictions* about the real system in the future. Modelling for prediction requires a careful encoding of observed qualities in the simulation program, verification, and validation through comparison with the real world.

Another way in which simulation has been used in science (particularly social science) is in order to generate 'would be worlds', alternative scenarios or hypotheses about the workings of different systems. Such simulations usually have a looser reference to the real world (or faithfulness to empirical evidence) and constitute a kind of experiment that helps understand the role of different parameters and explore relationships between them. Agent-based social simulation in particular, further makes the assumption that starting from simple rules and as little assumptions as possible one may generate complex behaviours or structures. This is remarkably useful for hypothesis generation, but caution is needed against using this kind of simulation to draw (unsubstantiated) conclusions about the real world. For more extensive treatments on the idea of modelling, simulation, and the role they play in prediction and explanation see Rosen (1991) and Edmonds (2005).

An important point to observe is that simulation is a deductive process: given some initial conditions (input) and a set of axioms or rules of operation (transformation of inputs to outputs), the program computes outputs. This can be considered as one of the strengths of simulation - although it seems to be conflicting with the idea of emergence. The key is to think that simulation, in comparison to deductive systems in general, is also generative; it generates data, structures or macroscopic patterns that may not have been envisaged before or only become manifested through the running of the program (for a discussion see Epstein, 1999). On the other hand, some researchers use simulation to perform inductions: the experimentation in this case is thought to provide empirical corroboration of the results and thus allows deriving generalised rules, even if these results do not come from observation and measurement of a real system. This particular double nature of simulation is what makes it a distinct and, for many, appealing type of methodology (Axelrod, 1997).

The present thesis uses simulation for theory development; in particular as a means to think through the dimensions and characteristics of coordination, and construct a coherent framework for its understanding. It is informative to consider this statement in more detail. The exploration in Chapter 2 led to the development of a theoretical view of multi-agent design as coordination, captured by the conditions of learning, distributed control and co-evolution. These conditions are quite abstract, and instantiating them within a model provides the means to make them specific. The formalisation of the model within a simulation program in turn provides a way to check the consistency of the theory. As Gilbert and Troitzsch (1999: 5) note 'the process of formalization, which involves being precise about what the theory means and making sure that it is complete and coherent, is a very valuable discipline in its own right'. This is quite different from both types of simulation explained previously. In comparison to the first type, simulation in this thesis does not have the aim of modelling a particular system to inform prediction, although it does seek to represent causal relationships. These causal relationships however, refer to an abstract, theoretical construction and not to a particular multi-agent design 'system'. The theory itself is derived in the first instance via a combination of intuition, observation, and careful examination of knowledge originated from other studies of design (i.e. review of relevant design research) and thus is thought to represent reality. In comparison to the second type, simulation in this thesis does not aim to generate alternative worlds or formulate novel hypotheses, although it is used to make deductive inferences and discover relationships among different processes or attributes. This helps re-think and refine the theory.

We can picture the role of simulation in this study as being part of a design process where the object of design is the theory of coordination. In this process, building the simulation helps to understand aspects of the theory that are ill-defined, discover critical parameters, and dispose of unnecessary assumptions. On the other hand, the theory itself is used to evaluate the instantiated model and its results and to assess the appropriateness of the language or representation chosen. Knowledge is therefore obtained not by analysis of the produced data and their statistical properties but through designing, building and experimenting with the simulation. In some cases, even running the simulation will be considered unnecessary when appropriate insight is gained by the act of construction itself. In the following chapters we will be able to see this continuous play between making theoretical adjustments according to insights gained from experimentation and reviewing the modelling and technical choices.

Chapter 4

MULTI-AGENT DESIGN AS COORDINATION: A MODEL OF DISTRIBUTED LEARNING CONTROL

Chapter 2 offered an examination of the concept of coordination in various domains, together with an investigation of critical features of multi-agent design. This led to the identification of some critical dimensions of coordination that capture the nature of design as a distributed decision making process. Through the dimensions of learning, decentralised control, and co-evolution we can start understanding coordination as a process that involves generation and reconciliation of distributed goals, but also generation and reconciliation of alternative solutions. This view not only extends the understanding of coordination beyond management to include creative design, but it also places creative design in a distributed, social setting. The current chapter aims to take this preliminary view of multi-agent design as coordination further by examining how the proposed dimensions can be expressed more precisely. The particular hypothesis put forward is that coordination can be conceptualised as distributed learning control. The chapter puts in place some basic notions and ideas that explain this hypothesis. The proposed conceptual model is elaborated further in the next chapter where different efforts for its realisation are presented and discussed.

4.1 MODELLING COORDINATION AS DISTRIBUTED LEARNING CONTROL: AN INTRODUCTION

The chief idea behind the view of (multi-agent design as) coordination in this chapter is that each agent embodies a control process which generates control actions (design decisions) in order to meet time-variant individual targets, despite endogenous uncertainties and exogenous disturbances coming from the actions of other agents. Learning corresponds to a process of capturing interdependencies among decision variables in order to improve the controlling ability of agents, as well as to inform the process of goal formation. Coordination in this sense is not an explicit mechanism but an effect of distributed learning control. To better explain this general idea it is important to first introduce the key notions of control, learning, and distributed control.

4.1.1 An introduction to control and distributed control

The idea of control, which has its roots in the domain of cybernetics (Wiener, 1948; Ashby, 1956), has been applied in various scientific areas: from the physical to the social sciences, and from biology and engineering, to economics, management and psychology. Its distinctive characteristic is goal directedness: it refers to actions taken by a system with the purpose of achieving or maintaining a target state, despite disturbances or perturbations coming from its environment. Control necessarily assumes a relation, or interaction, between a controlling system and a controlled environment. So, control systems are systems in which a controller interacts, by way of one or more controlling variables, so as to influence the state of a controlled object (also called 'plant'). At the heart of control theory lie two important complementary notions: controllability and observability. Briefly, controllability is a measurement of the ability of a system to manipulate the (states or outputs) of the controlled object, while observability is a measurement of the ability of a system to predict all the states of the controlled object on the basis of observations of its outputs.

It is customary to use diagrams to illustrate control systems. In these diagrams boxes usually represent components, devices, or processes (e.g. the controller, the control object...), while arrows are used in order to depict inputs, outputs, and interactions between components. A typical illustration of a simple control system is given in Figure 2.

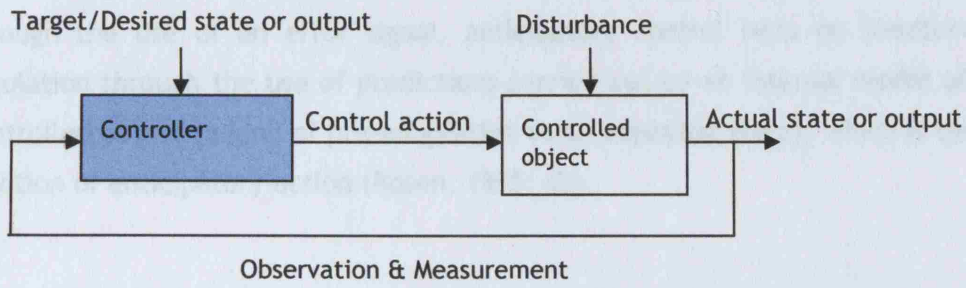


Figure 2 A simple control system.

4.1.1.1 Ingredients of control problems

Control problems arise in a variety of situations and their difficulty is due to a number of different reasons. A classic ingredient of control problems is that the controlled objects have their own dynamics and so their outputs cannot be changed instantaneously. Additionally, the range of available adjustments (control actions) that can be made to control an object is usually constrained. Other typical ingredients are that the characteristics of the controlled object are not known with certainty; and the state of the controlled object may be affected by uncontrolled, external and unpredictable inputs (known as exogenous variables). Moreover, the desired values for the state of the controlled object are often exogenous and unpredictable; current values of the state of the controlled object are uncertain; and measurements carrying information about current states are noisy (Jacobs, 1993: 1).

4.1.1.2 Control mechanisms: feedback and feedforward regulation

The most typical mechanism used in control is *feedback*. The controller attempts to steer the controlled object by adjusting a controlling variable (control signal) on the basis of measuring the difference between a desired value and the actual value of a controlled variable. This *error* measurement is fed back to the system: positively to change the variable in the same direction, or negatively to change the variable in the opposite direction. Exogenous sources may disturb the feedback signal, but the controller needs to achieve its objective notwithstanding. Traditional control was built on the notion of feedback, but *feedforward* control is another important variation. While classical control rests on feedback regulation through the use of an error signal, anticipatory control rests on feedforward regulation through the use of predictions carried out by an internal model of the controlled object (a kind of pre-adaptation to an expected state), which is tied to a notion of anticipatory action (Rosen, 1985: 42).

4.1.1.3 Adaptive control and neural network learning

When the parameters of the system vary in time, a control strategy is also needed that adapts itself to the changing conditions in order to be effective. This is the domain of adaptive control. In general, the efficiency of the controller can be improved by learning - gaining knowledge about environment behaviour and the available control variables. In the following, the focus will be on artificial neural network (NN) learning, typically used in adaptive control.

Understanding neural networks

Artificial neural networks are computational analogues of biological neural networks. The fundamental unit of an artificial neural network is a *neuron* which receives inputs and calculates outputs on the basis of a simple function. Information is stored in the *weights* of the connections between neurons. These weights among neurons change according to a process of *learning* or *adaptation*. It is useful to delve here a little bit on details about neural networks in preparation for the subsequent presentation of the computational models. For a much more detailed yet accessible overview of neural networks see Gurney (1997).

Figure 3 is an illustration of a simple neuron that receives a single input. In the figure, p represents an input to the neuron and w represents the weight (the strength of the connection); b is a bias (a constant parameter) customarily added to the weight (the sum is symbolised Σ); n is the summed output (also called net input); and f is the transfer function that produces a, the output of the neuron. The notation and illustrations here follow Hagan et al (1996) and are similar to those used in MATLAB (Mathworks, Inc), the software and modelling environment adopted for the experimentation.

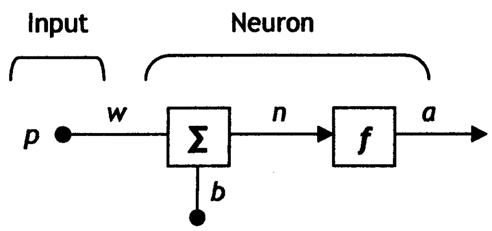


Figure 3 An illustration of a simple artificial neuron with one input.

The output of such a neuron is calculated as follows:

$$a = f\,(wp + b) \qquad \textit{Equation 1}$$

The transfer function f may be a linear or non-linear function of n, and may take different forms according to the task of the neural network. For example, the 'hard limit transfer function' sets the output to 0 if $n < 0$, or 1 if $n \geq 0$. Other transfer functions scale the output between 0 and 1.

Typically, a neuron will have more than one input $p_1, p_2,..., p_R$, each associated with a different weight $w_{1,1}, w_{1,2},..., w_{1,R}$, in which case the equation above (calculating the output of a neuron) becomes:

$$a = f\,(\mathbf{Wp} + b) \qquad \textit{Equation 2}$$

where **W** is a matrix containing the different weights and **p** is an array. A neural net will also typically have a number of neurons. The weight matrix **W** will then contain as many rows as the number of neurons and as many columns as the number of inputs to each neuron. Any number of neurons operating *in parallel* is considered to be structured in a *layer*. There are different neural network architectures corresponding to different ways to connect layers together. Feedforward neural networks are generally based on the idea that layers are connected in sequence, while recurrent neural networks contain feedback from output to input. Figure 4 shows a feedforward neural network with 2 layers.

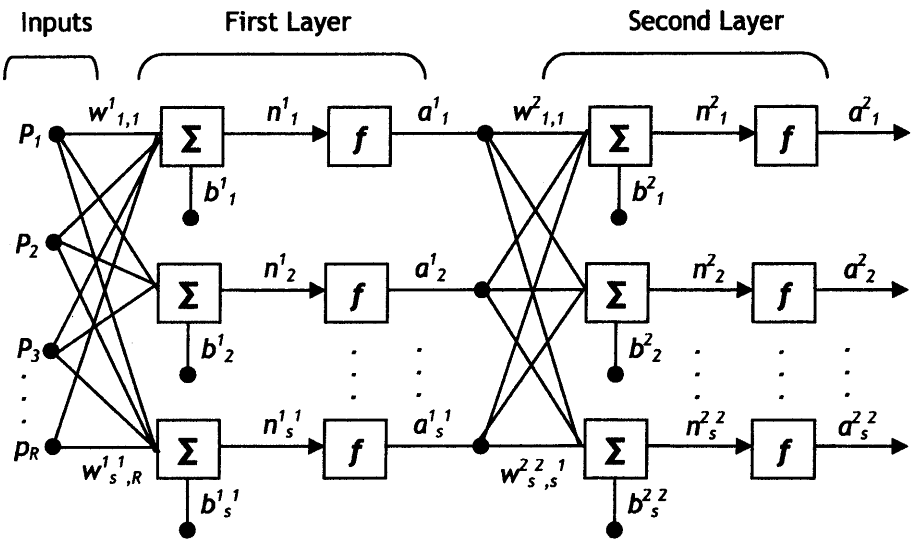


Figure 4 A two-layer feedforward neural network. The number of a layer is denoted by a superscript. Each layer has a different number of neurons S^1 and S^2 and a different number of inputs R and S^1 respectively (i.e. the number of inputs to layer2 is equal to the number of neurons in layer1). There is no feedback from layer2 to layer1.

The crucial aspect of neural networks is that computations are performed on the basis of adjusting weights and biases: the procedure for modifying weights and biases is called a *learning rule* or a *training algorithm*. Learning usually involves an iterative process whereby adjustments are made according to a number of training inputs. These training inputs may include the desired output explicitly: in this case, learning consists in making the outputs match the desired ones (*supervised learning*). In *unsupervised learning* by contrast, there is no a priori target for the network: in this case, the learning algorithm is a process by which the inputs are organised into clusters. In general then, supervised networks learn associations between a set of inputs and a set of outputs, while unsupervised networks learn to discover patterns within a set of inputs. Another typical form of learning is *reinforcement learning*. Reinforcement learning is similar to supervised learning, but instead of an explicit example of the desirable target, a reward is offered to evaluate the performance of the network: in this case the learning algorithm is a process of discovering and selecting actions that will maximise an expected reward.

Adaptive control based on neural networks

Artificial neural networks display many desirable characteristics required to address control problems: they are capable of learning by experience (from the data); they have the ability to map non-linear functions; they do not require deep understanding of the process or the problem being studied; they have the ability to generalize; and they are robust in the presence of noise (Kecman, 2001).

Neural network-based adaptive control is rooted in the idea of creating a model of the object to be controlled ('plant' modelling or identification) to be used in the control process, either directly (as a controller) or indirectly (to guide and support the control process). In general there are two basic ways to model a dynamical system by using neural networks. The first is called *forward modelling*. Forward modelling involves placing a neural network in parallel to the observed system and using the error between plant output and model output as a training signal (supervised learning). It is customary to use the letter *y* to refer to outputs (*yp* for plant output or system output, and *ym* for the model output) and the letter *u* to refer to inputs. A forward modelling structure is illustrated in Figure 5.

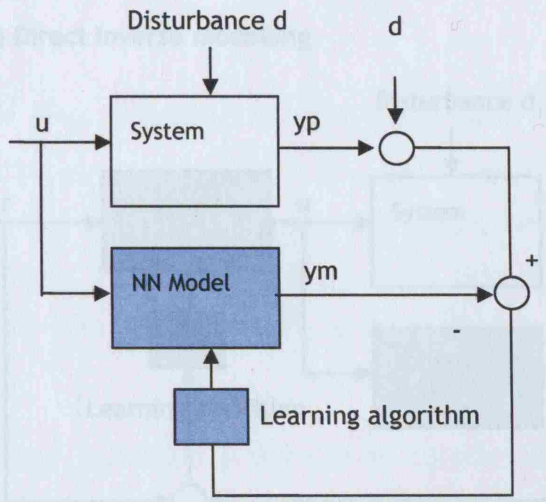
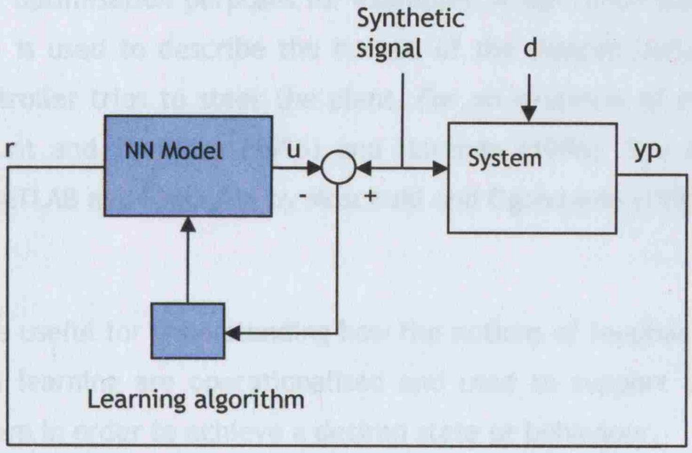
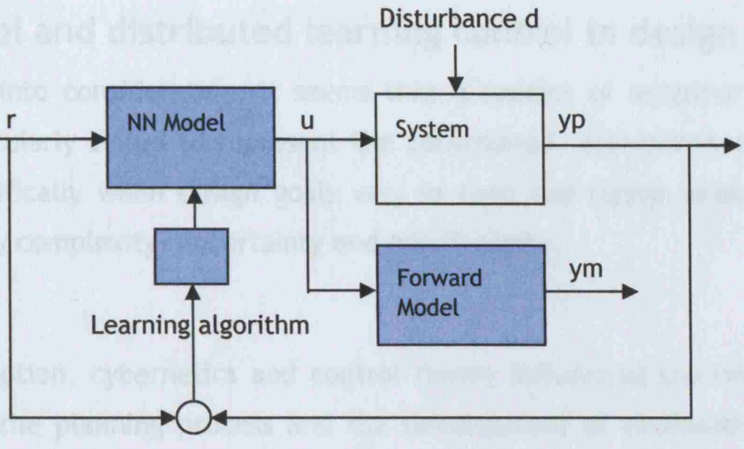


Figure 5 Structure for forward modelling of a system using neural networks

The second way to create a neural network model of a system is called *inverse modelling* where the target is to create a representation of the inverse dynamics of the system. In the simplest case of direct inverse modelling (Figure 6a), what the neural network receives as input is the output of the system produced in response to a synthetic signal. The network then is trained using the error between the neural network output and this synthetic signal. In specialised inverse modelling (Figure 6b), the neural network is placed before the system and what receives as input is a training signal which spans the desired output space of the system (a reference signal *r*). In this case, the error between the training signal and the system output is used for learning. Note that in some cases also a (forward) model of the controlled system is used to substitute the system.



a) Direct inverse modelling



b) Specialised inverse modelling

Figure 6 Structures for inverse modelling of a system using neural networks. Diagram a) represents the direct method, and diagram b) represents the specialised method.

There is a variety of control architectures which in effect use those neural network models in different arrangements, according to the nature of the control problem. For example, supervised control is like forward modelling (n.b the neural network acts as a controller) with the only difference that inputs and output targets are provided by humans. Another example is direct inverse control. In this case, the inverse model is used as a controller. The controlled system is placed in sequence with the controller so that the system output is identical to the network input (i.e. the reference signal represents the desired response). Some control architectures may use a combination of neural networks and other techniques or algorithms (for optimisation purposes for example). A reference source, or model of this source, is used to describe the bounds of the desired behaviour towards which the controller tries to steer the plant. For an overview of neural network control see Hunt and Sbarbaro (1995) and Narenda (1996). See also Advanced Control with MATLAB and SIMULINK by Moscinski and Ogonowski (1995).

These ideas are useful for understanding how the notions of feedback, adaptation, evaluation and learning are operationalised and used to support the activity of steering a system in order to achieve a desired state or behaviour.

4.1.2 Control and distributed learning control in design

Taking all this into consideration, it seems that a species of adaptive learning control is particularly suited to represent the constrained, goal-oriented activity of design, specifically when design goals vary in time and design problems are characterized by complexity, uncertainty and non-linearity.

From their inception, cybernetics and control theory influenced the creation of theories about the planning process and the development of methodologies for modelling (Forrester, 1969; Chadwick, 1971; McLoughlin, 1973; Ackoff, 1974; Beer, 1974). In addition to its scientific rigour, the great allure of the theory of cybernetics in this context was due to its ability to incorporate purposeful behaviour, account for dynamical processes, and take a holistic view of systems by paying attention to relationships and interactions. The same reasons apply in the adoption of these ideas for the understanding and modelling of design processes (Archer, 1970; Asimov, 1962).

In these studies, planning and design were equated with the process of control, while the planner, designer or decision-maker in general was considered as a controller. Archer for example (Archer, 1970) considered the design process as a kind of complex iterative control process which, roughly stated, involves generating and manipulating a set of (decision) variables in order to satisfy a set of objectives. His paper abounds with control diagrams.

Batty (1979) offers a valuable early critique of this singular control-centred view of design and planning and discusses some limitations in the way the notion of control was perceived in the view of planning (and design) as rational decision making. In particular, Batty proposes that the notion of control entails a duality between knowledge and action which was simplistically seen in these early works. A diagram of the relationship between knowledge and action in rational model of planning is given in Figure 7. This is an illustration of a control system with the controlled system on the left (reality) and the controller on the right (the planner). The planner was considered to be an expert with privileged knowledge about the city system and who was able to control reality through the planning process.

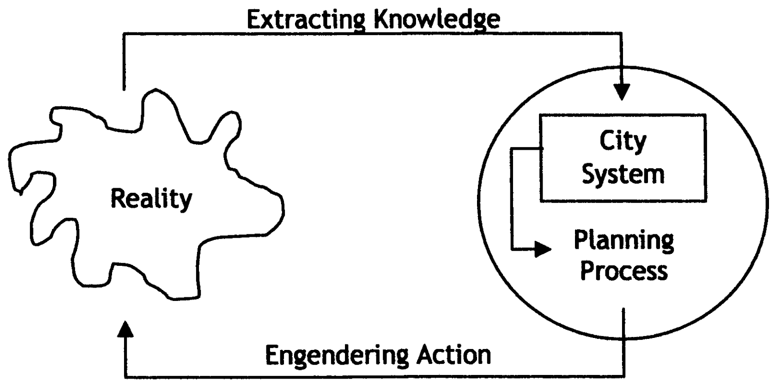


Figure 7 The relationship between knowledge and action in the rational model of planning, from Batty (1979). This is an illustration of a control system with the controlled system on the left (reality) and the controller on the right (the planner).

Batty discusses two main limitations in this view of control. The first has to do with the assumption of rationality, of the ability of the planner (or the planning system he/she represents) to extract perfect knowledge about reality. The second, has to with a failure to recognise that the planning process is not only one of the many control processes (and actions) that exist in the city, but it is also itself part of the system that it tries to control. One of the fundamental conclusions of this discussion is that a notion of decentralised control is necessary for understanding the complexity of planning and design processes, but also as a strategy for potentially producing better designs.

The lesson from this, as far the present thesis is concerned, is that we should strive for an understanding of design at an organisational level, by focussing on how goals, knowledge and activities are organised together to produce a consistent whole. This is important for design in general but even more important for multi-agent design in particular. For a more extensive critique about the use of the notion of control in design, and a detailed proposal for an alternative cybernetic view of design based on the notion of organisation see Zamenopoulos and Alexiou (2007b).

In sum, the main difference between traditional views of control in design and the view adopted in this thesis is the idea of distributing the design decision making process into different agents each with their own goals and learning capabilities. Therefore there is no central top-down controller able to control the overall goal and process. This view is obviously derived from conceptions of distributed artificial intelligence and multi-agent systems which we reviewed in Chapter 2. In this view, coordination is thought to emerge out of local actions of agents. While control refers to the actions of agents generated in order to achieve individual goals, learning refers to a process of knowledge formation aimed at improving the controlling ability of agents as well as their ability to discover and construct new goals for the design process. An informal representation of the idea distributed learning control is given in Figure 8.

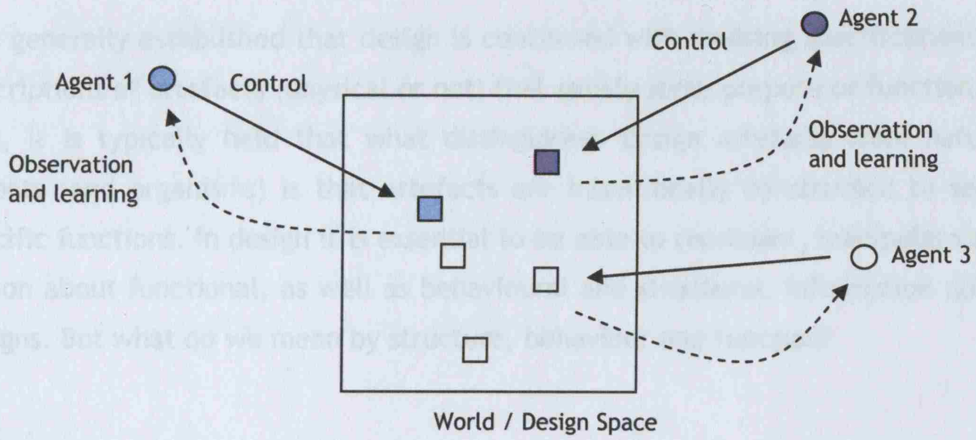


Figure 8 An informal illustration of the idea of distributed learning control. Coordination is thought to emerge out of the individual learning and control activity of distributed agents without any central control mechanism.

According to this picture, each agent therefore is thought to be a part of the distributed design process and is able to control only a part of the overall design. But because of the fact that agents' goals and actions are interdependent, each agent is also to a certain extent 'controlled' (constrained as well as enabled) by the actions of others. Observation and measurement of changes and effects of individual actions is important for improving individual performance, but also helps improve the performance of the distributed design system as a whole.

4.2 A CONCEPTUAL MODEL OF COORDINATION AS DISTRIBUTED LEARNING CONTROL

The aim of this section is to make this general picture of coordination as distributed learning control more specific. For this purpose and to discuss how the main conditions of multi-agent design identified in Chapter 2 are realised in the model, the Function-Behaviour-Structure (FBS) framework proposed by Gero (1990, 2000) is utilised. Before describing this framework, and how it is adopted and adapted in the thesis, it is instructive to make some notes about the broader context within which it is positioned.

4.2.1 Structures, behaviours and functions in design

It is generally established that design is concerned with devising specifications or descriptions of artefacts (physical or not) that satisfy some purpose or function. In fact, it is typically held that what distinguishes design artefacts from natural objects (and organisms) is that artefacts are intentionally constructed to serve specific functions. In design it is essential to be able to represent, manipulate and reason about functional, as well as behavioural and structural, information about designs. But what do we mean by structure, behaviour and function?

The most straightforward notion of the three is structure, which commonly refers to the form of an artefact and the organisation of its components. For example, in architectural design, structure would refer to the description of the form of a building and the way space is configured (including topological and geometric relationships between spaces).

The notions of behaviour and function are more complex and contentious. From a typically engineering perspective, behaviour refers to the operation of an artefact through which a particular function is satisfied, i.e. it describes how a particular function is achieved. For example, the behaviour of a door refers to the operation of sliding, swinging, or rotating in order to allow or prohibit access to someone in a room. Behaviour may also however refer to some observation or measurement applied over a particular structure. The energy consumption of a building is an example of a behaviour derived from the structure of the building and the occupation patterns of its inhabitants. Finally, behaviour may also refer to the change of structure in time, for example the growth of a city.

Function is the most contentious notion of all. It can generally be perceived as a description of what an artefact does: for example, a building provides a living space, a door allows/prohibits access, a window provides lighting and ventilation and so on. In this sense, function is closely related to behaviour, but this relationship can be perceived and represented in different ways. For example, in their review of functional reasoning in design, Umeda and Tomiyama (1997) identify two ways by which conceptions of function in relation to behaviour may differ. First, there are different views about the representation of behaviour through which function becomes expressed. For example, function can be represented through verb-noun pairs, input-output flow transformations, or as a transformation between input-output situations/states (Chakrabarti and Blessing, 1996). Second, there are different ways to perceive the scope of function: others represent function only as a part of behaviour ('abstracted behaviour'), while others also incorporate the designer's intention.

This last issue is particularly important. At the beginning of this subsection it was stated that artefacts are distinguished from natural objects because they are intentionally constructed to serve specific functions. A notion of purpose or intentionality seems to be intrinsic in the notion of function. An artefact of course may have unintended functions, so there is a difference between function seen as an activity and function seen as utility. Some reserve the term function to refer to the second case only. On the other hand, function can also be perceived as an assignment of a meaning, a name, or a label over a particular structure; in this sense function is an interpretation of structure that characterises what the structure is, rather than what it does. In this sense again, function may be an abstract concept, a need, an abstract activity or an abstract idea, which is both the driver and the final destination of design. For an interesting and comprehensive discussion about the meaning and role of function in biological and artificial (designed) systems see Buller (1999).

The position adopted in the thesis, is that function is indeed both the driver and destination of design, but it is not seen strictly as a description of what an artefact does. So behaviour is a kind of mediator between structure and function: the structure impinges on the behaviour of an artefact, the behaviour achieves the function, and on the other hand function entails behaviour and drives the generation and selection of structures.

4.2.2 Gero's FBS framework for modelling the design process

Descriptive and prescriptive models of the design process typically include the tasks of analysis, synthesis and evaluation. More specifically, the design process is customarily modelled as an iterative process where analysis, synthesis and evaluation are structured in a series of loops (that can be visualised as a kind of spiral) which moves from the more abstract to the more specific. This process begins with abstract functions or needs, and ends with a satisfying form or structure for the design artefact. Although other components may also be added in this general model (for example communication) and although different pictorial representations may be used, the core of the idea remains the same - compare for example Pahl and Beitz (1984), French (1992), Roozenburg and Eekels (1995). The reason for choosing Gero's FBS framework here is because it offers a way to understand the role of functions, behaviours and structures within this iterative design process.

In Gero's FBS framework, function is generally tied to purpose, although there are other variations (for instance in Rosenman and Gero (1998) purpose is differentiated from function). Gero also distinguishes two types of behaviour: structural behaviour which is directly derived from structure, and expected behaviour which is derived from function. Here we will use the version described in Gero (1990, 2000). In these papers Gero identifies a set of processes by which a designer moves from function to structure and thereby to an appropriate design description. The model is illustrated in Figure 9.

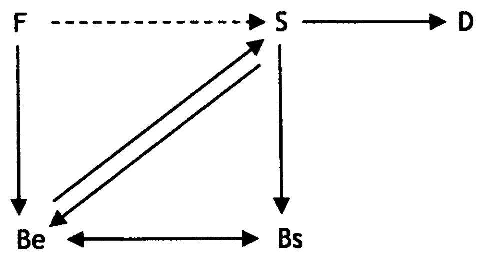


Figure 9: The FBS model adapted from Gero (1990). F denotes function, S denotes structure, B denotes behaviour (Be=expected and Bs=structural) while D stands for design description. Arrows indicate transformations between FBS. The double arrow denotes comparison, and the dashed arrow denotes an occasional transformation.

More analytically, these processes are: a) *analysis*, where the behaviour of a structure Bs is deduced from this structure S, b) *formulation*, where function F is mapped to expected behaviour Be, c) *synthesis*, where the expected behaviour Be is used to produce a structure S, based on knowledge of the behaviours Bs, d) *evaluation*, where the structural behaviour Bs is compared with the expected behaviour Be to determine whether the synthesised structure can satisfy the desired function, e) *reformulation*, where the range of expected behaviours Be can be changed, and through them the function, and h) design *description*, where the structure is developed into a description of the artefact to be constructed. In later papers (e.g. Gero and Kannengiesser, 2002), two other types of reformulation are also added (not included in Figure 9): one which refers to changes in terms of structure variables or ranges of values for them, and one which refers to changes in terms of function variables or ranges of values of them. For a summary see Table 5.

Analysis:	S → Bs
Formulation:	F → Be
Synthesis:	Be → S
Evaluation:	Be ↔ Bs
Reformulation:	S → Be S → S S → F
Design Description:	S → D

Table 5 Summary of the processes involved in design as discussed by Gero and Kannengiesser (2002)

4.2.3 Employing the FBS framework to propose a model of coordination as distributed learning control

From this presentation, the processes of analysis and synthesis can be taken to represent operations that concern the solution space (the space of possible structures), whereas formulation and reformulation concern the problem space (the space of possible functions). This gives a means to understand and model how a co-evolution of problem and solution spaces may be realised. The conceptual model of coordination introduced here develops this idea and recasts the processes of analysis, synthesis, formulation, reformulation and evaluation within the context of distributed learning control.

More analytically, a distributed design task involves a number of agents that act on a common world whilst trying to achieve their individual visions about this world. Thus, there is a direct interaction between an agent and its environment, which includes other agents. Each agent is considered to carry out two combined control-based activities depicted in Figure 10.

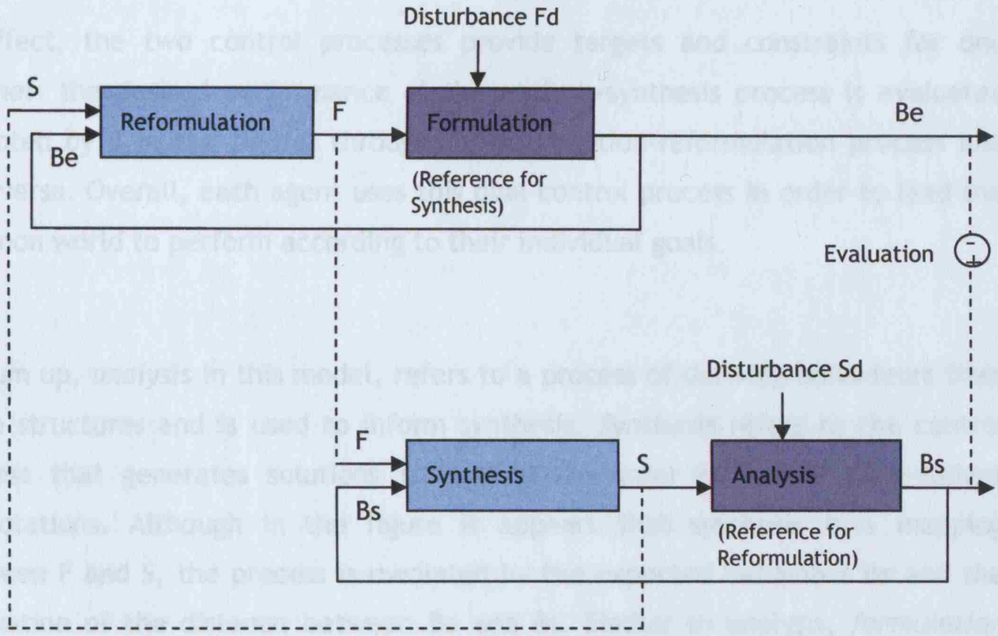


Figure 10: Coordination as a dual control process, one that corresponds to a synthesis-analysis-evaluation route and one that corresponds to a reformulation-formulation-evaluation route.

The first control activity alludes to an analysis-synthesis-evaluation route. The objective of each agent in this case, is to find a suitable path or class of structures S (control actions) that can lead the behaviours Bs to follow the expected behaviour Be (derived by a target function F), despite uncertainties and despite exogenous disturbances Sd produced by other agents' decisions. Here, knowledge of associations between proposed structures S and actual behaviours Bs, and the observed distance or 'error' between expected and structural behaviours, are both used to inform the selection of suitable control actions S.

The second activity alludes to an evaluation-formulation-reformulation route. The objective of each agent in this case is to find a suitable function F leading to such expected behaviour Be that can satisfy a given structural configuration S, despite uncertainties and despite exogenous disturbances Fd (i.e. other agents' goals). Here, again, knowledge of associations between proposed functions F and the overall derived behaviour Be, as well as the 'error' between this and the actual behaviour of the world, are used to inform the generation of suitable targets F.

In effect, the two control processes provide targets and constraints for one another: the desired performance of the analysis-synthesis process is evaluated (denoted by E in the figure) through the formulation-reformulation process and vice-versa. Overall, each agent uses this dual control process in order to lead the common world to perform according to their individual goals.

To sum up, *analysis* in this model, refers to a process of deriving behaviours from given structures and is used to inform synthesis. *Synthesis* refers to the control process that generates solutions (structures) in order to satisfy time-variant expectations. Although in the figure it appears that synthesis is a mapping between F and S, the process is mediated by the expected behaviour Be and the evaluation of the distance between Be and Bs. Similar to analysis, *formulation* refers to a process of deriving behaviours from functions, and is used to inform the reformulation process. *Reformulation* is the control process that aims to redefine the problem formulation (function) so that the expectations developed for a design object respect limitations posed by knowledge of available solutions. Again this is not a direct mapping between S and F, as the process is mediated by the actual behaviours Bs observed in the world and the error between actual and expected behaviours. *Evaluation* is used to inform both synthesis and reformulation and hence constitutes a measurement of the degree of 'matching' between the two control activities. The control signals St,..., St+n produced by this combined control constitute a set of evolving solutions (proposals) for the design problem in hand.

This conceptualisation recognises that each agent's goals and solutions may be conflicting and explicitly makes this part of the design process. Distributed control captures the idea that each agent needs to arrive at design solutions by avoiding conflicts (or disturbances) introduced in the world through the action of other agents. The process of solution generation based on learning control is a process of self-adaptation of agents that leads to coordination of their distributed descriptions.

4.2.3.1 How the three dimensions of multi-agent design as coordination are incorporated in the model

It is worth highlighting how the three conditions of coordination - decentralised control, learning, and co-evolution - are embedded in this model. The core of the proposed model is the idea that in a distributed design task, different agents are driven by different and often conflicting goals, and are only able to take actions and decisions locally. There is no single agent who can control the overall design process. Moreover, the goals of each agent change over time while the design process progresses. The model assumes that this can be modelled as a distributed control process, where each agent uses a control mechanism to achieve individual goals by avoiding conflicts and internal uncertainties. In fact, the design process is not considered to follow exclusively a route from desired functions to structures, but it also involves deriving appropriate functions that can reformulate the design problem in new, more effective, terms.

Additionally, for each agent to be able to derive appropriate structures or functions, it is necessary to create a model of how the world (its environment) reacts to different propositions. Learning not only refers to knowledge about the effects of design decisions and the performance of the designed artefact, but also to knowledge about the relationship between functions (goals) and structures (design solutions). Through learning and reciprocal control, the problem space and the solution space can be said to co-evolve together.

In this regard, the creative ability of the agents is largely related to the possibility to discover and learn novel interdependencies among FBS variables and facilitate the co-evolution between problem and solution spaces. The existence of disturbances is then desirable to the degree that they may constitute a source of variation and an opportunity for novelty for individual agents. Furthermore, the distribution of the design process in different agents, with individual knowledge and targets, implies the possibility for global phenomena to emerge through the interaction and self-adaptation of agents at the local scale. Coordination can thus be conceptualised as an emergent effect.

4.3 SUMMARY AND CONCLUSIONS

This chapter was concerned with the development of a model of coordination able to incorporate the three dimensions of decentralised control, learning and co-evolution, identified in Chapter 2. The idea behind distributed learning control is that agents basically perform a control process that generates actions (design decisions) in order to meet time-variant individual targets, despite endogenous uncertainties and exogenous disturbances coming from the actions of others. Learning in this setting corresponds to capturing interdependencies among decision variables in order to improve the controlling ability of agents, as well as to inform the process of goal formation. Coordination is thus conceptualised as a condition that is not explicitly modelled but rather emerges via the distributed actions of agents.

The model of coordination as distributed learning control was developed further using the FBS framework proposed by Gero. In this way, the design processes of analysis, synthesis, formulation, reformulation and evaluation were incorporated in the model. More specifically, coordination was modelled as a distributed dual control process of analysis-synthesis and formulation-reformulation. The two control processes inform one another through evaluation of the distance between actual, desired and expected structures and functions.

These ideas are further elaborated in the next chapter by way of a series of computer models and simulation experiments. The reasoning behind this is to use simulation in order to make the dimensions of coordination explicit and more precise. In particular, the purpose is to explore whether and how distributed learning control can be used as a model for capturing these dimensions, and reveal potential benefits and drawbacks.

Chapter 5
BUILDING COMPUTATIONAL MODELS OF COORDINATION AS DISTRIBUTED LEARNING CONTROL

The previous chapter presented a conceptual model of coordination which embodies the basic dimensions of multi-agent design. Coordination was specifically modelled as distributed learning control. In this chapter we turn to computational modelling and simulation as a way to realise the proposed conceptual model and the main assumptions behind it, and also evaluate how this model captures the notion of coordination and the identified conditions of multi-agent design. The chapter presents two different realisations (two versions of multi-agent design as distributed learning control) developed using MATLAB-SIMULINK (Mathworks, inc) technology. Findings from this experimentation are used in order to refine the conceptual model, and identify important issues for further exploration and development.

5.1 COORDINATION AS DISTRIBUTED LEARNING CONTROL: VERSION 1

5.1.1 Building the distributed learning control model

5.1.1.1 The design problem

For specificity, a particular design problem was considered: that of devising a simple land use and layout plan for a hypothetical urban development. Such problems, typically involve various groups of agents with different goals and responsibilities. Each agent manipulates a partial component of the overall description. This reflects the idea that agents may convey domain-specific knowledge about the design problem, they may have limited capabilities to control the overall task, and their proposals may change in time according to changing situations and new knowledge gained in the process.

In line with the general stance of this thesis it is assumed that plan descriptions are dynamically generated and modified through the interaction between human actors and artificial agents. Plans are composed by objects introduced and manipulated by agents on the basis of desired functions or goals. Plans therefore work as an interface among human and artificial agents and constitute the space where conflicts are expressed and observed. In the version of the distributed learning model reported here plans are represented as a virtual reality (VR) space. The model considers that the way human users manipulate objects in the VR world constitutes a knowledge source for the artificial agents. See Figure 11. For the experimentation however simple inference systems were used to represent users and model their design decision making. This will be discussed in more detail later.

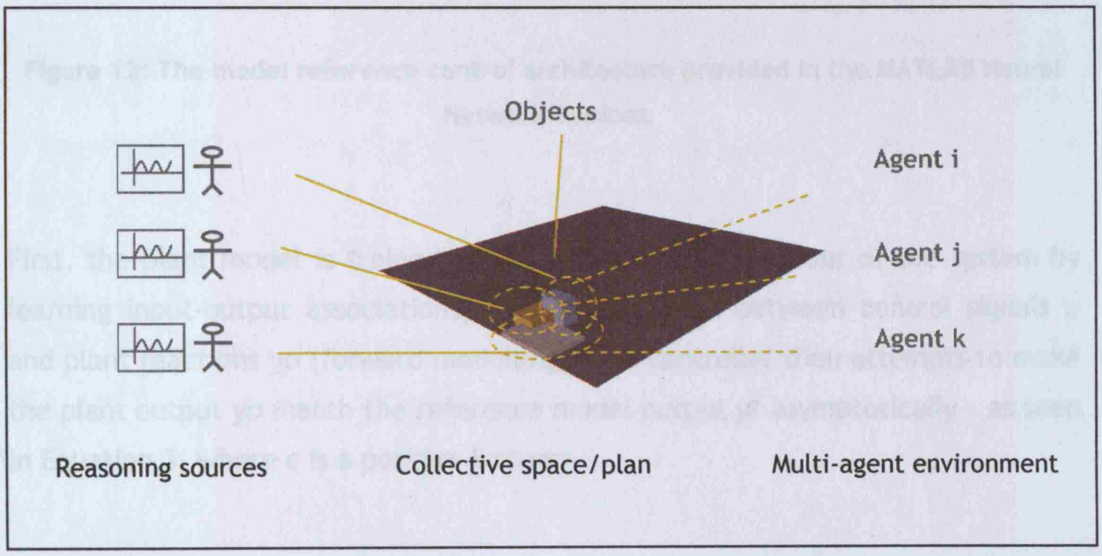


Figure 11: The general setting of the model. It assumes a number of agents (human or artificial) which act on a collective space. Each agent manipulates an object within this space. The way human users (or their models) manipulate the objects constitute the knowledge sources for the artificial agents.

5.1.1.2 The control architecture

For this experiment, the 'model reference control' architecture was chosen. This architecture uses two neural networks: one to act as the controller (the system that controls) and another to act as a 'plant' model (a model of the system to be controlled). The architecture also contains a reference model, which provides an explicit goal for training the controller (Figure 12).

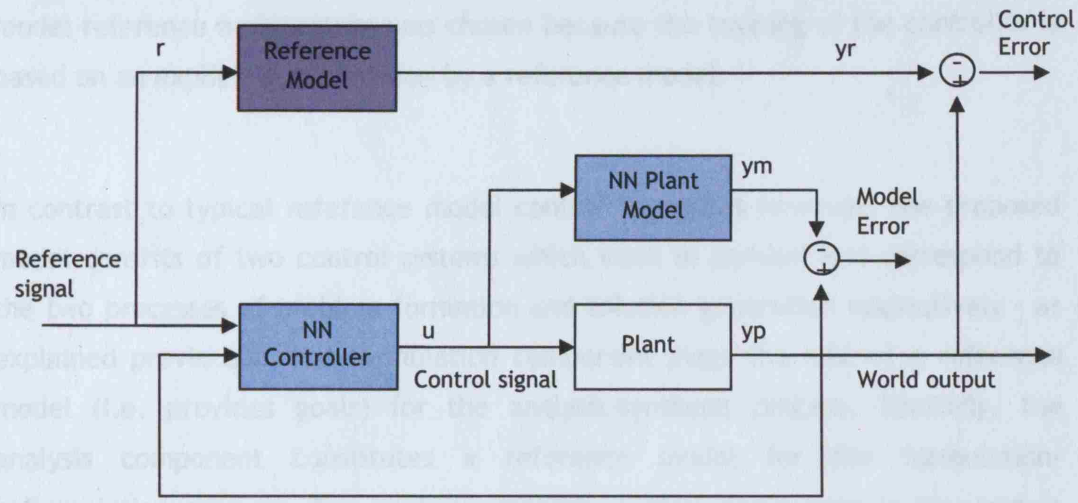


Figure 12: The model reference control architecture provided in the MATLAB Neural Network Toolbox

First, the plant model is trained to approximate the behaviour of the system by learning input-output associations, that is, mappings between control signals *u* and plant reactions *yp* (forward modelling). The controller then attempts to make the plant output *yp* match the reference model output *yr* asymptotically - as seen in Equation 3, where ε is a positive integer:

$$\lim_{t \to \infty} |yp(t) - yr(t)| < \varepsilon \qquad \textit{Equation 3}$$

For a more detailed presentation of model reference control see Landau (1979).

It is useful to see one by one the different elements of the system, as adapted to serve the proposed conceptual model of coordination.

If we take as plant the space with the three cuboids described above, then these three objects can be thought of as representing three agents with different development goals. The plant model (which we will call world model from now on) serves to identify the behaviour of those objects as agents manipulate them to achieve their goals, and this knowledge is used to guide the control process. The model reference architecture was chosen because the training of the controller is based on an explicit goal provided by a reference model.

In contrast to typical reference model control structures however, the proposed model consists of two control systems which work in parallel and correspond to the two processes of problem formation and solution generation respectively - as explained previously. The formulation component plays the role of a reference model (i.e. provides goals) for the analysis-synthesis process. Similarly, the analysis component constitutes a reference model for the formulation-reformulation process. The analysis-synthesis control component is pictured in Figure 13 and the formulation-reformulation control is pictured in Figure 14. Note that in the analysis-synthesis control process the error between the behaviour Bs of the world and the behaviour Bs derived by the world model is used in order to train the world model: this is a typical feedback control loop. The error between the behaviour of the world and the reference behaviour Be derived by the formulation process however constitutes a performance criterion and is used to adjust the control parameters. The same idea obviously applies to the formulation-reformulation control process.

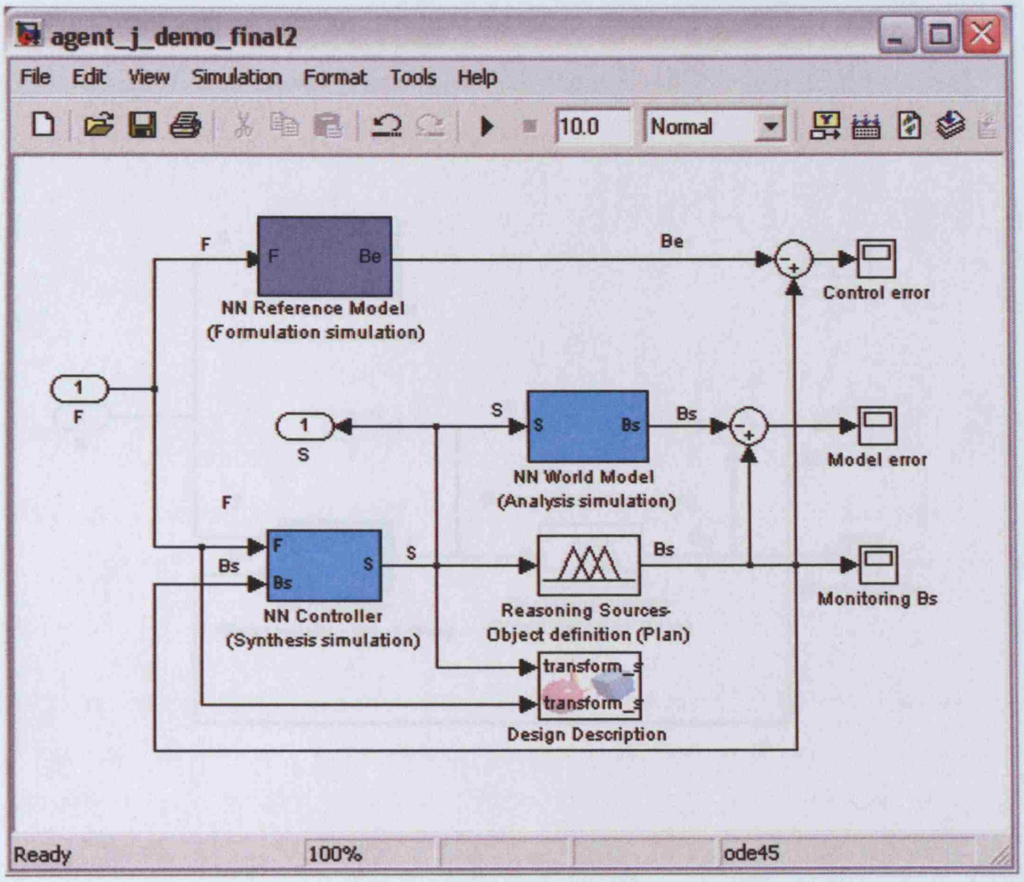


Figure 13: The proposed Analysis-Synthesis control structure.

It is also important to notice that this architecture does not correspond to a single control agent responsible for overseeing the design process. It can either be considered as an abstract representation of the overall design decision making process performed by distributed agents (a representation of the overall coordination process), or equally, it can be considered as a representation of the mechanism that drives the actions of each individual agent (a representation of the design decision making process within an agent). The two alternatives are equivalent in that in both cases the conflicts and inconsistencies, which arise because of the distribution of goals and knowledge, drive the collective development of the dual process of problem and solution construction.

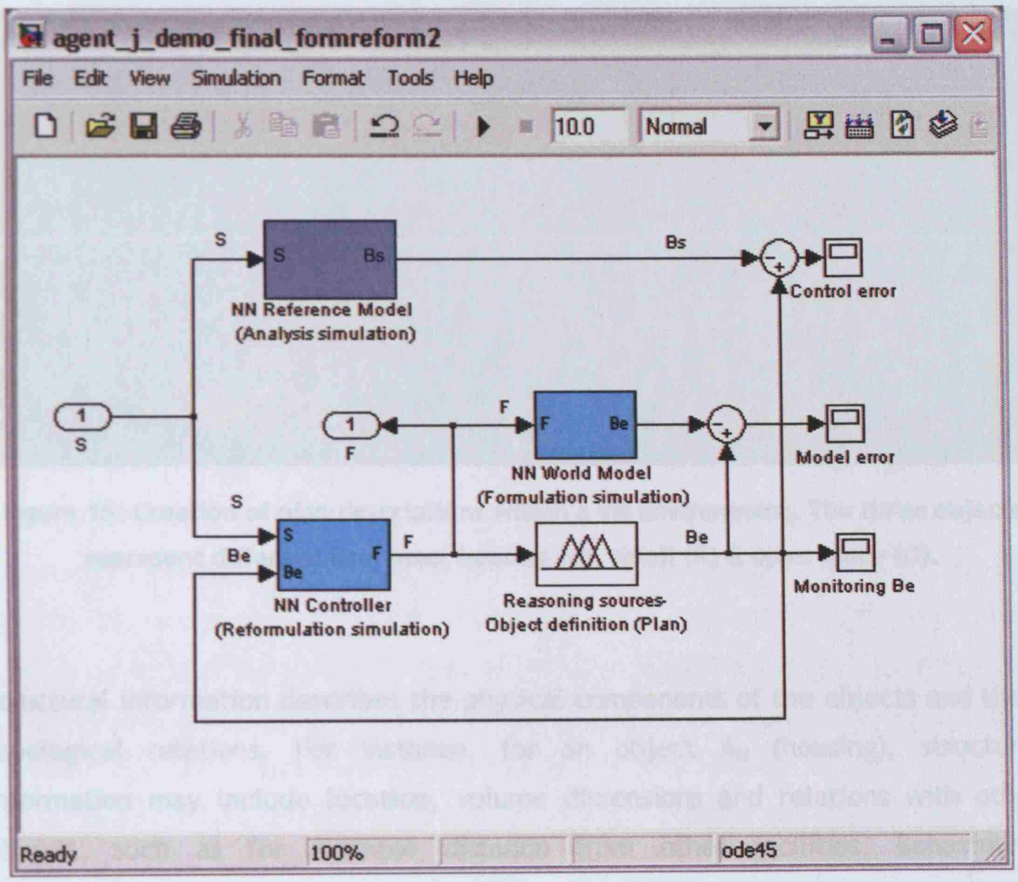


Figure 14: The proposed Formulation-Reformulation control structure

5.1.1.3 Representing functions, behaviours and structures

The above architecture is still quite generic. Functions, behaviours and structures are represented as abstract classes of information. To implement this model, we need to consider how to express these classes as specific variables.

We can consider the existence of three objects, each controlled by a different agent. Each object can be specified as a matrix $A_i = [F_i, B_i, S_i]$. The overall plan description is the matrix $P = [A_i]$ of all these objects. Let us also consider that functional information represents the ontology of each object which we can interpret as a different land use: housing (denoted H), retail (denoted R) and open space (denoted O). See Figure 15.

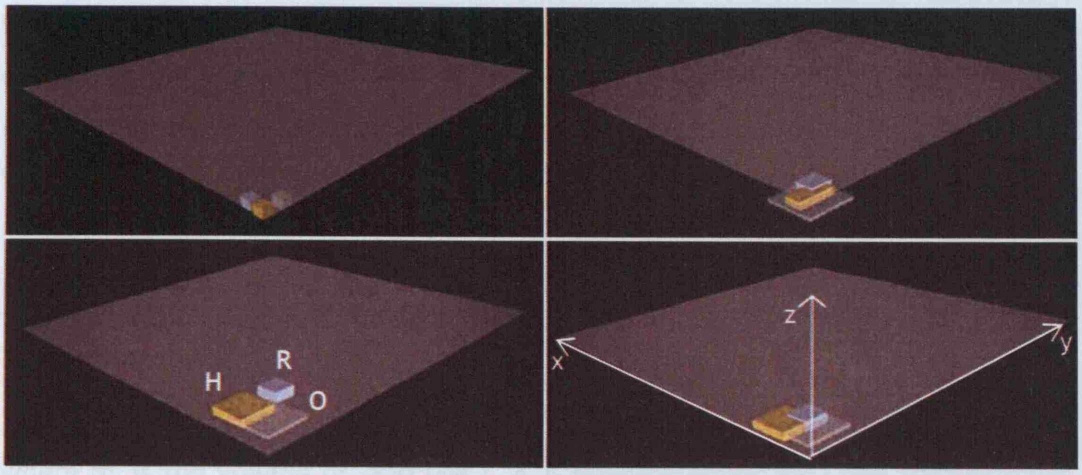


Figure 15: Creation of plan descriptions within a VR environment. The three objects represent different land uses: housing (H), retail (R) & open space (O).

Structural information describes the physical components of the objects and their topological relations. For instance, for an object A_H (housing), structural information may include location, volume dimensions and relations with other objects, such as for example distance from other facilities. Behavioural information specifies the way each object reacts to changes of its state and its environment in order to reach its intended function. For instance, behaviours may describe land use attractiveness (tendency of land uses to be attracted to - or repelled by - other facilities) and change of distance between objects. As discussed, behaviour may also refer to observations or measurements applied over structures, for example cost (derived in relation to land value and floor area).

5.1.1.4 Modelling the reasoning sources

Let us now discuss some modelling exercises developed to explore how it is possible to represent human reasoning and model different choices about the structure, behaviour and function of objects in the VR world.

Models of spatial interaction in geography and planning typically use some form of gravitation or attraction-repulsion concept to explore location-allocation problems and model flows and interactions between spaces. The same idea was used in the thesis in order to model location changes (movement) of the cuboids as a function of the attractiveness between different land uses. In specific the Newtonian equation of motion was used to model the 'moving behaviour' of a land use j as follows:

$$m_j x_j'' = \sum_{i=1}^{n} k_{ij}(x_i - x_j) \qquad \textit{Equation 4}$$

where m_j is the volume of land use j, x_j is position of the land use, and k_{ij} is the interaction matrix between land use j and i (that defines the strength of the adjacency relationship between land uses). The parameter x_j'' refers to the velocity of movement. This was implemented by considering that the three cuboids are connected by springs and the desired adjacencies are represented as forces that operate at the centre of the cuboids. The sum of the distances between land uses multiplied by the strength of the connection between them (the stiffness of the spring) gives the total force applied on the cuboids. The outputs are new locations (structural information). A similar principle can be applied to model the change of volume dimensions assuming that springs are placed between the walls (the surfaces) of the cuboids.

A fuzzy logic system was also built to model human reasoning. Fuzzy logic (introduced by Zadeh in 1965) is a methodology used to turn expert qualitative knowledge into workable algorithms. It is useful to introduce some general principles of fuzzy logic here before presenting the model. For comprehensive introductions to fuzzy logic see Kosko (1994), and Klir and Yuan (1995). Briefly then, fuzzy logic refers to a possibilistic kind of reasoning where propositions are neither true nor false, but are both valid to a certain degree. The basic notion behind fuzzy logic is the notion of a fuzzy set. In traditional set theory we assign a cut-off point to express whether an element is a member of the set or not. In fuzzy sets, membership is a matter of degree - so we assign a membership function to express the *degree* to which a member belongs to a set. For example, the set of houses which are close to some given place of work is described by a membership function that steps from 0 (close) to 1 (far) at a particular point, say

1Km (Figure 16 left). The corresponding fuzzy set would look like the Figure 16 on the right, where places are evaluated as close or far to a certain degree between 0 and 1 (for instance 500m are 'very close' and 1.5Km are 'quite far').

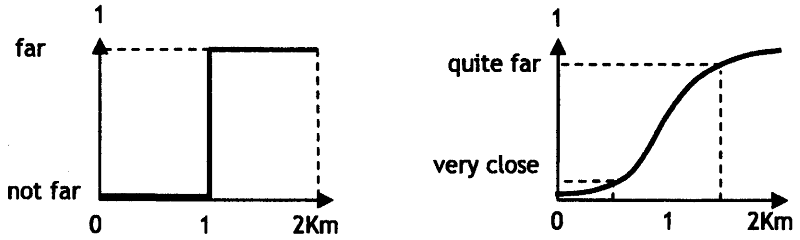


Figure 16 Membership functions for a crisp set (left) and a fuzzy set (right) describing close distance between home and place of work.

In fuzzy logic, logical statements also become a matter of degree. Figure 17 shows how the truth operations AND, OR and NOT work for two-valued logic (upper part) and fuzzy logic (lower part). The logical operators AND and OR between two sets A and B are substituted by the operators MIN and MAX, while NOT A becomes 1-A.

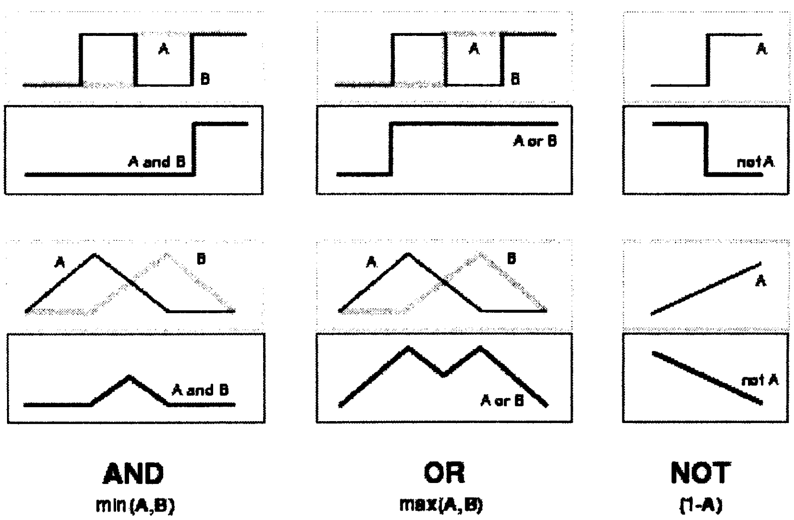


Figure 17 Figure showing how the logical operations AND, OR and NOT work for crisp sets (upper part) and fuzzy sets (lower part) (MATLAB documentation).

Inferences are formed using IF-THEN rules (fuzzy propositions expressed in natural language). For an example from architectural design, a rule may associate certain structural characteristics of a building with an expected behaviour: *IF west-facing surface is large and windows are big THEN room temperature is high in the summer*. A typical fuzzy inference system will include more than one fuzzy IF-THEN rules, which are combined to give a final evaluation.

In the thesis, a simple fuzzy inference system was built to model attractiveness between cuboids, and represent qualitative evaluations about the fitness of a specific location based on criteria of proximity with neighbouring facilities. This system was implemented using the Fuzzy Logic Toolbox in MATLAB. More specifically, the system has as inputs the fuzzy sets housing (H), retail (R), open space (O) and volume (Vol), and as outputs the fuzzy sets proximity to housing (ProxH), proximity to retail (ProxR), and proximity to open space (ProxO). This is illustrated in Figure 18. Housing, retail and open space represent the function of the cuboids, while volume expresses structural information. Proximity expresses behavioural information. All the input and output sets are described by Gaussian membership functions. As an example, the housing membership function is illustrated in Figure 19.

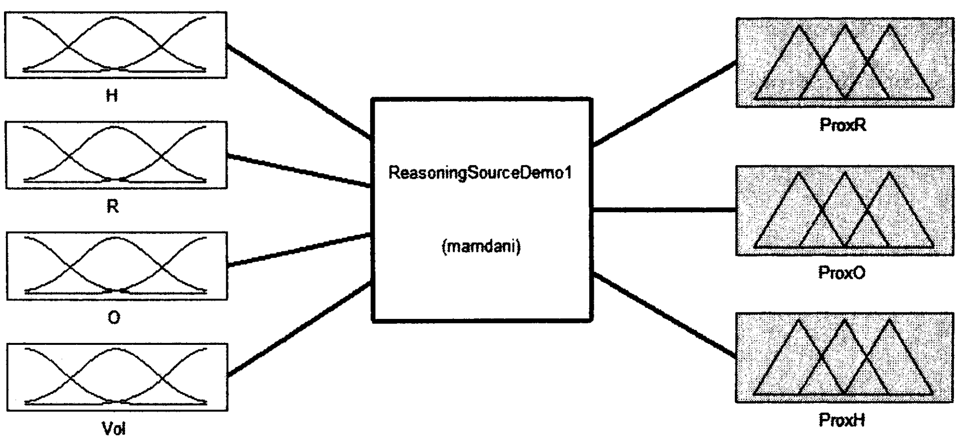


Figure 18 The fuzzy inference system used to model the behaviour of the cuboids. H=housing, R=retail, O=open space, Vol=volume, ProxH=proximity to housing, ProxR=proximity to retail, ProxO=proximity to open space.

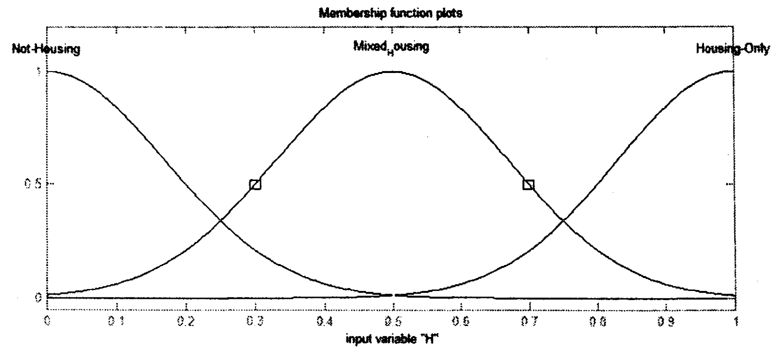


Figure 19 The membership function for housing (Gaussian).

Two rules were created to model the behaviour of land uses and particularly tendencies of attraction and repulsion among them. More specifically, the first rule states that *IF there is a high degree of housing, a high degree of open space and the (cuboid) volume is small, THEN (the desired) proximity with retail is high and the proximity with open space is high*. The second rule states that *IF there is a high degree of retail, a high degree of open space and the volume is large, THEN (the desired) proximity with open space is high and the proximity with housing is low*. The two rules represent different development goals. The first rule expresses the idea that it is desirable for houses to be close to retail facilities and open (un-built or recreational) spaces. On the other hand, when there is a very big open space available large retail facilities will tend to locate there and in distance from housing (so that land costs are kept low). The rules are shown in Figure 20.

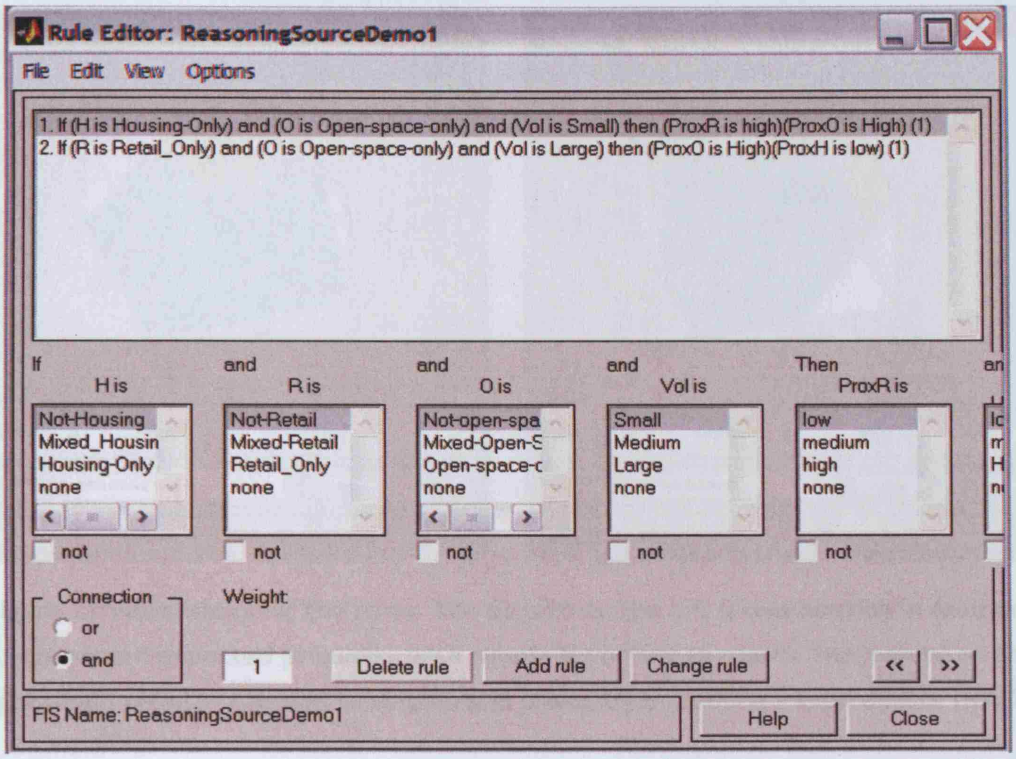


Figure 20 The two fuzzy IF-THEN rules built in order to represent the behaviour of the housing and retail cuboids.

This represents a conflict situation as we can consider that the agent that controls the housing cuboid will tend to move it close to the retail cuboid, while the agent that controls the retail cuboid will tend to move it far. The conflicting rules can be visualised using the 'surface viewer' tool provided by MATLAB (Figure 21). The picture on the left shows housing in relation to its volume dimensions (horizontal plane) and the expected proximity with retail (vertical axis) according to rule 1. The picture on the right is the corresponding diagram according to rule 2 where retail is mapped in relation to volume in the horizontal plane while the vertical axis represents proximity to housing. The rules express conflicting goals about the configuration of space and so the two surfaces look like they are the negative of one another.

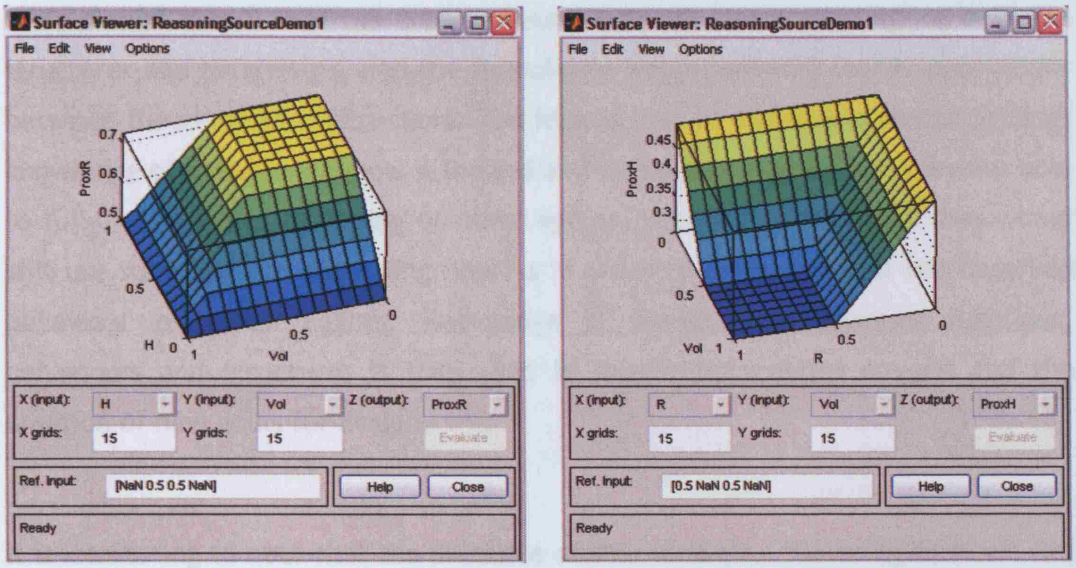


Figure 21 Visualisation of the rules. The picture on the left shows housing in relation to volume and expected proximity with retail - according to rule 1. The picture on the right shows retail in relation to volume and proximity to housing - according to rule 2.

The whole fuzzy inference system works as follows. The system receives as inputs values for housing, retail and open space that range between 0 and 1 and represent the degree to which a land use is present. The fourth input is an evaluation of the building volume (that ranges from 0 to 1 from small to big). Then for each rule the values are mapped on the membership functions, the logical implications are applied, and a fuzzy set is produced as output. The outcomes of the two rules are aggregated according to an aggregation method. The final step is to 'defuzzify' the result so as to produce a single output value. The most common defuzzification method is the centroid calculation, which returns the center of the area produced by the aggregation method.

Both the model that is based on the Newtonian law of motion and the fuzzy system were used in order to explore how to represent functions, behaviours and structures and model human reasoning about their association (including conflicts). The inputs and outputs of these models are inputs to the two neural networks that aim to develop a model of the world (the plan) and the activity of the cuboids[4]. Recall that the analysis neural network learns associations between structures and behaviours, and the formulation neural network learns associations between functions and behaviours. The idea is that in distributed design settings knowledge of individual agents is limited and we cannot assume that they are able to fully 'model' the reasoning of other agents, yet we consider that they would still use some kind of modelling process in order to make sense of the observed behaviour of other agents. Knowledge of associations between functions, behaviours and structures is then used to inform the control process and the creation of new goals for design.

It is interesting to note that the presence of the world (the reasoning sources) was implicit in the conceptual model of distributed learning control, but became explicit through the process of building the computational model (compare Figures 10 and 12). This led to the realisation that the analysis and formulation components in Figure 10 should therefore be considered to incorporate the world, the place where the generative activities of agents are expressed, alongside the learning mechanisms (that create a model of this world). We will come back to this issue later, in the second version of the distributed learning control model.

5.1.2 Results and reflections

In this model of coordination as distributed learning control the capacity to generate designs is attributed to the distributed control and learning functions. The problem for each controller is to generate structural or functional formulations and hence steer the overall design description towards individual temporal targets. The controlling ability is tightly linked to the neural networks, which are able to gradually construct new knowledge about FBS *interdependencies* from the cases/examples presented to them. In this regard, the creative ability of the agents is largely related to the possibility to discover and learn novel

[4] The focus on these models was the interaction among the three objects within a neutral (void) space. More realistic models could include interactions that are extended beyond the three objects alone (e.g. with existing buildings) and pose further restrictions and requirements. In modelling terms, there are two different directions for achieving this: to incorporate a model of the environment in a knowledge base for the agent, or to equip agents with the ability to recognise their environment at any given time.

interdependencies among FBS variables. Creativity also depends from the construction of new FBS *variables* and new goals for the individual design activity of agents which facilitate the co-evolution between problem and solution spaces. The existence of disturbances is then desirable to the degree that they may constitute a source of variation and an opportunity for novelty for individual agents.

As the control and learning functions were operationalised in a computational environment however, certain technical problems emerged with important theoretical implications for the understanding of multi-agent design using the abstraction of distributed learning control. These problems prevented the full implementation of the particular version of the model and although simulation results were not derived from it, important insights were produced from the process of building it. These are the important results of the experimentation. Let us explore the problems and discuss the theoretical insights in more detail.

5.1.2.1 Modelling learning in multi-agent design

We have discussed that it is essential for coordination to endow agents with the ability to create a model of the world (i.e. acquire knowledge about the designed object and its behaviour, as well as knowledge about other agents' actions and their effects). This knowledge is useful for the generation, evaluation and control of distributed design solutions.

Reflections about the technical implementation

Technically, the adopted reference model control architecture, like most architectures which contain a system identification phase, uses an offline training algorithm. More specifically, once the neural network is trained at a sufficient level with a sufficient amount of data, it can then be used online for real-time control. Similarly, if the controller does not reach its target within sufficient time, then it should be succeeded by the goal reformulation process (which is also implemented as a neural network learning mechanism). And so the process goes, until a coordinated solution can be found. The conceptual model of coordination does not make assumptions about when the different processes may alternate, but instead considers that these are mutual processes which continuously inform one another. However, in order to implement the model, it is necessary to introduce some sort of external performance criterion for switching between the learning process and the control process, as well as for switching between synthesis and reformulation. This is necessary even in cases where online learning can be used. The problem is that we do not have a theoretical basis for introducing such criteria, and hence define in advance what is the sufficient amount of data for training, or the sufficient time for the control process, etc.

One alternative considered for structuring the succession of the control and the goal-reformulation process was to implement the reference model as a pattern recognition or classification network. The idea is that again goals can be formed from observations of the controlled object, but in this case the learning process is based on assessing these observations in terms of similarity between them and the initial goal. This solution may relieve us of the task of having to devise criteria for goal reformulation, but it is still an arbitrary mechanism.

By producing various simulation models one can experiment with different settings so as to find an optimal setup, and define performance benchmarks for these processes of control, learning and goal reformulation. This can offer some useful insights about how the coordination process may work in reality: for example, one can derive observations and general measurements about the optimal duration of learning relative to control, or the average frequency of goal reformulation. We should also, however, be cautious about the way we interpret, translate or generalise such results. In a computational implementation it is not possible to refer to learning (or any other process) in an abstract way. Specific choices have to be made for the model, and the results and conclusions will have to refer specifically to these choices. For instance, in the experiment, artificial neural networks were chosen for modelling learning; neural networks have a long tradition in machine learning and are used not only because of their technical features but also because they are generally considered to be plausible biological metaphors for human learning. The results should then refer to this particular type of learning, and can only be generalised to a certain extent. Moreover, it is true that all learning processes which are implemented computationally depart from reality in many ways: for example, the time frames of neural network learning, or evolutionary learning, are completely different in real design tasks than in computer simulations. These are, of course, ordinary questions in science, particularly experimental science. The important question here is to understand what sorts of assumptions can or should be made about learning, and what kind of exploration is necessary for the purposes of the thesis.

Assumptions about learning in design research

Let us explore for a while the issue of learning as it appears in the design literature. Researchers focus on understanding the peculiarity of learning in the domain of design in order to compare and contrast the properties of design reasoning with relation to other categories of human reasoning, such as scientific reasoning (Cross, 1985). Understanding how learning occurs in the design process, and identifying learning types and styles of designers, has important implications for design education and practice, but also for the development of appropriate CAD systems. For some relevant literature see Demirbas and Demirkan (2003), Felder and Silverman (1988), Kvan and Yunyan (2005), Lawson et al (2003). The area of machine learning in design is also an important focus of investigation for those interested in developing intelligent design systems. For a representative sample of this work see three special issues in the Artificial Intelligence in Engineering Design and Manufacturing Journal (AIEDAM) edited by Maher, Brown and Duffy (1994), Duffy, Brown and Maher (1996), Duffy, Brown and Goel (1998), and for some more recent advances see for example, Liu, Tang and Frazer (2001, 2002), Reffat (2000), Sim and Duffy (2004). Although there is an agreement among design professionals and researchers that learning is a crucial aspect of design (particularly creative design), and despite some empirical evidence that certain types, or styles, of learning may be more characteristic, more important, or more successful than others, there is to date no comprehensive, generally agreed, theory about the learning processes that are, or should be, in play during design. Also it is reasonable to assume that different learning processes or styles will be active in a distributed design task.

Moreover, the requirement of this thesis is that the abstraction and model of coordination are developed in such a way as to account for design as a distributed social process. Again, as we saw in Chapter 2, learning is considered a critical dimension of multi-agent design. However, although some work exists that looks at learning in teams (e.g. Shih, Hu and Chen, 2006; Stumpf and McDonnell, 2002; Wu and Duffy, 2004), most of the literature is focussed on individual learning. Even theories and models that do pay attention to the role of the environment or situation within which design occurs are still predominantly focused on the individual designer as the unit of analysis. An important and useful exception is research which adheres to the ideas of distributed cognition[5]. Very generally, distributed cognition considers that the locus of cognition is not the individual mind, but a wider system composed of agents interacting with each other, and with external representations, tools and artefacts (e.g. Salomon, 1993; Hutchins, 1995; Hollan et al, 2000). Proponents of distributed cognition in design consider that the process of design is effectively formed and developed within such a socio-technical context (e.g. Arias et al, 2000; Fischer and Nakakoji, 1997; Fischer, 2000; Perry, 1997, 1999; Zhang and Norman, 1994; Norman, 1993). In these, alongside the notion of distribution, the notion of learning features as an important aspect. From the above list, Fischer and colleagues in particular, highlight the importance of learning in solving complex design problems and maintain that social creativity essentially rests on learning as 'collaborative knowledge construction'.

Edwin Hutchins who established the idea of distributed cognition, also initiated a comprehensive research programme including empirical studies (Hutchins, 1995), as well as computational modelling (Hutchins and Hazlehurst, 1995). In design however, most research is focussed on ethnographic studies of groups, and/or the development of computational tools with an emphasis on the sort of equipment and interfaces that can support communication, interaction, and reflective

[5] It is worth noting that alongside distributed cognition, approaches such as situated action and activity theory also represent an effort to overcome the traditional focus on cognitive agents by incorporating the socio-cultural context of work. Although each approach has its merits, distributed cognition is closer in spirit to the present thesis because it places the notion of distribution at its core and adopts a systemic view of design phenomena. For general overviews and comparisons of the various approaches, as well as pointers to relevant literature one may consult Nardi (1995) and Susi and Ziemke (2001). For an overview in the context of design in particular see Perry (1997).

reasoning (i.e. HCI research).

The issue of social learning

From this brief review, it becomes apparent that a proper appreciation of **social learning** in design is needed. In cognitive science, social psychology, social science, education and other domains, there are two distinct ways in which social learning is usually considered: a) as individual learning influenced by social context, and b) as learning of social entities (i.e. learning at a 'social level'). In the first case, the focus is on understanding how individuals learn through their participation in social groups, and how social structures may guide, constrain or facilitate individual learning. In the second case, the focus is on understanding how social entities themselves (such as teams or organisations) acquire, construct and share knowledge. Salomon and Perkins (1998) distinguish six different meanings of social learning: 1) active social mediation of individual learning (i.e. through instruction, tutoring etc), 2) social mediation as participatory knowledge construction (i.e. situated or constructive learning), 3) social mediation by cultural scaffolding (i.e. learning through use of tools and artefacts), 4) the social entity as a learning system, 5) learning to be a social learner, and 6) learning social content. The authors further pinpoint a very important question, that of understanding how 'individual and social aspects of learning interrelate and interact in synergistic ways' (ibid pp 2). This is a question of connecting the 'cognitive, acquisition-oriented' view of learning with the 'situative, participatory' view. In the context of the present thesis this seems to be indeed a very significant clue for reconciling the typical cognitive view of design with the conception of design as a social process.

Outside design research, modelling and simulation of social learning has been the focus of many studies (e.g. Boyd and Richerson, 1985; Zentall and Galef, 1988; Flinn, 1996; McElreath et al, 2005). Social learning is usually perceived as an imitation process and modelled as copying and acquisition of new behaviours or strategies in an evolutionary way. Social learning is recognised as an important feature of (intelligent) agency, and there is a general agreement that social phenomena are the outcome of a mutually constructive interplay between individuals, social entities, and environmental parameters. Despite this agreement, the question of how we can model this mutual influence between individual-cognitive processes and social processes is still open. For some relevant literature see Conte and Paolucci (2001) and Gilbert et al (2006).

Individual and social learning in the proposed model

Exploring the interrelation between social and individual learning is largely undeveloped in design, but it is extremely important for the purpose of this thesis. Since distributed design decision making is fundamentally a social process, it is imperative to understand the role of social learning and the way individual goal-driven learning and social learning co-exist.

The model presented here proposes the existence of two kinds of learning. On the one hand the model incorporates learning processes which are used for the purpose of control. This includes a process that captures knowledge about the world and informs the control process and a process by which the controller learns to achieve a specific target. The two processes work together so that they enhance the ability of each individual agent to control its environment (i.e. the ability to control depends on the comprehensiveness and precision of the world model). Note that the elicitation of knowledge about the world is motivated by the desire to achieve a particular goal. On the other hand, the model also incorporates a learning process for creating a reference model: this knowledge is again captured from observations of the world, but it represents a tool for evaluating and adapting individual goals. This process can be thought as a kind of social learning, where agents observe their environment and become attuned to it. As we saw the two kinds of learning processes inform one another: the control process contributes to the formation of the reference model and the reference model provides new targets for control. The suggestion of the model is that the existence of these two types of learning is important for understanding how individual and social learning may co-exist, becoming an integral feature of coordination. This does not in any way offer a final, or complete, resolution to the question of what is learning in multi-agent design and how it happens, and so further work is certainly needed. We will however come back to this issue again later in this chapter, as well as in subsequent chapters.

5.1.2.2 Distributed learning control and creativity

The main hypothesis put forward here was that coordination can be formalised using the mechanics of distributed learning control. We have already extensively explored how learning and decentralised control (two of the three dimensions of multi-agent design) are embedded in the model. The definition of coordination laid out in this thesis holds that co-evolution of the problem and solution spaces (the third dimension of multi-agent design) is a necessary condition and indeed one that is tightly linked to creativity. The second important aspect of coordination as approached here is its conceptualisation as an emergent state and process, which again relates to the issue of creativity. It is time to consider co-evolution and emergence in more detail.

Modelling co-evolution

To discuss how co-evolution is approached in the context of distributed learning control it is instructive to make a comparison with the alternative model of co-evolution proposed by Maher et al (Maher, 1994; Maher and Poon, 1995; Poon and Maher, 1997; Maher, 2000). Maher suggests that co-evolution is an appropriate model of design in cases where the focus of the design process may change and the space of requirements (problem space) is developed together with the solution space through mutual interaction. She contrasts co-evolution to the approach of design as search, where the problem requirements are well-defined in advance and the focus of design remains unchanged, and to design as exploration, where a change of focus in the requirements space may direct the search process to different parts of the design solution. This model of co-evolution is illustrated in Figure 22.

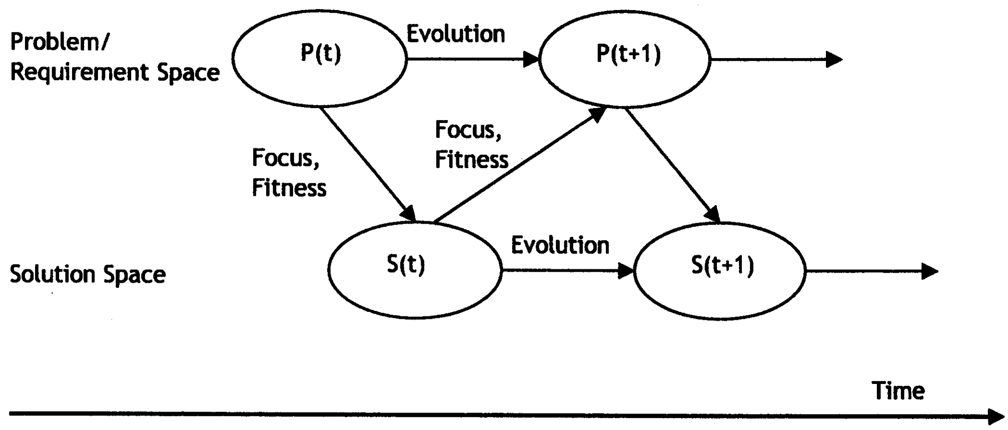


Figure 22: The model of design as co-evolution proposed by Maher and Poon (1995). The model incorporates two search spaces, a problem space P and a solution space S, which evolve horizontally in time and provide goals for one another. The downward diagonal arrow represents a search process by which a problem leads to a solution and the upward arrow a process where a solution refocuses the problem.

The main idea behind the model, is that the two spaces provide the focus for one another: the downward arrows represent the evaluation of solutions using the requirement space as a focus (the design goal), and the upward arrows represent the evaluation of requirements using the solution space as the source of focus (now the solution space provides a goal for search in the problem space). The horizontal dimension represents evolution over time.

In computational terms the co-evolutionary process is divided into two phases that inform one another iteratively. Each phase corresponds to a search process implemented as a simple genetic algorithm (i.e. in each phase, the focus or fitness function remains unchanged). In fact, the authors describe two alternatives for implementing the co-evolutionary algorithm: by using a single composite genotype, or by using two different sets of genotypes and phenotypes. The process starts from some initial problem requirements that provide the goal for searching the solution space (first phase). When convergence is reached and if the termination condition is not satisfied, then the focus shifts and a new search process starts - this time in the problem space (second phase).

The authors observe that unlike what happens in typical genetic algorithms, convergence and termination are distinct processes. Convergence is the criterion for changing the focus, or stopping the search in one phase (for example no new genotypes can be found). The termination criterion by comparison specifies when the overall process should stop. Various termination criteria are discussed, for example time constraints or when some sort of equilibrium is reached (see Poon and Maher, 1997: 321). Another important question is related to the change of focus. The change of focus is expressed by the transformation of the fitness function. Although change is desirable, it seems reasonable that some kind of continuity or coherent association with the original goals should be maintained. Here the proposition is to expresses fitness in a way such that the initial (problem or solution) requirements are concatenated with the new (best) requirements in each phase. The balance between keeping or changing goals is ultimately defined by the designer of the program.

The similarities and differences between the model of distributed learning control and co-evolution can now be discussed more lucidly. Both models consider two interacting mechanisms that provide goals to one another and generate alternatives that satisfy a desired correspondence between problem and solution formulations. While in Maher's model the basic mechanisms are driven by the evolutionary laws of inheritance, selection and random variation, the basic mechanisms in the model presented in this thesis are driven by the principles of feedback, learning and control. The most important difference however, is that co-evolution proceeds by adopting one goal at a time (whether this refers to the problem or the solution space), whereas the model proposed here considers a distributed goal system where different goals are explored in parallel. In other words, the model of coordination described can be said to extend the concept of co-evolution so as to account for the mutually constraining and generative processes that are developed in distributed design decision making settings. It is worth noting that although the relation between agents is modelled as a kind of reciprocal control, the effects of control actions can lead to divergence, as well as grouping, between the agents, their goals, and the solutions they generate. In this approach, it can also be argued that creativity is expressed at two different levels: at the level of the individual agents, who are able to generate and explore new problem and solution spaces, and at the level of the distributed system as a whole, where new collective problem and solution spaces are created.

It is also important to take notice of an important issue arising in both cases. Both models endorse the idea that (creative) design is an intentional or goal-oriented process, but also that the formation of the goals themselves is part of designing. In the co-evolution model, goal formation is achieved through exploration of the problem and solution spaces and the use of a transformation criterion, such that a desired balance between old and new goals is achieved. In the coordination model, goals are formed through a dual learning process, such that novel structural and functional features are generated and incorporated by way of generalisation (or pattern formation). In realising any of the two models in computational terms, we are faced with a similar problem: the criterion, or mechanism, for deciding the balance between keeping or changing goals is arbitrary. It is (externally) defined by the designer of the program. Arguably, one could suggest that a meta-level process is established within the system that effectively evolves this mechanism (the fitness function, learning mechanism, or any other method we could potentially devise and use). Although this would solve the problem of how to incorporate in the model the definition of the criterion that drives the goal formation process, we will be left with the problem of how to define the meta-level process itself in a non-arbitrary way - and so on ad infinitum. The essential difficulty is that we do not have a theory to tell us what the process of goal formation really consists in, so as to be able to reproduce it computationally. This is not an issue where the objective of the computational model is to solve a particular problem; in this case any mechanism is appropriate and valuable so long as it achieves its objective. The issue of course arises when the modelling task aims to represent or replicate the design process itself; in this case, even if the mechanism works this does not offer verification that the mechanism is 'the' one, but only that it is a sufficient one. However, examination of alternative mechanisms through simulation is indeed extremely valuable because it allows exploration and analysis of the various hypotheses and helps make the underlying assumptions transparent and precise.

Design and emergence

Creativity and emergence are two interlinked concepts since they both refer to the creation of something new. They are both desirable when they contribute to finding a solution to a problem that is elegant or economical, or in general, they lead to the creation of an object with added value. In the design context, emergence is usually associated with the spontaneous discovery of some new attribute of the design description or artifact and is predominantly studied in relation to the generation and visual recognition of shapes (see for example Knight, 2003). The present thesis approaches the issue of emergence in design in a wider way, using insights from multi-agent systems and complex systems science.

The typical view of emergence in multi-agent systems defines it as a new pattern or behaviour that is formed at the macro level from the interaction of agents at the local level. In multi-agent systems, emergence is usually desirable when it leads to behaviours that can achieve a certain task or goal without the existence of a central mechanism directing the activities of agents towards this end. In this sense, we can consider that coordination is exactly the desirable emergent effect of the activities of the distributed agents. We may also argue that in this way coordination becomes an indication of the creative ability of agents. Castelfranchi maintains that for 'self-organizing emergent structures' to appear, there needs to exist some kind of feedback from the collective phenomena to the individual mind. This feedback can be facilitated either through evolutionary/selective mechanisms or through some form of learning (Castelfranchi, 1998: 179). Along these lines we can consider that interaction and learning are the keywords to creative complex systems and these have been the focus in the development of the distributed learning control model.

As mentioned previously, in the model presented, creativity is associated with the ability of agents to discover new values and new relationships among FBS variables and therefore enhance the problem and solution spaces. Another way to express creativity, which would more closely reflect the notion of emergence, is to consider how agents may extend the original definitions of FBS variables. Such an action requires agents to be able to modify the definition and complexity of the objects they manipulate. This could be achieved for instance by incorporating new function variables (Mixed Housing, Recreation etc), or by adding new objects or even by sub-diving the initial objects (cuboids) - see for instance Figure 23. This would also introduce novel interdependencies and constraints for the multi-agent design process.

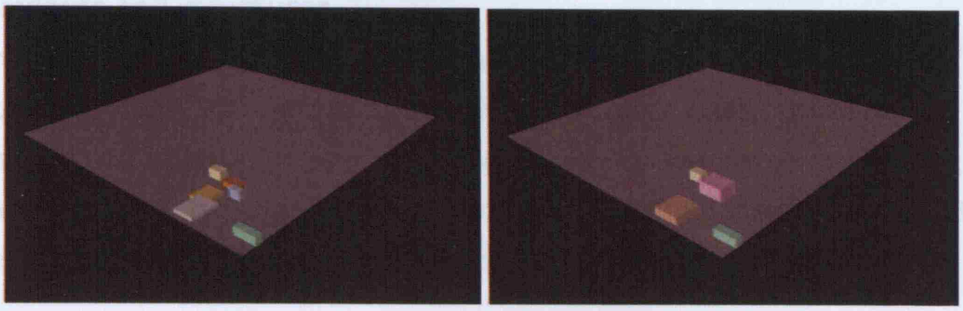


Figure 23 Emergence involves redefining the original FBS variables by adding or merging different cuboids and/or introducing new functions. Coordination then suggests a process of organising agents, their goals and the objects they manipulate.

In the introduction to the notion of distributed learning control it was discussed that each agent is able to control parts of the overall design and is also to a certain extent 'controlled' (constrained as well as enabled) by the actions of others. Re-defining FBS variables, means therefore re-defining the interdependencies among agents and the way they are grouped together to achieve their goals. The type of emergence suggested here implies that coordination involves a kind of hierarchical structuring (organisation) among agents, their goals, and the objects they manipulate. To understand coordination in these terms requires a deeper understanding of emergence and the formulation of a theoretical framework that is apposite to the problem of multi-agent design decision making. The issue is an important one and is therefore reserved for further exploration in Chapter 7.

5.2 COORDINATION AS DISTRIBUTED LEARNING CONTROL: VERSION 2

5.2.1 Building the distributed learning control model

5.2.1.1 The design problem and the representation of functions, behaviours and structures

In the second experiment an alternative, quite simplified, interpretation of the conceptual model was explored. The design problem under consideration was transferred to an architectural scale and defined as a question of devising the layout of a building. Again, the setting involves three agents, each with a different (initial) goal for the configuration of the building plan.

The space where the actions of the agents are expressed was realised by using the representation proposed by Steadman et al (Steadman, 1998; Steadman and Waddoups, 2000). The 'archetypal building' is an abstract representation of a large class of rectangular built forms. This abstract form is composed from three types of spaces defined according to the type of lighting they incorporate. The archetype therefore contains spaces that are naturally lit, spaces that are artificially lit (like corridors) and spaces lit from the top (top-most courtyard floor and courtyards). The archetypal built form is illustrated in Figure 24.

Steadman et al showed that from this abstract form and by using a series of transformations (basically subtraction) a very great variety of rectangular forms can be produced, corresponding to different types of buildings found in the real world. The nucleus of the archetype is a subset of the overall form and contains one courtyard. The authors suggested an encoding of this form by using a 14-digit binary string, of which the first 7 digits represent positions in a horizontal axis (x axis) and the remaining 7 digits represent positions in a vertical axis (y axis). By assigning a value of either 0 or 1 in any such position, one can add and subtract strips to create different forms (see Figure 25).

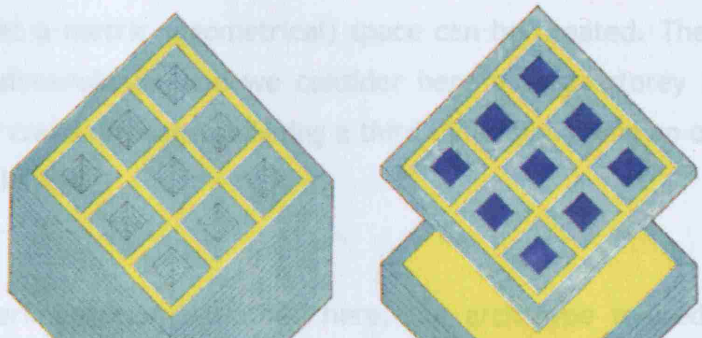


Figure 24: The archetypal built form by Steadman (1998). Yellow represents artificially-lit space (e.g. corridors), green represents naturally side-lit space, and blue represents naturally top-lit space.

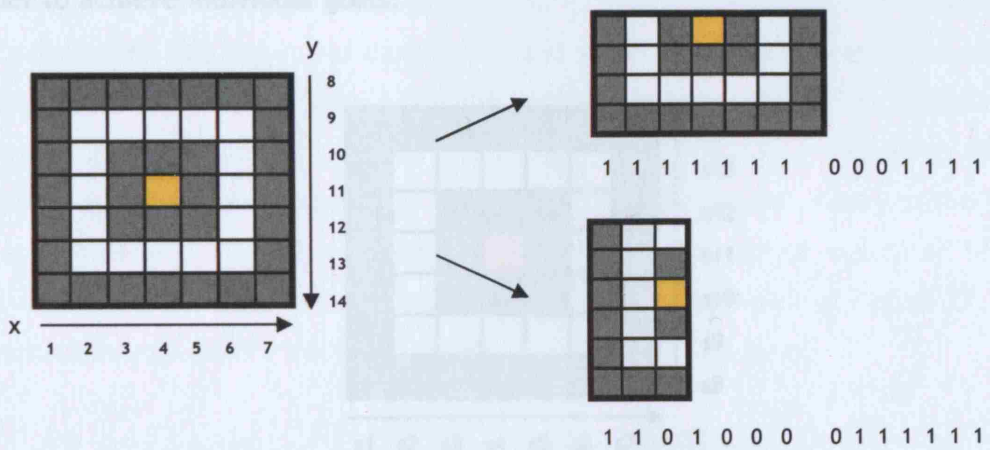


Figure 25 The binary encoding of the archetypal building proposed by Steadman and Waddoups (2000). The 14-digit binary array can be used to describe the geometry of rectangular built forms as shown, for example, on the right (where 1 represents the presence and 0 the absence of a strip in the x and y axes). Here white refers to artificially-lit space, grey to side-lit space, and yellow to top-lit space.

The representation adopted for the experiment described here (Figure 26), is slightly different from the original in that counting of the y positions (or strings) starts from the bottom. Also, more importantly, in this experiment, each position is not defined by a binary value, but instead takes a value in the range between 0 and 1, so that a metric (geometrical) space can be created. The representation used is two-dimensional, and we consider here a single-storey building layout, although one could envisage assigning a third dimension along an orthogonal axis z to represent height.

For the experimentation described here, the archetype is used as a space of possibilities within which agents' proposals can be developed. An important reason, for which the archetype was used, was the fact that it captures functional and structural features within a single representation. As mentioned, by definition of the archetype, the structural (topological) dimensions are defined according to their function (lighting). This was considered desirable for the experiment because it made it possible to simplify the control architecture. Behavioural information therefore is captured as change of functional and structural characteristics in order to achieve individual goals.

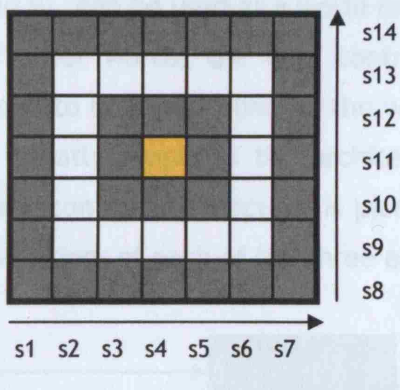


Figure 26 The representation of the archetype used in the experiment, adapted from Steadman and Waddoups (2000).

It should be noted that in contrast to the computational model in Version 1 where goals were represented in terms of evaluations or if-then rules about structural and functional relationships, goals in this model are represented directly in terms of archetype configurations. The idea is that different configurations express in principle different preferences about the organisation of space in relation to lighting, but these preferences are not explicitly articulated. This allows to not only simplify the control architecture, but also to express situations where initial ideas about architectural solutions come in the form of rough plan layouts. The knowledge behind the process of formulating and reformulating goals is tacit and distributed learning control here can be thought as a kind of 'design reasoning without explanations' (see Coyne, 1990). The downside of this representation of goals is that it makes the semantics of the design object implicit, and seems to place a somewhat unbalanced emphasis on the manipulation of abstract forms as the core of the design process.

5.2.1.2 The control architecture

The model was built in MATLAB, this time using the program's editor (not the SIMULINK library) to design the neural networks and the control algorithm.

The architecture adopted uses two neural networks: one to act as a controller and another to act as the world model. Since functional and structural information is encoded using the same representation, the same neural network (which learns associations between S, F and B), can be used as a world model in the analysis and reformulation processes. In other words, the dual control architecture of the previous experiment is reduced to one. Additionally, the world model is also used as a reference model. This greatly simplifies the architecture described in the previous experiment. The new control architecture is pictured in Figure 27. The architecture represents the workings of each of the three agents.

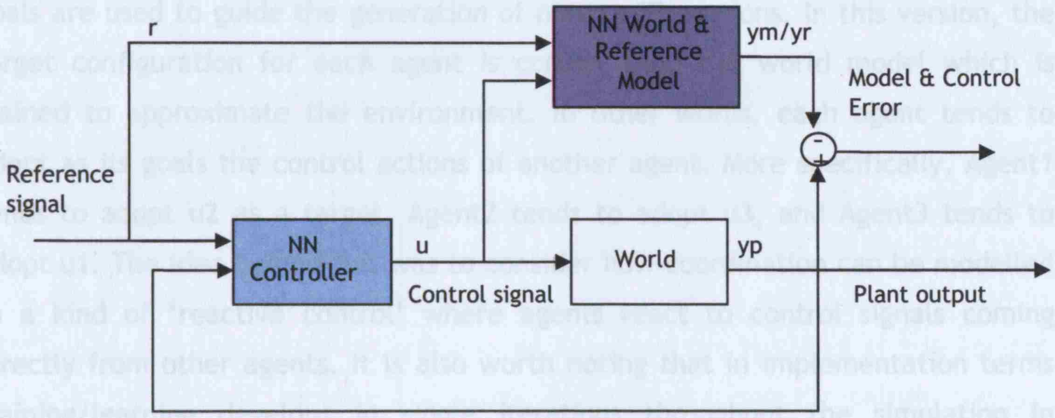


Figure 27: The control architecture used in the experiment, which represents the workings of each agent. The World Model is also used as a Reference Model.

More specifically, the experiment is set up as follows. The three agents interact with each other in a sequential way (Figure 28). Each agent uses the first neural network (NN Model) to learn input-output pairs of descriptions, that is, associations between its own proposals (control actions) directed to the next agent and the replies from the agent two steps down the line. So, what constitutes the observable world for one agent is a composite of the two other agents. Namely, the world model of Agent1 is formed by learning associations between u1 and u2. Similarly, the world model of Agent2 is formed by associations (u2, u3) and the world model of Agent3 by associations (u3, u1). The second neural network which acts as a controller essentially learns the inverse of the world model.

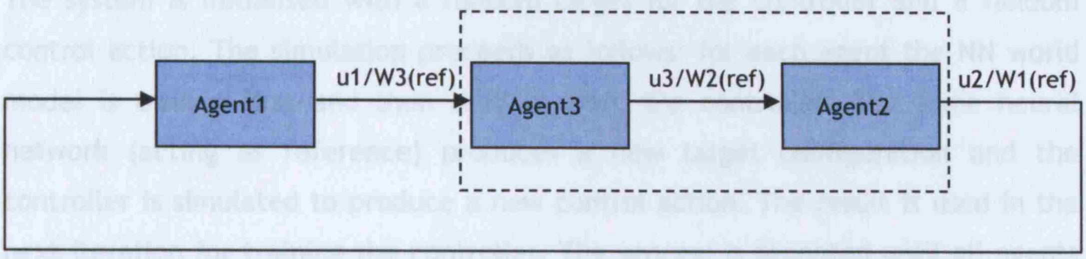


Figure 28: An illustration of the sequential connection between agents. The world for one agent is a composite of the other two.

An important difference with the previous model is found in the way goals are formed. As mentioned, an agent uses the world model also as a reference model, namely in order to generate new goals for the configuration of space. The new goals are used to guide the generation of new configurations. In this version, the target configuration for each agent is coming from the world model which is trained to approximate the environment. In other words, each agent tends to adopt as its goals the control actions of another agent. More specifically, Agent1 tends to adopt u2 as a target, Agent2 tends to adopt u3, and Agent3 tends to adopt u1. The idea behind this was to consider how coordination can be modelled as a kind of 'reactive control' where agents react to control signals coming directly from other agents. It is also worth noting that in implementation terms training/learning develops in single iterations throughout the simulation in contrast to the previous model where the learning process had to be carried out before the control process.

More specifically, the inputs for each agent are presented as 14-digit arrays corresponding to geometrical configurations (instances of the archetypal building). Both the NN Model and the NN Controller are standard 2-layer feedforward networks that have 32 neurons in the first layer and 14 neurons in the 2nd layer (to correspond to the dimensions of the output array/archetype). The hyberbolic tangent sigmoid ('tansig') transfer function is used for the first layer and the log sigmoid ('logsig') transfer function for the second layer (which squashes the output values into the range 0 to 1)[6]. For detailed information about the neural networks, in MATLAB code, see APPENDIX 1.

The system is initialised with a random target for the controller and a random control action. The simulation proceeds as follows: for each agent the NN world model is trained first and then used to train the controller. The same neural network (acting as reference) produces a new target configuration and the controller is simulated to produce a new control action. The result is used in the next iteration for training the controller. The process is repeated until all agents have achieved their targets - for the simulations a specific simulation time of t=100 was set, which was sufficient for the completion of the process. The MATLAB code of the model is also provided in APPENDIX 1.

5.2.2 Results and reflections

Figure 30 shows results from a simulation run for time t=100. The control actions of the three agents are displayed in a sequence: u1, u2, u3. The vertical succession denotes time (simulation cycles). In the first cycle (t=1) the control actions result from the simulation of each agent's controller given an initial world W (Figure 29) and three random inputs - one for each agent. In the second cycle (t=2), the results are obtained after the controller is trained for the first time using the control actions and targets/expectations from the previous cycle. Due to the control architecture, in all subsequent iterations the topology of the control action of each agent remains the same, although the specific values change.

[6] The tansig transfer function is $a = e^n - e^{-n} / e^n + e^{-n}$ and the logsig transfer function is $a = 1 / 1 + e^{-n}$ *where n is the net input.*

Compare for example the control actions for t=2 and t=37 in the figure. From t=37 to t=100 also the geometry (the values) of the control actions remains the same for all agents. This precisely reflects the control algorithm as described above, which determines how the control actions of one agent become targets of another.

This dead end situation is due to a very important reason. The setting of the experiment equates the world at any given moment with the control action of each agent active at that moment. Hence, the world is not a common space where the control actions (and design proposals) of agents can be collectively expressed with all their mutual or conflicting effects. In particular, the notion of disturbance introduced in the conceptual model at the beginning of this chapter is completely eliminated. In this sense the task of creating a world model (a model of each agent's actions) is duly straightforward. This brings us back to the discussion about the importance of the world as a generative component. The role of disturbance in this respect was overlooked in the experiment. Although disturbance is a source of noise in the face of which agents have to learn to plan their control actions, it is also an important source of variation. Coordination is significantly reliant on the existence of disturbance, **a direct effect of the fact that design is distributed.** Other agents may be causes of conflict, but may also bring new knowledge, new resources, and new opportunities for creativity. In the model of coordination proposed here, disturbance works to an extent as a counterpart of the idea of random variation which we encounter in evolutionary and co-evolutionary models of design.

Additionally, by conflating the world model with the reference model in the control architecture the agents are forced to adopt as their goals the control actions of others. In effect, agents simply exchange goals and no progress with the design coordination process can be made. This experimentation however was valuable, so far as it clearly highlighted the importance of incorporating a separate component for providing reference goals, and modelling the coordination process in such a way that a balance between changing and keeping previously formed goals is maintained appropriately.

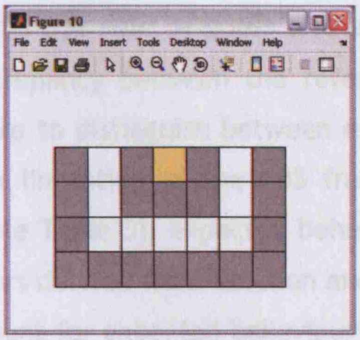
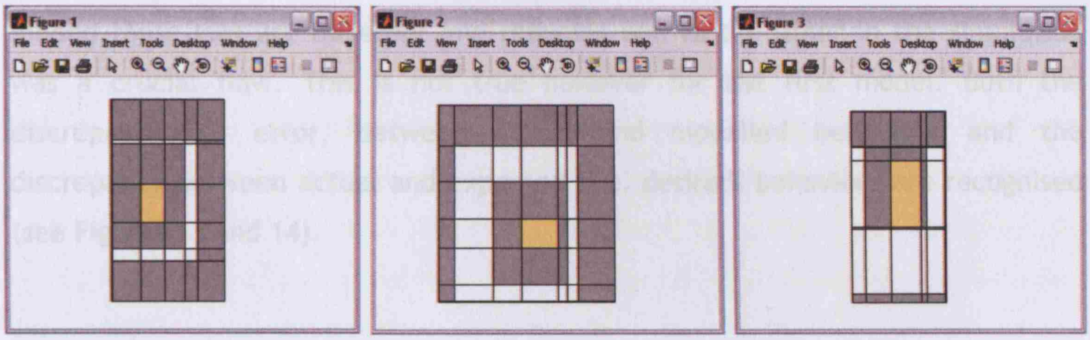
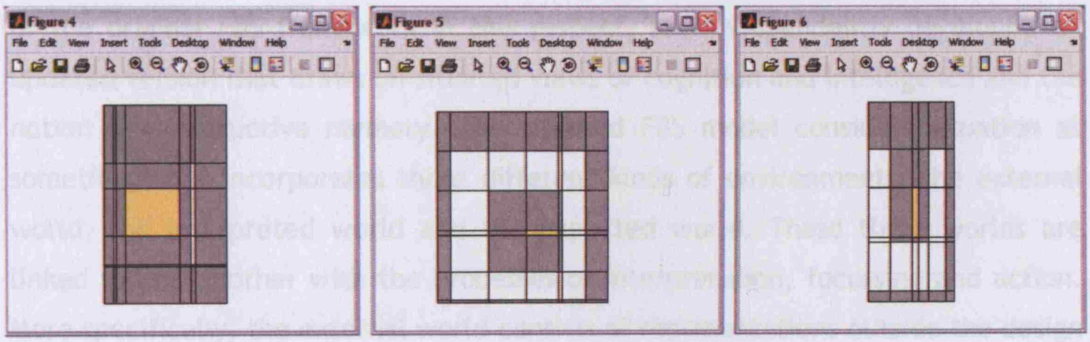


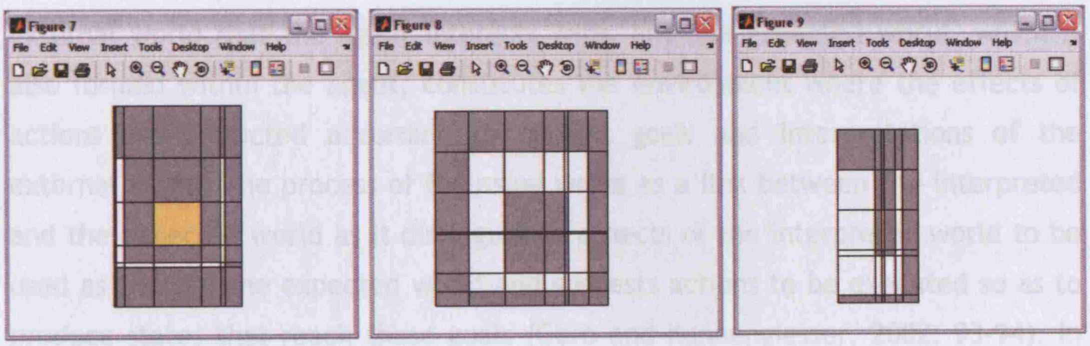
Figure 29 The initial world provided to the three agents for the simulation.



Results after 1st cycle (t=1). The figures represent u1, u2 and u3 from left to right.



Results after 2nd cycle (t=2).



Results after 37th cycle (t=37). The control action of each agent has remained the same.

Figure 30: Results from a simulation run for time t=100. The figures illustrate the control actions of the three agents, given an initial world (Figure 29), and for times t=1, t=2 and t=37, from top to bottom.

5.2.2.1 Coordination and the role of goals and expectations

The omission of the discrepancy between the reference model and the world model is linked to a failure to distinguish between expectations and goals and is more deeply rooted in a limitation of the FBS framework used. As described previously (see for example Table 5), expected behaviours in this framework are considered to be behaviours derived from function and are therefore uniquely tied to goals. There is no account for expected behaviour as behaviour derived from a model of the world, or knowledge about the effects of other agents' actions. Indeed these two are different and reducing one to the other in the simulation was a crucial flaw. This is not true however for the first model. Both the discrepancy, or error, between actual and modelled behaviour and the discrepancy between actual and expected (i.e. desired) behaviour are recognised (see Figures 13 and 14).

Gero and Kannengiesser (2002, 2004, 2006) have indeed realised the shortcomings of the original FBS framework in this respect, and independently developed an updated version that draws on situated views of cognition and intelligence and the notion of constructive memory. The situated FBS model considers situation as something that incorporates three different kinds of environments: the external world, the interpreted world and the expected world. These three worlds are linked to one another with the processes of interpretation, focussing and action. More specifically, 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; and the expected world, which is also formed within the agent, 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 (Gero and Kannengiesser, 2002: 93-94). In this framework the formation of expectations is considered fundamental both for the creation of internal representations and for the construction of memories. This idea extends the original framework with an additional set of processes thus considered to be involved in designing.

It is worth noting that 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. This notion is consistent with Schön's (1983) arguments about reflection in action. The situated FBS model is primarily focussed on reflective reasoning as a characteristic of individual design agents, although an attempt has been made to account for social aspects of design (Gero and Kannengiesser, 2003). In that paper the ability of agents to interact socially is considered to originate in their ability to create internal representations of the world and potentially adapt their interaction according to this knowledge. The discussion offered, however, is quite limited since the specifics of agent-to-agent interaction are not explained or modelled in detail. By contrast, the coordination model proposed in the present thesis, uses the idea and mechanics of distributed learning control in order to account for conflicts and interdependencies developed between agents.

5.3 REVISITING THE MODEL OF COORDINATION AS DISTRIBUTED LEARNING CONTROL

Taking into consideration the arguments explored above, it seems necessary to revise the conceptual model presented at the beginning of this chapter with a more appropriate diagrammatic representation, where the difference between expected and desired behaviours is made explicit (Figure 31). The new representation also makes clear the idea that the analysis and formulation components incorporate the world, the place where the actions of agents are expressed to generate designs.

In the coordination model, therefore, expectations express knowledge about future states of the world and are formed on the basis of observations of current and past states of this world[7]. The acquired knowledge is used to guide the generation and selection of control actions able to satisfy given goals or desires.

[7] As each agent has a partial view of the world and bounded capabilities for learning, this knowledge is of course partial and incomplete. Perhaps a better characterisation would be to refer to beliefs rather than knowledge.

The reference model is also created on the basis of knowledge about the world and observations about the consequences or effectiveness of actions. But while expectations are used to infer control actions, goals are used as a way to evaluate the boundaries of the control actions or, in other words, to delineate desired possibilities for the world that is being designed. As discussed, the creation of goals, in parallel to the creation of design solutions, is a very vital requirement for creative design in general and coordination in particular, where these processes are distributed. We should keep in mind that goals in this model are formed both for the synthesis (solution formation) and the reformulation (problem formation) processes. What seems to be especially important, however, is not only to understand and model the transformation of goals in the distributed design process, but also to give an account of how the function that evaluates these goals is modified in the process (i.e. the mechanism that decides when and how these goals should be transformed).

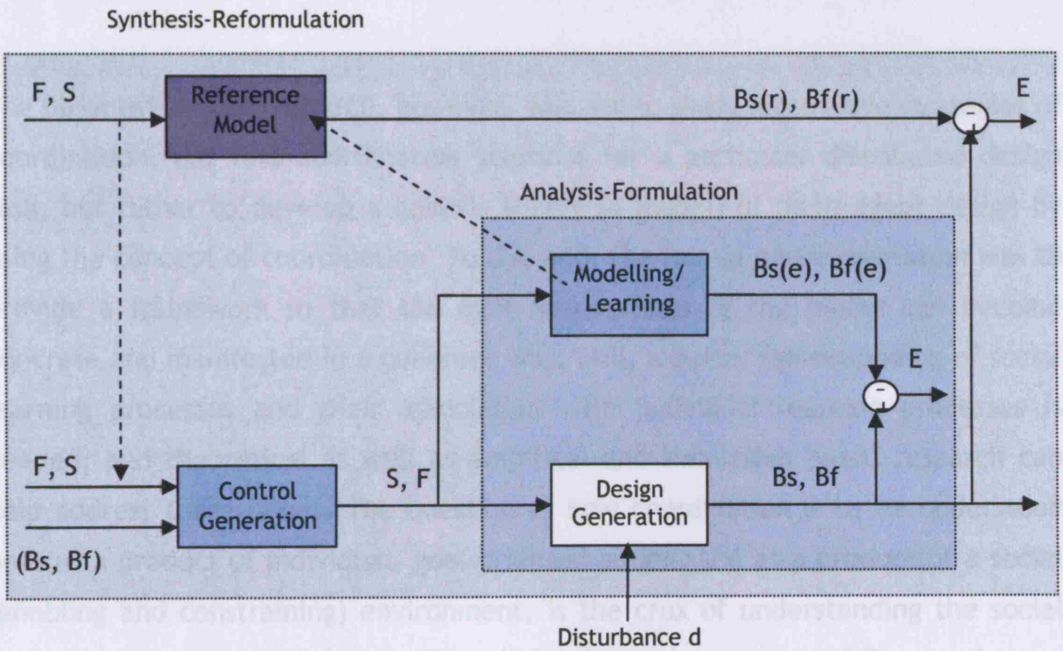


Figure 31 The updated model of coordination as distributed learning control. F=Function, S=Structure, Bf=Behaviour derived from function, Bs=Behaviour derived from structure, Bf(e) and Bs(e)=Expected Behaviours, Bf(r) and Bs(r)=Desired Behaviours, E=Evaluation.

As already mentioned, in the model of coordination proposed here, the reference model, which is responsible for the transformation and evaluation of goals, is developed on the basis of a kind of social learning. That is, the knowledge of agents is formed through observations of the world within which their own as well as others' actions are expressed. Although neural networks were employed in the experiments, other types of learning mechanisms could be appropriate and are worthy of exploration. The question of how the social context may influence agents' decisions and goals is crucial and considering alternatives for expressing this idea would be valuable. For example, an interesting experiment would be to include a representation of the social network of agents (a representation of the way agents are connected) and develop a process such that this network influences the formation of agents' knowledge and goals and (possibly) vice versa. This experimentation would be useful in exploring various hypotheses about successful modes of social learning, successful strategies for goal adoption, or factors for achieving better coordination in different cases.

The objective of the research, however, was not to study alternative strategies of coordination, nor find coordination solutions for a particular distributed design task, but rather to develop a generic theory (a model) of multi-agent design by using the concept of coordination. To this end, the role of experimentation was to provide a framework so that the main assumptions of the model can become concrete and manifested in a coherent way. Still, a better understanding of social learning processes and their association with individual learning processes is needed, and theoretical as well as empirical and simulation based research can help address these issues. The question of how coordination is to be understood both as a product of individual, goal-oriented action, and as a product of a social (enabling and constraining) environment, is the crux of understanding the social character of multi-agent design. This will be the subject matter of Chapter 6.

5.4 SUMMARY AND CONCLUSIONS

The reasoning behind this chapter was to use simulation in order to make the dimensions of coordination previously identified explicit and more precise. In particular, the purpose was to explore whether and how distributed learning control can be used as a model for capturing these dimensions, and reveal potential benefits and drawbacks. These ideas were elaborated by way of a series of computer models and simulation experiments. The experimentation highlighted various implementation problems but more importantly some conceptual issues.

In the first experiment, the dual control process was modelled using a model reference control architecture. The first insight gained from this experiment was that control, learning and generation of design alternatives are vitally linked to each other. Both control and learning constitute generative functions so long as they contribute to the creation of new interdependencies among design variables and new goals for the design process. Generation is made possible by the fact that agents act in a common world, and this common world is a source of knowledge and opportunities for creativity. This highlighted the importance of including in the model of coordination an explicit linkage to the external world. The experimentation also helped draw attention to the importance of better understanding learning as a crucial element of coordination, particularly as to its role in facilitating and constituting design as a social process. Additionally, a discussion on the relation between coordination, emergence and co-evolution was also developed. This discussion offered a clearer understanding of how co-evolution is embodied in the model, and exposed the need to investigate the concept of emergence in a way apposite to the problem of multi-agent design decision making.

In the second experiment, a simpler form of the control architecture was explored. This eliminated the initial idea that a separate component is needed to act as a reference model. It also eliminated the idea that agents act in a common world, thus disposing of the notion of disturbance. Both were proven to be serious flaws. The experimentation corroborated the previous discussion about the importance of incorporating the influence of the external world in the model. This is a space where the actions of agents are commonly expressed and although it provides agents with a source of conflicts, it also provides them with opportunities for new knowledge, and the creation of new design solutions. The experiment also made it possible to understand that a clear distinction between goals and expectations is conceptually necessary. Moreover, by distinguishing between goals and expectations, and the learning processes that formulate them, we may start understanding how social and individual goal-driven behaviours become intertwined in a distributed design task.

Finally, based on the lessons learned from the experimentation, a revised model of coordination was proposed. In all, two main issues for further development were identified. The first is that a clearer understanding of the social character of multi-agent design is needed, particularly through understanding of the interplay between individual and social learning. The second is that an appropriate understanding of the concept of emergence is needed, one which is specific to multi-agent design. These two issues are explored in the following chapters.

Chapter 6

THE SOCIAL CHARACTER OF MULTI-AGENT DESIGN: COORDINATION AS A MICRO-MACRO LINK

In Chapter 2 a review of coordination in different domains was offered with the aim to create a framework for developing a concept of coordination suitable for multi-agent design. It was discussed that such a concept should include purposeful action, as well as an account of the social character of the design process. Learning, distributed control, and co-evolution were proposed as characteristic dimensions which can help us formalise the main principles of coordination. The experimentation in Chapter 5 revealed that the learning processes linked to the formation of goals and expectations in the proposed model of coordination are important for understanding the interplay between goal-oriented behaviour and social learning. The current chapter explores in greater depth the question of how design can be understood both as purposeful construction which leads to the creation of socially recognised artefacts, and as a product of social activity, which is formed and constrained by social structures. The problem of understanding how the macro level of society is related to the micro level of individual agents is a fundamental one in sociology and it is crucially linked to the concept of emergence. The chapter discusses this issue in order to unveil the theoretical and epistemological difficulties with establishing a view of design as societal ability and point towards possible avenues for resolution.

6.1 AGENTS, SOCIETIES AND THE MICRO-MACRO LINK

Understanding and defining how the macro level of society is causally related to the micro level of individual agents is an age old question in sociology, sociology of knowledge and social psychology, but has also recently been raised in the domains of distributed artificial intelligence, multi-agent systems and social simulation. In fact, the question of resolving or reconciling micro and macro, agency and structure, is at the very heart of sociological theory as it refers to the problem of explaining social reality and its origins. The term 'micro-macro link' reflects a dichotomy between approaches that ascribe primacy to individual action (typically represented by rational economic theory) and approaches that ascribe primacy to social order and collective structures (typically represented by structuralism). The problem is to explain the ontological status of social structures, where they come from, and how they are sustained, but also how individuals take part in the process, and whether and how they shape or are shaped by society. The investigation is fundamentally linked to an understanding of emergence as a way to explain the interdependence between the two. We will focus here on emergence only as it relates to this question, and reserve a more rigorous exploration of the concept for the next chapter.

6.1.1 Approaches to the micro-macro link in social theory

Alexander and Giesen (1987: 14) distinguish five major approaches to the question of the micro-macro link, which correspond to different positions about the nature of agency (rational versus interpretive) and the source of social patterns (individual versus collective) - see Figure 32. More analytically, approaches (1) and (2) both take an individualistic view, but the first stresses the rational and objective character of action, while the second puts emphasis on subjective psychological factors and processes of interpretation. On the other hand, approaches (3), (4) and (5) take a collectivist position but have different attitudes towards the role and status of individual agents. At one end, approach (5) considers that social control shapes and constraints action, and denies any subjective perception of order, while approach (4) accepts subjectivity and purports that socialised individuals reproduce society by translating the existing structure into the micro realm. Approach (3), which actually rests somewhere in between the other approaches, gives analytical autonomy to the micro level by recognising that the socialised individual re-creates society during the process of reproducing it. The various contributions to the synthesis of the macro with the micro usually build on distinct combinations of these approaches.

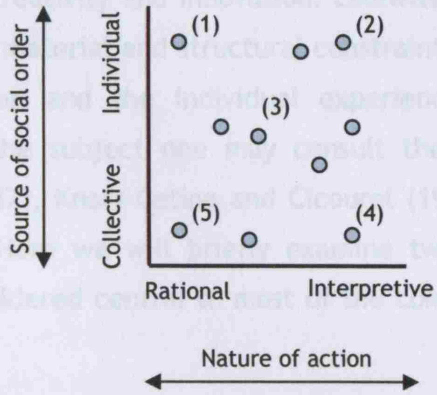


Figure 32 Different approaches to the micro-macro link in social theory correspond to different positions about the nature of agency (rational vs interpretive) and the source of social order (collective vs individual).

There are also many conceptions of what the micro and the macro actually are in analytic terms, each coupled with a different understanding of the processes of creation, change and reproduction. Münch and Smelser (1987) distinguish various interpretations of the two levels. These may range from cases where micro refers to individuals and small social units, to cases where the micro is expressed as interactions, empirical indicators or psychological propositions. Similarly the macro may be perceived as the level of populations, or as a level that refers to repeated experiences, large scale interactions, or constraints of interaction. The authors also review various works on linking micro and macro. In attempts at linkage that move from the local scale to the global scale, they identify several modes of transition: a) aggregation, b) combination (of microinteractions with other variables), c) externalisation, d) creation, sustaining or reproduction of macro, and e) conformity. Different modes of transition are identified also in the move from the global scale of society to the micro level: a) internalisation, b) limiting laws and rules (by means of instruments like contracts, division of labour, or institutionalisation), and c) repression.

Finding a link between the micro and the macro or a middle ground between agency and structure signifies a need on one hand to explain the stability of societal structures despite individual action or violation, and on the other hand to explain variability, creativity and innovation. Likewise, it is important to account for the existence of material and structural constraints and influences, as well as for purposeful action and the individual experience of reality. For a more complete view of the subject one may consult the collections of articles by Alexander et al (1987), Knorr-Cetina and Cicourel (1981), and also Huber (1991) and Ritzer (2000). Here we will briefly examine two theories that are widely referenced and considered central to most of the contemporary discourse around this issue.

6.1.1.1 The theories of Giddens and Bourdieu

The first theory that will be explored here is Giddens' theory of structuration (1984). The focus of investigation in this case is neither individual experience nor society as a totality, but human social practices which are (self-) reproduced in space and time. Central to this theory is an understanding of the 'duality of structure'. Structure is seen as both the pattern of social practices and as a 'virtual order' of rules and transformative relations. So structure consists in a set of rules and resources, better perceived as 'structural properties' that constrain and at the same time enable social activities.

Giddens uses a stratification model to establish agents as knowledgeable actors in this framework (Figure 33). Agents reflexively monitor their actions and aspects of the contexts within which they move, have a 'theoretical understanding' of the grounds of their activity, and embody a potential for action (motivation). Through these processes agents reproduce the conditions that make their practices possible and continuous in space and time. So the relation between agent and structure is also dual: 'the structural properties of social systems are both medium and outcome of the practices they recursively organize' (Giddens, 1984: 25). This is key to understanding the stability of structures. However, Giddens also argues that while agents reproduce the rules and resources on which their activities are based, agents are also always *bounded* in their capacities. The *unintended consequences* of actions play an important role in bounding, or *conditioning*, further activities and thus constitute an important source of structural variation.

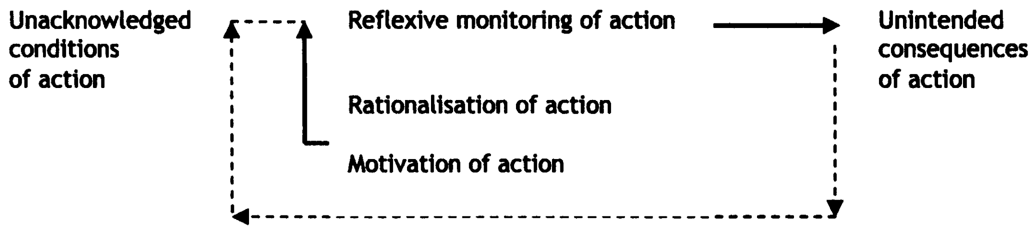


Figure 33 An illustration of Giddens' (1984) stratification model showing how agents reflexively monitor their actions and establish an understanding of the grounds of their activity. Through these processes agents reproduce the conditions of their practice. The unintended consequences of actions play a role in conditioning further activities.

The second theory we will consider here is that of Bourdieu (Bourdieu, 1977; 1984). Bourdieu's key concepts, habitus and field, are both concerned with relations. The field constitutes a space where different forms of capital (power or resources) are distributed and represented as *objective (historically defined) relations between positions*. The habitus on the other hand, is a relational configuration which results from the *internalisation of objective relations* within individuals, in the form of 'mental and corporeal schemata of perception, appreciation and action' (Bourdieu and Wacquant, 1992: 16). These schemata are a form of embodied knowledge which disposes individuals to perceive and act on the field according to their positions in it. The relation between field and habitus is obviously dual. The field feeds and conditions the habitus, but cognitive activity (the habitus) (re)constructs the field as a meaningful world. Habitus is thus an *open* system of dispositions acquired through previous experience that can be used as a source for the construction of a variety of social realities. It is important to note that although individual and collective action is constructive, the categories of thought applied in construction are historically constituted and thus lie in an objective (social) realm, that is, agents understand the social world because they are socially produced and structured.

Both Giddens and Bourdieu are primarily interested in explaining why (and how) agents happen to reproduce society while being knowledgeable and deliberate, and how society conditions action. They both prefer to adopt a notion of structure as a set of transformation rules, conditioning properties, or structuring mechanisms rather than equate it with a rigid delimiting entity or constraint. They also both take important steps towards conceptualisations that seek to dissolve the dichotomy between agency and structure by guaranteeing a form of embeddedness of actors in the social world.

We can understand from these examples that the dichotomy between agency and structure is akin to the dichotomy between system and environment. In the discussion about emergence in complex systems we saw that the dissolution of this dichotomy is crucial for understanding, and that such a dissolution can be achieved by shifting focus onto organisational characteristics and descriptions of phenomena. Our social theorists here also attempt to develop organisational descriptions of social systems. Giddens's description of the reproduction of social practices is reminiscent of descriptions of self-organised or autopoietic systems. Similarly, Bourdieu adopts the stance of 'methodological relationism', and takes (power) relations and their representations in human experience as his unit of analysis.

6.1.1.2 Micro-macro link and emergence

Now let us return to the concept of emergence and reflect on what these two theories have to offer to the discussion. Giesen (1987) discusses how the concept of emergence becomes central to the problem of linking the micro and macro levels. He identifies three problems: the problem of *descriptive emergence* (finding a scientific description of how the micro and macro are linked), the problem of *practical emergence* (accounting for the practical inconsistencies or discrepancies between micro and macro as experienced in everyday life) and the problem of *explanatory emergence* (providing an explanation for the independence of the two levels and their change processes). The crucial point to hold on to is that considering emergence in social systems means that an explanation is sought both as to how the macro-level is created, and also as to how the macro-level becomes autonomous and how it feeds back into the micro-level (has causal powers over the micro). This is captured by the notion of downward causation mentioned earlier that calls for incorporation of top-down, constraining processes into our models and theories of emergence.

The structuration theory of Giddens and the habitus-field theory of Bourdieu are particularly useful in understanding how the macro (society) can have causal powers over the individual agents without it being an external absolute form of control or constraint; but their account of the micro part is somehow limited to concepts of re-production or re-construction. In other words, the causal effects of intentional action do not receive substantial attention. What is more, none of the two approaches actually appear to target or address emergence specifically.

It is interesting to contrast this situation with the situation of research in DAI and MAS. The same problematic about linking micro and macro seems to have developed from a completely opposite direction. Starting from a very strong tradition of exploring emergence by adopting the bottom-up approach of micro-sociology and focussing exclusively on agency and individual (local) action, a considerable literature has developed which seeks to incorporate mechanisms and explanations for downward causation. It is because of the traditional focus in MAS on the micro that the concept of downward causation becomes so important.

6.1.2 Approaches to the micro-macro link in multi-agent systems and social simulation

Conte and Gilbert (1995 - Introduction) consider the issue of emergence to be central for computer scientists involved in the simulation of social phenomena and the creation of artificial societies. They point out three problems with emergence in MAS: a sub-cognitive bias (the focus on reactive or sub-symbolic systems has meant that emergent phenomena that can be produced by cognitive agents have been ignored), a behavioural bias (non-behavioural emergent effects have not received appropriate attention), and an individualistic bias (the study of social phenomena has tended to follow only the direction from the micro to the macro). Although properties and phenomena that pertain to the macro domain, such as social norms and conventions, have been considered in the literature, it is often the case that they are formalised so as to correspond to a constraint that does not feed back into individuals or become part of their apparatus. For an overview of related work and main issues regarding the understanding and modelling of social order see (Conte and Castelfranchi, 1995; Conte and Dellarocas, 2001).

Schillo et al (2000) focus specifically on the question of how the micro-macro problem is perceived in DAI and they identify four misconceptions about social phenomena which they unearth on the basis of knowledge derived from a review of sociological discourse. One of the problems they identify is that most studies reduce macro-level design to mechanism design (design of interactions among agents) and do not deal with the full variety of sociological concepts covering interactions among different levels of organisation (from teams, to institutions, to society as a whole). The second criticism is that DAI studies focus solely on emergent behaviour (as global behaviour or pattern) which is not sufficient as an account for the structuring properties of macro-level entities that affect the micro level. The third problem is that analysis usually proceeds by aggregating values from the micro to the macro which excludes the use of proper structural interpretations (causal relationships, measurements of trust or influence etc). Their final criticism is against the view that populations of agents make artificial societies; this is intended to highlight that a collection of agents is not sufficient for the notion of society and an account of macro aspects is necessary. The authors also argue that a computational simulation is a sort of hybrid society since humans (the designers of the simulations) provide artificial agents with their capacities and goals. In short, the issue seems again to be a limited understanding, representation and modelling of the macro and its effects on the micro, and a lack of conceptual and methodological tools for addressing the linkage between the two.

Gilbert (2002) documents a variety of emergence types according to the way the relations between micro and macro are treated. The first type refers to the emergence of a feature at a macro level which comes as a consequence of the local rules obeyed by individual agents (*first-order* or *bottom-up emergence*). The second type of emergence is *downward causation* and refers to the case where a macro feature constrains or influences individual action. Finally the third type, which is called *second-order emergence*[8], occurs when agents are able to detect emergent features and adjust their action. Gilbert and others (e.g. Gilbert and Troitzsch, 1999; Gilbert, 1995, 2000; Conte and Gilbert, 1995; Castelfranchi, 1998) have emphasised the need to advance social simulation by taking into account the reflexivity of human institutions and societies and the ability of people 'to recognise, reason about and react to human institutions, that is, to emergent phenomena' (Gilbert and Troitzsch, 1999: 11). This sort of classification clearly points to a consideration of different levels of emergent phenomena that mediate the transitions between micro and macro. Moreover, the acknowledgement of second-order emergence shifts attention to exploring, and including in simulations, aspects of communication and language as well as abilities of recognition, interpretation and learning.

Sawyer, in a series of recent papers (2000, 2001, 2003, 2004a, 2004b), also offers a comprehensive review of MAS simulations and their relation to sociological theories and explanations. He focuses in particular on the question of downward causation, or how macro properties may become ontologically distinct from the micro interactions that realise them and may in turn take part in causal relationships. Despite the name, his version of downward causation is equivalent to second-order emergence as discussed above, given that internalisation of macro patterns is also here considered to be an essential requirement. Sawyer terms his philosophical stance 'nonreductive individualism'. His criticism of MAS simulations is exactly that they do not include downward causation: even when some form of social structure and influence is incorporated, this is usually pre-wired in the system and is invariably not internalised by individual agents. The requirements for modelling this kind of emergence are therefore first, ensuring a 'bi-directional dialectic' between micro and macro and second, developing a complex communication system that can support reflexive processes and particularly reasoning about the communication process itself.

Castelfranchi (2001) suggests that in order to reconcile micro with macro teleology it is necessary to develop theories of cognitive agents (in terms of goals

[8] The term second-order emergence appears to have been used first by Steels (1994).

of beliefs) alongside models of emergent functionality (in terms of unintended mechanisms). Castelfranchi proposes that the link between micro and macro necessitates the development of a theory of social function. Let us consider Castefranchi's view in more detail.

6.1.2.1 Castelfranchi's theory of social function

Castelfranchi (2001) view considers that an emergent macro-structure, i.e. social order, exists objectively when it has a causal effect. A 'strong' conception of emergence in fact requires that a macro-structure reproduces *because of* these effects. Castelfranchi essentially frames the question of linking micro and macro as a question of defining and explaining social *function*: how can emergent effects reproduce through intentional action, but 'independent of the agent understanding and pursuing these effects'? (pp 20) The model illustrated in Figure 34 explicates how certain unintended effects of intentional action are reproduced because they reinforce the beliefs and goals that caused them.

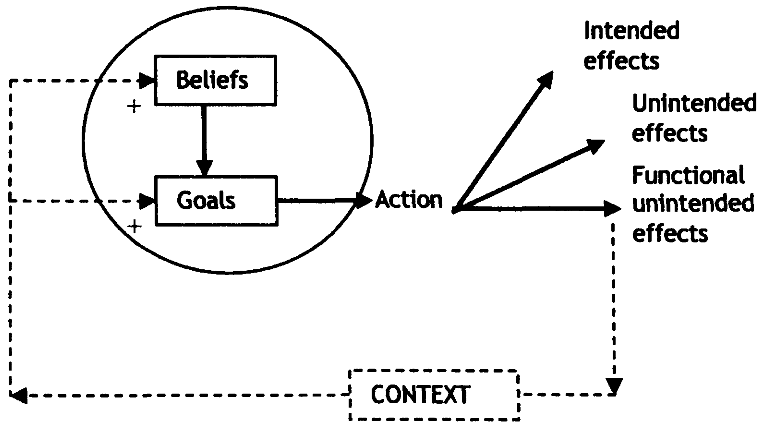


Figure 34 Castelfranchi's (2001) model showing how unintended effects may acquire a social function because they feedback into individuals: they reinforce the beliefs and effects that caused them.

Castelfranchi's view is similar to Giddens's in proposing that unintended consequences of actions play an important role in conditioning (here *selecting* or *reinforcing*) further activities. But he additionally dwells on aspects of individual cognition and action. He further argues that the appropriate mechanism for linking intentional action with playing social functions is some form of 'learning without understanding'. He discusses three different mechanisms through which such a form of learning can be achieved. The first is reinforcement learning, where the beliefs and goals of agents are reinforced (either the beliefs and goals are confirmed, or the association between beliefs, goals, and the context within which they operate, is strengthened); the second is restating, where the contextual conditions that lead to the action are sustained or re-produced; and the third is emotional reinforcement, where possible goals are 'pre-selected' because of learned emotional reactions to anticipated future scenarios.

Castelfranchi's treatment, which builds upon evolutionary conceptions of function, offers a way to study the relationship between agents and the collective macroscopic reality. Not only does it allow us to account for the circular causality between micro and macro, but also for the independence of one from the other. This is the reason why he is particularly focussed on the question of how cognitive intentional agents may participate in the reproduction of social order without being aware of doing so. He distinguishes two ways in which agents may be playing a social role: the convergent way (functionality impinges on the intended effect) and the divergent way (functionality is derived from the side-effects of intended action). Nevertheless, in the model illustrated in the figure, only the second type of behaviour is incorporated. Moreover, while he recognises the importance of relating and distinguishing intentionality and function (he considers this to be 'the hard problem'), he doesn't show in his model how social order may be created/reproduced not only through functional unintended effects, but also through deliberate creation/reproduction (i.e. <u>the relation</u> between intended effects and functional unintended effects a propos the creation of social order is not discussed).

Coming from a cognitive science background, Castelfranchi is understandably focussed on issues of individual cognition and action, and is primarily concerned with the development of (formal) models of individual and social cognitive agents. His work is part of ongoing research and debate about the necessary characteristics and capabilities of social agents, especially with a view to implementing them computationally. For an overview of typologies, models, architectures and technologies for social agents see also: Sun, 2001, Carley and Newell, 1994; Ekdahl, 2002; Jennings and Campos, 1997; Verhagen, 2000; Dautenhahn, 1998, 2000).

This discussion points back to the issue of learning discussed in Chapter 4 and which became a central component of the conceptual model of coordination and the accompanying simulations. Here, learning appears as a critical issue in relation to the micro-macro debate. The question translates as follows: how can we reconcile individual, constructive learning with social learning? By the term social learning we refer to the general idea that the social and cultural context influences learning (e.g. Lave and Wenger, 1991; Wenger, 1998), or that learning is mediated in various ways (e.g. Salomon et al, 1991; Salomon and Perkins, 1998), but we also refer to types of learning that are based on imitation as well as observational or vicarious learning (see Bandura, 1977, 1989; Conte and Paolucci, 2001). While on the one hand proactive individual learning is regarded as one of the hallmarks of agency, on the other hand social learning appears to be a necessary requirement for an explanation of how emergent properties feed back into individuals.

6.1.3 Summary and insights for multi-agent design

In sum, in this section the problem of the micro-macro link was introduced for its relevance to understanding and modelling multiagent design as a social process. We looked at different conceptions of emergence in sociology both from the perspective of social theory and from the perspective of social simulation. The main insight from this review was the need to reconcile first and second order conceptions of emergence, particularly through an appreciation of the role of learning. Design as a social process inherits the requirement for comprehending emergence in this way. It is however instructive to note that there exist some additional requirements for design as a special case of social process. The first obvious characteristic is that design is an action for change. Social activity has many different manifestations and, for example, can consist in plain regulation or preservation of order; design in contrast is driven by the anticipation of creating something new and it is therefore by definition directed towards change. Another characteristic is that while some of the approaches reviewed above consider that social structure inevitably pre-exists individual agents, in design we have to reject this assumption. By this of course it is not meant to say that agents do not embody socially formed norms, conventions, knowledge, beliefs or goals, but only that in any collaborative design undertaking the social structure needs to be created in the process and in response to the particular task in hand.

In the following section, the model of coordination elaborated in Chapter 5 will be used as a framework for developing an understanding of how the micro-macro link can be seen in the context of multiagent design.

6.2 ANALYSIS OF THE MICRO-MACRO LINK IN THE MODEL OF MULTI-AGENT DESIGN AS DISTRIBUTED LEARNING CONTROL

The model proposed in Chapters 4 and 5 had as its basis the idea that we can model multi-agent design as coordination using the mechanics of distributed learning control. The action of change performed by each agent is in essence a control action that tries to guide the world towards a state that corresponds to individual goals and beliefs (and also individual and global constraints). Beliefs are created and re-created through observation and learning of action-effect associations (World Model) and provide the stock for discovering appropriate control actions. Goals are also formulated and re-formulated on the basis of knowledge about the world, as well as about the ability to control (Reference Model). That is, agents reformulate their goals according to constraints or opportunities they discover in the world. The world is a synthesis of the different control actions bound together under resource limitations and other global constraints. To allow comparisons with models reviewed in the previous section, the model is represented in Figure 35.

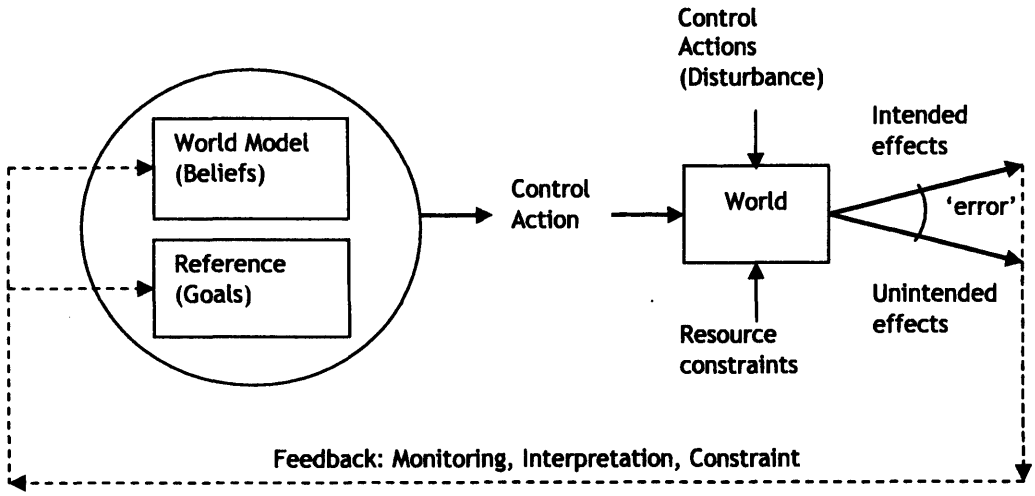


Figure 35 An illustration of the model of multi-agent design presented in chapter 4 of the thesis. Agents operate over a common world by producing control actions formed on the basis of goals and beliefs. Both intended and unintended effects have functional role and their relationship is established through the notion of error.

6.2.1 Comparison with Castelfranchi's model

There are some noteworthy similarities and differences between this model and the one proposed by Castelfranchi. The two models have in common the idea that individual action, which is guided by goals and beliefs, is the basis for the creation of the macro-level, while the macro-level becomes autonomous as unintended effects feed back into individuals through a learning process. The coordination model however, by definition, focuses on the interrelationship between multiple agents and explicitly considers this to be a causal factor for the creation of macro-level structures. Hence, in contrast to Castelfranchi we have here a representation of the world as a field where convergent and conflicting actions are manifested. So agents interact through this world[9] and use knowledge about it to guide their future action.

The model of coordination considers that both intended and unintended (collective) effects contribute functionally to the creation of the micro-macro

[9] As already mentioned this represents a kind of indirect or tacit communication.

dynamics. From the point of view of an individual agent, the perceived distance (or error) between intended and unintended effects motivates action, but it is also used as a metric of the limitations and constraints over that action. In reality, agents are bounded in two ways: one is related to the limitations of available resources, laws, and other external constraints, and the other is related to the ability of an agent to learn, interpret or internalise the external world and therefore form expectations and predictions about it. Although in the simulation experiments described in Chapter 5 only a particular type of (neural network) learning was examined, it should in principle be possible (and it is in truth desirable) to incorporate other forms of learning such as Castelfranchi's 'unaware' type of reinforcement learning, Bandura's vicarious learning, or evolutionary learning.

6.2.2 What is micro and macro in the model?

It is essential here to clarify what is meant by social structure in the context of this work. As mentioned previously (see paragraph on Münch and Smelser pp 133 this thesis), there are many conceptions of what the macro is in analytical terms. The coordination model is focussed on (distributed) design decision making and is therefore primarily concerned with the processes that play their part in the formation of the design space (problem and solution spaces) and the characteristics of this space. In this sense, the study of coordination is the study of how the different dimensions of the problem and solution space are created and organised. The model particularly addresses the question of how distributed goals, requirements and knowledge about a design solution (micro-level) are synthesised into an actual collective configuration (macro-level). Clearly, the way by which agents interact and get grouped together or differentiated (in terms of goals and beliefs), is reflected in the organisation of the design space at the macro-level. So micro and macro do not exactly correspond to the problem and solution space, but rather to their expression in an individual and a global scale respectively.

To better understand this it is useful to recall the discussion in Chapter 5 about the importance of including in the model of multi-agent design, a representation of the common world within which agents interact. As discussed, agents represent individual goals and desires about the world and control parts of the overall configuration. Agents also create a model of this world, by developing knowledge and expectations about the effects of their actions and the responses from other agents. As the world is the product of the (control) actions of agents, every change in the world reflects a change in the relationship between agents.

We can illustrate this idea with the help of Figure 36: let us consider Agent 1 and Agent 2, each manipulating an object within the design space. By changing, for example, the position of the objects from *a* to *c* and from *b* to *d*, agents produce changes in the design space. This means that each control action, although it is applied on a different object, it also affects other objects in the world (as for example, the distances between objects change). Therefore, each control action changes the relationships among agents and the objects they manipulate, even if there is no direct interaction between them.

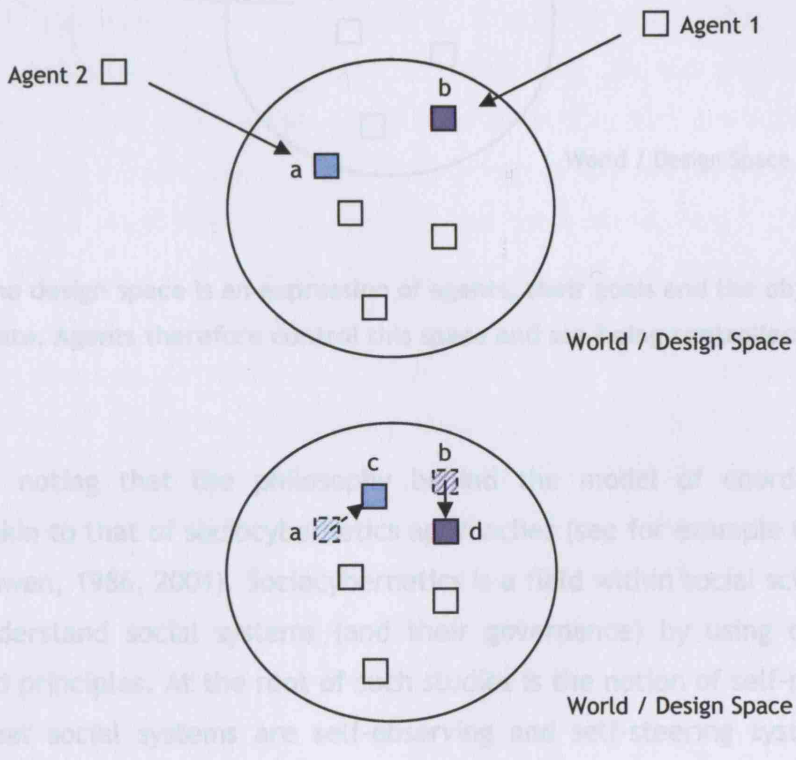


Figure 36 An illustration of the design space as perceived in the model of distributed learning control.

The main point for consideration in this model of design, is that agents are both inside and outside the world; they control it and at the same time are being controlled by it (Figure 37). The social structure developed among agents, is an unintended effect that emerges out of the individual controlling activity of agents, but it is observed by the individuals and acquires the function of enabling and constraining the formulation of their goals and actions. Internalisation of the world and its emergent structures occurs either because agents proactively observe and gain knowledge about the world in order to be able to control it, or because the effects of these structures persist long enough to be understood as undisputable entities (they are there). In the model of coordination proposed in the thesis the micro builds the macro, and vice versa, through the process of distributed learning control.

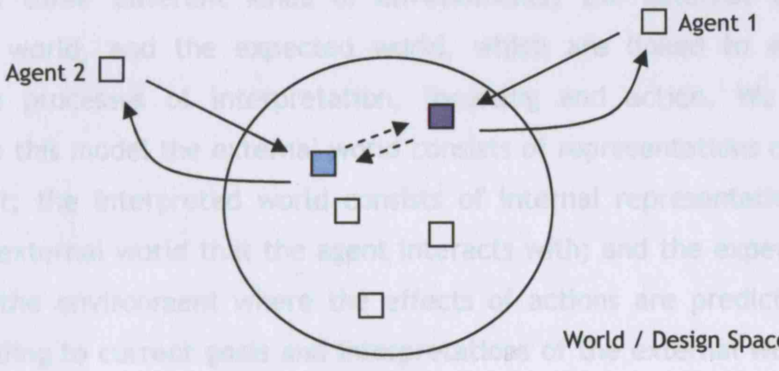


Figure 37 The design space is an expression of agents, their goals and the objects they manipulate. Agents therefore control this space and are being controlled by it.

It is worth noting that the philosophy behind the model of coordination is somewhat akin to that of sociocybernetics approaches (see for example Geyer and van der Zeuwen, 1986, 2001). Sociocybernetics is a field within social science that tries to understand social systems (and their governance) by using cybernetic theories and principles. At the root of such studies is the notion of self-reference: the idea that social systems are self-observing and self-steering systems. The notion of self-reference is a key for design, but the study of design differs from sociocybernetics in that design is not merely about (self-) governance and control, but it is essentially about the anticipation of future constructs and changes.

6.2.3 Comparison with relevant studies in design and planning

In the review of coordination in Chapter 2, we inspected the different conceptions of and ways to study design as a group, collaborative or social process. Although there is a growing body of studies interested in social aspects of design, these are usually concerned with descriptions of work practices or of interactions between designers within a particular setting. Even studies who follow the advances in MAS and social simulation do not specifically account for the problem of reconciling micro and macro causation.

An example is the work of Gero and Kannengiesser (2002, 2004, 2006) on modelling situated design agents, referred to in Chapter 4. Their model incorporates three different kinds of environments, the external world, the interpreted world, and the expected world, which are linked to each other through the processes of interpretation, focussing and action. We saw that according to this model 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; and the expected world constitutes the environment where the effects of actions are predicted by the agent according to current goals and interpretations of the external world. While interpretation is a link between the external and the internal world, the process of focussing works as a link between the interpreted and the expected world: it distinguishes aspects of the interpreted world to be used as goals, and suggests actions to be executed so as to reach these goals (Gero and Kannengieser 2002: 93-94). As already discussed, there are evident similarities with the model of coordination, yet the model proposed by Gero and Kannengiesser is mostly focused on issues of individual cognition, and does not account for agent-to-agent interactions and their role in the formation of the external world. By contrast the coordination model proposed in this thesis, uses the idea and mechanics of distributed control in order to account for conflicts and interdependencies developed between agents.

An example where the problem of the micro-macro link is explicitly considered within a design context is Sosa's thesis (2005). The thesis is motivated by the recognition of a general lack of understanding about the way in which accounts of creativity as individual ability can be integrated with accounts of innovation as a society-wide phenomenon. To work out this problem, Sosa uses computational simulations to explore individual and situational factors of influence between individuals as design agents and the society as a collective evaluation system. Although the general inspiration is similar (dealing with the micro-macro link), the context and aims are different. Sosa is not interested in modelling coordination between micro and macro expressions of the design space, but he is rather interesting in exploring the factors that contribute to a micro-to-macro influence and vice-versa - for example, how individual action may trigger adoption of innovative artefacts, and how societal evaluation may influence the generative design process.

The problem of linking individual action with social structure is a more familiar one in urban planning, where social theory is a more integrated part of the discourse. We have encountered the seeds of this problem in Chapter 2 when discussing the issue of individual versus collective rationality. In planning the question seems to be multi-fold. It not only refers to the problem of understanding and modelling the interplay between individual action and situational, spatial, or societal factors and constraints. It also refers to the problem of understanding the role of bottom-up versus top-down decision making, and dealing with the question of whether one should design for individuals or for collectives (i.e. how to plan and for whom). However, in the context of urban modelling the main focus is not on societal factors like norms or conventions, or on the interplay between individual and institutional or societal constraints, but on the role of spatial and temporal constraints and factors.

From the point of view of this thesis, a relevant example can be found in Portugali (2000). Portugali discusses the view that, in contrast to the traditional position of rational economic theory, individuals' behaviours do not directly follow from their intentions. Following Festinger's theory of cognitive dissonance (1957) he argues that the gap between a person's intentions, wishes and values and what is their actual, demonstrated, behaviour creates a tension 'which eventually will have to be resolved either by change of behaviour and action, or by a change of intentions and value system' (pp 144). Through a series of simulations he explores how this dissonance may lead to individuals adopting a new cultural identity which is manifested spatially (socio-spatial segregation). In these studies, the dissonance is explained as a result of the fact that the emerging order (spatial pattern) 'enslaves'[10] individual behaviour, and corresponds in a way to Giddens's concept that unintended consequences condition individual action. The thesis reported here is in agreement with the statement that the gap between intentions and behaviours (or better, intended and unintended effects) is a significant element for the understanding of the micro-macro bridge. For Portugali however the process underlining these phenomena is self-organisation. Self-organisation in this case translates to the idea that local interactions between agents may lead to the creation of a global pattern that influences individual behaviour (or choice). This idea has become quite common in urban modelling, particularly in relation to the study of spatial segregation, the origins of which can be found in Schelling's contribution (1969, 1978). For a relevant discussion see Batty et al (2004). The interesting difference is that Portugali's account moves closer to the notion of second-order emergence, as the resulting pattern impacts on the value system of individual agents and leads to adoption of a new identity.

It is particularly relevant to reference here O'Sullivan and Haklay (2000) who give an overview of various agent-based models used in planning and other domains of social science. They argue that the majority of these models adopt an individualistic view of the social world which does not account for the macro to micro direction of influence. On this basis, they call for a more careful consideration of the effects of this stance, and a greater engagement with social theory for advancing the potential of these models.

Byrne (2005) also criticises agent-based simulation for a simplistic view of social phenomena, which only focuses on simple rule-based individual behaviour and

[10] This notion is due to Hermann Haken's theory of synergetics (see for instance Haken, 1997). The 'slaving principle' means to describe the circular dynamic process by which local interactions within a system give rise to one (or a few) order states that then enslave/determine the behaviour of the parts.

does not account for historical processes or for the macro-social consequences of individual action. It is worth noting, that although complexity is starting to be recognised as a promising framework for bridging micro and macro (see Urry, 2005), this research programme is still incipient. The work in this thesis (especially through this and the next chapter) attempts to scratch the surface of how advances in MAS, cognitive science, and complex systems science in general can help deal with this issue as it becomes manifested in the context of design.

6.3 SUMMARY AND DISCUSSION

This chapter explored in depth the question of how design can be understood both as an individual and a societal process. This problem, which in sociology takes the form of a debate about the micro-macro link, or integration of agency and structure, was examined in detail in order to derive some basic requirements and characteristics for multi-agent design. The main insight from this was the need to reconcile first and second order conceptions of emergence, particularly through an appreciation of the role of learning. The chapter also offered an analysis of the coordination model proposed in the thesis, explaining how it deals with the micro-macro link. Comparisons with other models, both from the social simulation domain and from the design domain, show that the model offers an original framework for addressing the micro-macro link, which is specifically appropriate for multi-agent design as a distributed decision making process.

There are certainly issues requiring further consideration and elaboration. For example, how different types of learning (individual, social, evolutionary learning etc) can be integrated. Currently a clarification of the role of different types of learning and their interrelation remains an open question, as most studies focus exclusively on one or the other type of learning. Some insight could be gained from certain cognitive science studies where hybrid types of learning and hybrid cognitive architectures are explored (see for example Sun, 2001, 2006). Another issue for investigation is how a more sophisticated view of the micro-macro link can be elaborated, that does not just focus on the two opposite poles (micro and macro) but considers the micro-macro linkage as a multi-level process. Such a development would be useful for better understanding the processes and effects of organisation and the role of meso-level agents, like for instance institutions. The issue is also related to an effort of breaking free from a naïve conception of the micro as simple and the macro as complex. Knorr-Cetina (1981) tells us that this dichotomy is inappropriate from a sociological perspective, and reminds us that complexity is not only a matter of size, but it relates to our choice of the unit of analysis (and its variety), as well as the knowledge of the observer, i.e. the knowledge available for the analysis of a particular phenomenon. The following chapter aims to delve into these issues, and set the foundations for understanding coordination as a multi-level organisation process, by drawing on knowledge from the field of complex systems science.

Chapter 7

A COMPLEXITY PERSPECTIVE ON COORDINATION: UNDERSTANDING EMERGENCE IN MULTI-AGENT DESIGN

At the outset of this thesis a link between design and complexity has been assumed and was followed through as a matter of methodology as well as epistemology. The main assumption behind adopting a complexity view of multi-agent design - reflected in the very choice of the concept of coordination as an abstraction - was that it enables us to focus on systemic rather than individual abilities, and explore relational or organisational properties. In the previous chapters we embarked on unravelling the constraining and enabling relationships and interactions between agents, as well as with the external world, that are important for understanding and characterising multi-agent design. We also gained some insights about the organisational principles that make multi-agent design a social activity. In all the preceding discussions we saw that emergence is a crucial aspect of coordination related to concepts such as co-evolution, creativity and sociability. In order to develop a truly organisational perspective of coordination it is necessary to investigate the concept of emergence more rigorously. In this chapter we are going to review how emergence has been perceived and defined in the context of complex systems, in order to identify key issues for an understanding of emergence in multi-agent design. This treatment will form the basis for developing an organisational perspective of multi-agent design as coordination that extends beyond the model of distributed learning control and places it within the general context of complex systems science.

7.1 THE CONCEPT OF EMERGENCE IN COMPLEXITY SCIENCE

Emergence is a critically debated issue in many disciplines, from psychology and cognitive science, to biology and physics, to social science and philosophy. It has been central in great scientific and philosophical discussions about the mind-body problem, holism, irreducibility and the independence of sciences. As a concept, emergence has been linked with 'high level' functionalities and capacities like cognition, autonomy, intelligence, language, life and sociality, and recently has become the centre of attention for the science of complex systems. In fact, emergence has come to be regarded as the epitome of complex systems, so definitions of emergence are very closely linked to definitions of complexity. For more comprehensive reviews see O'Connor and Wong (2002) (focus on philosophical issues) as well as Bonabeau et al (1995a) and Corning (2002) (focus on AL and complexity).

The concept of emergence is customarily associated with the dictum 'The whole is more than the sum of its parts'. The dictum was originally articulated by Aristotle (*Metaphysics 10f-1045a*), later adopted by general systems theorists (Bertalanffy, 1968), and is now embraced (in different degrees and forms) by complexity scientists. In reality, almost every single word in this statement has been, and continues to be, the subject of rigorous debate and dispute (even the word 'sum': see Kubik, 2003). The appeal of the concept of emergence comes from the demarcation of a particular distinction (or relation) between wholes and parts (or between the macro and the micro); this is also where all the difficulty with the concept comes from. The statement considers that parts and wholes must be somehow different (summing up the parts doesn't simply produce the whole, hence the whole is qualitatively different from the parts), yet there must be a relation between them such that the parts somehow make up the whole (they are 'parts of' it).

7.1.1 Epistemological and ontological types of emergence

Variations on the main theme include the idea that the properties of the system as a whole cannot be reduced, deduced or predicted from (knowledge of) the properties of the parts. The question of **reduction** may have epistemological or ontological status (Peacocke, 2003: 189-190). *Epistemological* reduction asserts a position on the nature of knowledge about complex systems such that laws and concepts used to describe macroscopic patterns and properties are considered reducible to laws and concepts used to describe microscopic entities and properties. Emergence in this case is when such a reduction is not attainable and therefore emergence materialises as a characteristic of the limitations of human knowledge or the descriptive apparatus employed. We can distinguish some views here that focus on irreducibility from others that focus on unpredictability (O'Connor and Wong, 2002). Stephan (1999) names these two different types of emergence 'synchronic' and 'diachronic' respectively.

Basing the argument on the grounds of knowledge raises the issue of whether emergent patterns and properties are 'in principle' irreducible or unpredictable. Would a complete knowledge of parts and their properties eliminate emergent phenomena? Could emergent phenomena be explained or described a posteriori even if they could not have been predicted a priori? These are questions related to ontological assumptions. *Ontological* reduction asserts that complex entities can be completely dismantled to their most basic (physical) constituent parts. Ontological emergentism by contrast denies this assertion and attributes genuine novelty and distinct characteristics at the macroscopic level. There are various subtle arguments against this version of emergence (see for example Kim, 1999), but we can roughly distinguish two main difficulties. First, if no extra distinct characteristics can be shown to exist at the macro level (i.e. the macro properties supervene[11] over the micro properties), then emergent phenomena become epiphenomena. On the other hand, considering that emergent phenomena cannot be derived from their constituents makes the argument somehow extra-scientific.

7.1.1.1 Weak emergence

To avoid the conundrum, Mark Bedau (1997) proposes adoption of the concept of weak emergence (in contrast to strong emergence). His definition (for a system S) states that a 'macrostate P of S with microdynamic D is weakly emergent iff P can be derived from D and S's external conditions but <u>only</u> with simulation' (ibid -

[11] The term supervenience was first used by Morgan (1923) to express the way emergent properties bear upon their base properties. The contemporary meaning of the term is attributed to Davidson (1970) who used it to express the dependence of mental characteristics from physical ones (i.e. a response to the mind-body problem in philosophy). For more on supervenience see also Kim (1993) and McLaughlin (1997).

emphasis added). This version of emergence is weak because it applies to a wide range of phenomena and considers that emergent properties are in principle predictable. The main idea is that although macrostates are fully and solely constituted from microstates and their dynamics, they cannot be derived in any other way except from simulating them directly (i.e. they are computationally irreducible). On the other hand, understanding general principles and laws of macro level patterns and phenomena requires empirical observation at the macrolevel and therefore macrostates are considered to be somehow autonomous from underlying processes. In a similar vein, Darley (1994) adopts a definition of emergence in relation to computational and predictive complexity where he evokes the connection between computational irreducibility and formal undecidability.

The problem of emergence finds others completely rejecting (ontological) emergentism. Epstein (1999) takes the reductionist computationalist view of emergence so far as to doubt the validity of the concept altogether. Coming from an agent-based modelling perspective, he considers that emergent phenomena are in principle explainable and deducible from descriptions of agents or individuals (the parts) and their interactions. He therefore sees agent-based computational models (of 'generative' science) as explanatory devices of macrolevel phenomena: 'if you didn't grow it, you didn't explain its emergence' (ibid: 43). So 'mysterious' macro phenomena that others consider as emergent, are seen in this case as a by-product of the limitations of current knowledge about the systems we study. In this sense macro phenomena may be linked to unpredictability but not to unexplainability.

7.1.1.2 Strong emergence

Strong notions of emergence require that the macroscopic level of a system has some distinct properties and characteristics despite being derived from the microscopic level. To do this, one must explain how the system as a whole possesses causal powers and particularly how it impinges on the bottom level components. So, another solution to the conundrum is to actually find an explanation for downward causation[12]. We will come back to the issue of strong emergence later in this chapter. It is important however to highlight again at this point the centrality of the problem of untangling the relationship between micro and macro properties and phenomena, or between bottom-up and top-down processes and causal influences.

7.1.2 Phenomenological emergence

It is also useful here to consider another customary understanding of emergence which associates it with the 'appearance' of a 'new', 'global' pattern or behaviour formed from the 'bottom-up' and out of 'local interactions'. Again the problem is whether 'new' is defined in objective or subjective terms and similarly whether the perceived global, high-level patterns and behaviours are intrinsic to the system in hand or just epiphenomena. The notion of 'appearance' is sometimes used by hard core reductionists to indicate that emergence is an effect (trick) of human perception or psychology while in reality phenomena are completely reducible to - and computable from - parts and their interactions. In this sense emergence is often linked to surprise (Ronald et al, 1999). Others use the notion of 'appearance' in way to make the reductionist stance a bit weaker and preserve some of the magical appeal of emergence. Yet, others use it to show that emergence depends by the existence of an observer at a fundamental level. It is not an epiphenomenon but it is contingent upon the interaction of a system with its environment (an observer).

7.1.3 Emergence as creation of new observational and descriptive categories

In fact there is a considerable literature which associates emergence with the creation, or the necessity to create, and use, new observational and descriptive categories. For example, Pattee (1988) clearly links emergence to the activities of observation and measurement. In particular, he distinguishes three levels of

[12] The notion of downward causation is attributed to Donald Campbell (1974) and is exactly used to mean the causal influence of the whole to its parts (how the parts are determined or constrained by the whole).

emergent behaviour. The first, syntactic emergence, is associated with symmetry breaking and chaotic dynamics as exemplified in the works of René Thom, Ilya Prigogine and Benoit Mandelbrot. The second level of emergent behaviour, semantic emergence, is associated with creativity (both at the genetic and the cognitive level) and consists in the assignment of new meanings to completed measurements and observations. Finally, the third level, measurement itself, is associated with the creation of new measurement devices.

We mentioned previously that emergence and complexity represent in many ways the two faces of the same coin. Complexity is also defined in relation to an observer, observational frame or reference frame. The conviction that (complex) systems are conditionally defined relative to an observer goes back to cybernetics and general systems theory. For example, both Ross Ashby (1962) and Heinz von Foerster (1984) examined in depth the ideas of organisation and self-organisation as characteristics of complex systems and emphasised the role of the observer. They both contended that self-organisation is a deceptive term as in reality a system can only increase or improve its organisation by being coupled with another system with which it interacts or exchanges energy (its environment). Maturana and Varela (1980) who studied living organisms (autopoietic systems) and their high level cognitive abilities, also maintained that living systems 'cannot be understood independently of the part of the ambience with which they interact: the niche; nor can the niche be defined independently of the living system that specifies it' (ibid: 9).

7.1.3.1 Emergence relative to a model

Robert Rosen's (1985, 1991) elaboration of the concepts of complexity and emergence within the framework of the modelling relation is very instructive in this context. The modelling relation as a scientific endeavour refers to 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 in order 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 natural world (Figure 38). The crux of the idea is that the natural world is constituted by a set of perceivable qualities, and linkages between qualities, which we call observables: 'As such, then, a natural system from the outset embodies a mental construct (i.e. a relation established by the mind between percepts) which comprises a hypothesis or model pertaining to the organization of the external world' (Rosen, 1985: 47). Rosen associated complexity with the concept of error (the discrepancy between a system and its model) and related the appearance of bifurcation (emergent phenomena) with our ability to produce enough independent encodings to fully describe a given natural system. For a more detailed treatment of modelling relation see also Cariani (1989).

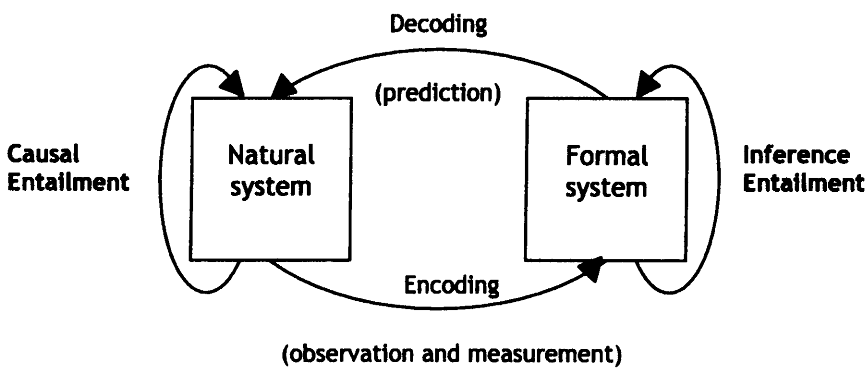


Figure 38 Rosen's (1985) diagram of modelling relation.

Casti (1986) also adopts a view of 'complexity as a *latent* or *implicate* property of a system, a property made explicit only through the interaction of a given system with another' (pp 146). He explicates the idea by pointing out the existence of two levels of complexity: design complexity, which is the complexity of the system in relation to the observer, and control complexity which is the complexity of the observer relative to the system. The complexity of a system depends crucially on the nature of the observables that describe it, the observables of the system that observes it and their modes of interaction. Thus, the complexity of a system S in relation to an observing system O corresponds to the number of non-equivalent descriptions (i.e. descriptions that are not reducible to each other) that O can generate for S.

Cariani (1989, 1991) capitalised on the ideas and epistemological reflections attested by cybernetics researchers, as well as Pattee and Rosen, in order to develop an observer-centred conception of emergence. He distinguishes between three different views of emergence: a) computational emergence, which adopts the view that emergent forms arise from local computational interactions, b) thermodynamic emergence, that sees emergence as the formation of new stable physical structures from noise and fluctuations, and c) emergence relative to a model, which views emergence as the realisation of new functions that then lead to the creation of new semantic and syntactic relations, new observables. Focussing on the question of design of adaptive devices that can exhibit emergent behaviour he argues that all evolutionary strategies used in simulations are incapable of creating new primitive structures (semantic or syntactic relations) because they are all restricted to the formal computational domain. The crucial point made is that 'Our simulated organisms are likewise part of informationally closed computational cycles, purposely insulated from the nonsymbolic world outside the simulations. This is exactly why we cannot make measurements on the physical world through programs and simulations that we specify completely' (Cariani, 1991: 791).

7.1.3.2 Emergence and hierarchical organisation

Another view which associates emergence with the creation of new descriptive and observational categories is provided by Baas (1994). Baas sees emergence in relation to hierarchical organisation, as the creation of higher-level structures through the mediation of observational mechanisms. Central to his argument is the study of emergence by considering three basic notions: structures (as the primitive entities), observational mechanisms for evaluating, observing and describing structures, and interactions among entities. He offers a definition of emergence which can briefly be described as follows: a property P is emergent at a certain level (S^2) - which is constructed from the set of primitive entities and interactions among them - if the property can be observed (and described) at this level but not at the level below it (S^1) using the same observational mechanisms. The definition captures the idea that although the higher-level structure is constructed by the interaction between entities at the lower level, new observational mechanisms are needed in order to describe the property P. Baas further distinguishes deducible or computable emergence, where the observational mechanism is an algorithm or deductive process, from observational (non deducible) emergence where the observational mechanism is a semantic meaning function or a truth function. It is also interesting to note that Baas considers observational mechanisms to play the role of some type of selection process that guides evolution towards higher order structures and he explicitly links this process to the subject of design.

It is interesting to note that the concept of levels, the view that natural (physical, biological, cognitive and social) systems are hierarchically[13] organised, seems to be inherent in any expression of emergence - even the elementary distinction between parts and wholes is an immediate assumption about the existence of levels. The understanding of complex systems as hierarchical structures has its roots in ideas originally exposed by Koestler (1967), Simon (1969), Allen and Starr (1982) and others, but many contemporary researchers explicitly adopt a multi-level view of complex systems (e.g. Baas, 1994; Emmeche et al, 1997, 2000; Damper, 2000; Corning, 2002; Lane, 2006). The concept of levels offers a natural basis for the study not only of how complex systems are structured and evolved, but also how observation and control can become part of the equation.

7.1.3.3 Emergence and relative complexity

A unification of the emergence-relative-to-a-model view and the view that binds emergence with the creation of high level structures is suggested by Bonabeau and Dessalles (1997) via a concept of relative complexity. The work builds on algorithmic definitions of complexity. In general, algorithmic complexity is defined in relation to the effort needed to describe an observed system. This effort is typically measured as a function of the computational resources needed

[13] The relation between hierarchy and complexity is somewhat baffling and a short explanatory note would be useful here. The word hierarchy is typically used to refer to a system in which every entity has a 'rank' and an ability to control the entities below it. Usually all power is concentrated to a single entity at the top. This in reality is only one type or meaning of hierarchy; the word is also used to refer to other kinds of organisation like networks, where power may be equally distributed to the different entities.

to complete a description - for example, in terms of time and space needed, or in terms of the size of the description. Bonabeau and Dessalles use a version of algorithmic complexity that associates complexity with the length of the shortest algorithm required to describe a system (which can be measured in relation to the number of axioms and rules needed to produce a description). The existence of a minimal description (realized as a program, algorithm, or formal system) is often perceived as an analogue of the most economical hypothesis that explains a phenomenon. In order to capture syntactic and semantic aspects of emergence, Bonabeau and Dessalles offer a definition of relative complexity which includes a clear distinction between observational and descriptive tools. *Relative complexity* C, denoted by C(S/D,T), is defined with respect to a set of observational tools (observables or detectors) D and a set of descriptive tools (relations between observables) T used to compute a description of the detected structures S. Complexity is then a relative concept defined as the difficulty of decomposing (or describing) a system S when certain detectors D are employed together with a theory T about the interdependence between the observables. In this context, emergence is linked with a decrease of relative complexity. This decrease reflects a shortening of the overall description caused by the activation of a higher level detector (a new observable at a higher level of abstraction, or a more general model) which substitutes lower level observations.

7.1.4 Conclusions: organisation, complexity and emergence

Overall, we can reason that the most important insight from these discussions about the existence and creation of hierarchical structures, the importance of inter and intra level interactions and the complementary relationship between systems and their environments, is that a system's organisation is indeed what explains and determines its complexity. Although organisation is itself an elusive term we can generally assert that it constitutes a universal attribute of a system (for a detailed treatment of this idea see Zamenopoulos 2007, and Zamenopoulos and Alexiou, 2007b). Crucially, organisation expresses not only how things are related but also how things work together. Organisational descriptions therefore focus on describing structural and functional properties jointly. This is an epistemological and methodological assertion. To quote Rosen's ideas again, if we focus on 'particles', emergence becomes an epiphenomenon - 'a particle, or any unit of structural analysis, does not (indeed, cannot) acquire new properties by being associated with a larger family of such units'. On the contrary, a functional unit 'changes as the system to which it belongs changes' (Rosen 1991: 121). In order to understand and explain emergence it is necessary to have an understanding of the organisation and causal processes occurring between system entities.

Making the step of attributing causal powers not (only) to the components themselves, but also to interdependencies and interactions among them and with the external world, is an important move towards the reconciliation of holism with reductionism. This view is adopted for example by Corning (2002) who refers to emergence as an effect associated with 'contexts in which constituent parts with different properties are modified, reshaped or transformed by their participation in the whole' (pp 24). Campbell and Bickhard (2001) in their take on emergence also maintain that constituents alone are not sufficient and that causal power must be attributed to their organisation and to the relations with external factors/elements in the environment that are necessary for self-maintenance.

The purpose of understanding and defining emergence is to be able to replicate it (exploit the positive effects, i.e. high-level functionalities such as cognition or intelligence) and/or harness it (avoid the negative effects, i.e. errors, breakdowns, unpredictability, etc). This effort however invokes a conundrum: If we cannot properly define it or understand it then it is useless as a concept; if we can define it, then it is not emergence... That is why many are content with weak definitions of emergence - they are interested in what <u>can</u> be achieved by computational methods and get some practical handle on it ('engineerable', or 'computable' emergence). Stephanie Forrest (1990) for example focuses on how emergent computation can be exploited for constructing efficient and flexible computational systems. Similarly, the themes of 'engineering emergence' and 'engineering complex systems' have recently drawn a lot of attention (see Braha et al, 2006; Brueckner et al, 2006; and Johnson, 2006).

There are several treatments which suggest that complexity or emergence are incomputable, some of which we saw above. Rosen (1991) for example, offered a proof built around an explication of the difference between mechanisms and living systems which is related to Gödel's incompleteness theorem. For relevant discussions and criticisms see (Ekdahl, 2000; Chu and Ho, 2006). Even if we agree with Rosen's and others' arguments and proofs that emergence is an incomputable phenomenon, the view adopted here is that we can nevertheless have a useful definition of it.

The only viable route for explaining, and possibly achieving, emergence is taking proper account of interactions and causal influences between system and environment. In the case of computational systems the environment consists of other computational systems and models, but also individuals and organisations that are engaged in their design and use. Consequently, our unit of study should include humans and computers together. The way knowledge, decisions and actions are distributed and organised accounts for the effects and abilities of the system on the whole. **What makes the whole be more than the sum of its parts is its organisation.** However, this is not an argument against working to develop computer models and systems that are capable of higher level functionalities - such as self-reference, anticipation, or learning - because in order for computers to be able to take part in complex interactions they should reflect an appropriate level of complexity themselves. This follows Casti's argument about complexity as a relative quality discussed above. Casti argues that the enterprise of understanding and managing complex systems involves equalisation of the control and design complexities; however in each different system and each different case there is an appropriate (in fact absolute) level at which control and design complexities are to be equalised, so that an efficient level of complexity is maintained overall (Casti, 1986: 166-168). This can be achieved through both feedback and feedforward (anticipatory) control strategies.

We have hopefully established here the importance of approaching emergence from an organisational perspective that appreciates the contingent and relative character of complexity and the importance of focussing on causal relationships and interactions. In Chapter 6 we also established coordination as a 'tool' for reconciling micro and macro aspects of emergence in our view of multi-agent design. In the following we will first revisit the coordination model of distributed learning control to see what sort of causal relationships are assumed and in which form the aspect of emergence as creation of new observables is incorporated. We will then discuss how a more elaborate view of multi-agent design as coordination can be developed drawing on Bar-Yam's (2004a, 2004b) work on multiscale variety in complex systems, which suggests links between the concepts of coordination and organisation.

7.2 A COMPLEXITY PERSPECTIVE ON MULTI-AGENT DESIGN AS COORDINATION

This section delves into the question of how coordination is linked to emergence and complexity with the aim of explaining coordination as a process by which collective design solutions emerge in multi-agent design settings. In the first part, the model of distributed learning control presented in Chapters 4 and 5 is revisited in order to explicate the suggested links between emergence and complexity. In the second part, a generalisation of the concept of coordination as an abstraction of multi-agent design is attempted. The objective is to take some first steps towards establishing an organisational definition of coordination and to expose directions for future research.

7.2.1 Revisiting the model of coordination as distributed learning control

In the previous section, the notion of emergence was associated with the creation of new observational and descriptive categories which engender changes in the relative complexity of a system. To better understand emergence in these terms, it is useful to consider the mechanisms or causal processes that produce such changes. One way to do this is to investigate the existence of different patterns of reasoning. Peirce (1934, 1958) who worked extensively on questions of scientific explanation and epistemology, has developed a theory about how different patterns of reasoning may be used to offer causal explanations of phenomena. His theory about deductive, inductive and abductive reasoning was picked up early on in design theory by Lionel March (1976) and later also developed by others (Coyne, 1988; Goel, 1988, Roozenburg, 1993; Roozenburg and Eekels, 1995). These ideas will be used for the analysis of the coordination model.

7.2.1.1 On the patterns of reasoning in design

Let us first have a brief look at the different types of reasoning and their definitions. *Deduction* refers to a type of reasoning that proceeds from a general law to a less general conclusion, which in logic is most commonly represented by modus ponens (if p then q is true; p is true; therefore q is true). It is a process that essentially derives theorems from a given rule or hypothesis. March (1976) remarks that deduction is an analytical process and can be associated with what in design is commonly considered as decomposition. By contrast, induction and abduction correspond to synthetic processes. *Induction* refers to a process of formulating generalised statements or laws from particular observations. It hypothesises the existence of a general class of phenomena based on a number of examples. Finally, *abduction* refers to a process of creative reasoning from effect to cause or, in other words, the production of explanations (the preconditions) for given facts. Roozenburg and Eekels (1995) further distinguish two different types of abduction, corresponding to Peirce's distinction between *explanatory* and *innovative* abduction (or '*innoduction*'). In the first case, reasoning proceeds from effect to cause *given* a particular premise, whereas in the second case both the cause and the theory or law used for inferring that cause are *unknown*. For a summary of these ideas see Table 6.

	Analytic	Synthetic		
	Deduction	Induction	Abduction (explanatory)	Innoduction (innovative abduction)
Premises	$p \rightarrow q$ p	p q	$p \rightarrow q$ q	q
Conclusion	q	$p \rightarrow q$	p	p $p \rightarrow q$
Pattern of reasoning	from general to particular	from particular to general	from particular to particular	from general to general

Table 6 A presentation of the different patterns of logical reasoning - adapted from Roozenburg and Eekels (1995).

In the coordination model introduced in Chapters 4 and 5 these different patterns of reasoning are assumed by different processes or components linked in a specific way. Untangling this association between the different types of reasoning will help us understand how multi-agent design is related to emergence and complexity, and how collective design solutions emerge out of the coordination process.

7.2.1.2 Analysing the patterns of reasoning in the distributed learning control model of coordination and their effects in complexity

The analysis offered in this section also uses the concept of relative complexity proposed by Bonabeau and Dessalles (1997). In particular, it is suggested that the notion of relative complexity can be specified here by considering that each agent has an observational and descriptive capacity which is tied to the existence of the different components or functions of the distributed learning control model. More specifically, each agent has a capacity to observe functional, behavioural and structural variables; and a capacity to describe and generate relationships between these variables. The observational capacity of each agent depends from the variety of the FBS variables; so the set of observables available to an agent (which Bonabeau and Dessalles call detectors) is increased or decreased depending on the number (the variety) of the FBS variables. The descriptive capacity of each agent is related to the capacity to model relationships between FBS variables and generate possible and desired configurations. So the descriptive capacity of agents depends from the comprehensiveness of the world and reference models and the extensiveness of the produced descriptions. The definition of relative complexity given by Bonabeau and Dessalles (1997) is translated as follows C=(W/FBSobservations, Descriptions(Causes, Rules)).

To develop this idea in more detail it is useful to proceed by focussing on a different part of the distributed learning control model at a time. For clarity the model is analysed using Figures 39 and 40 where the synthesis-analysis and formulation-reformulation components are depicted separately.

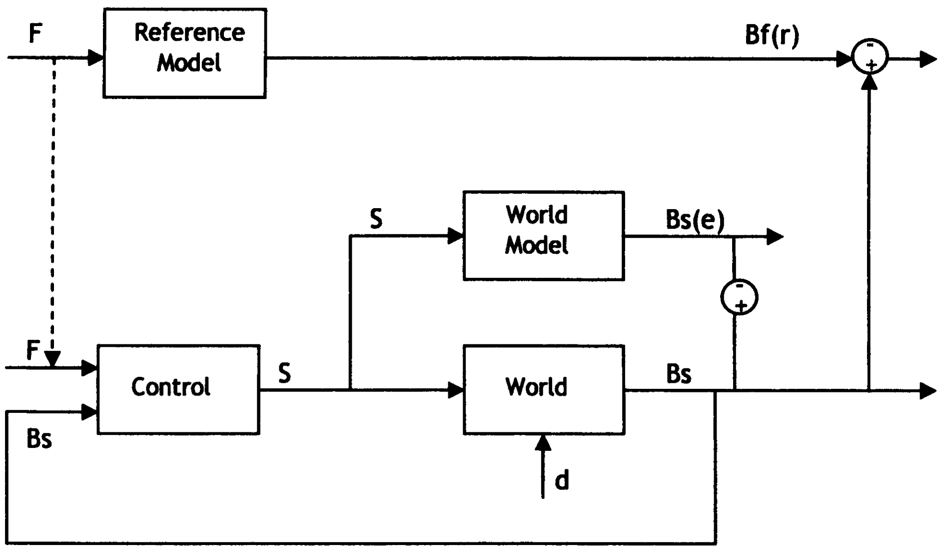


Figure 39 Analysis-Synthesis of S.

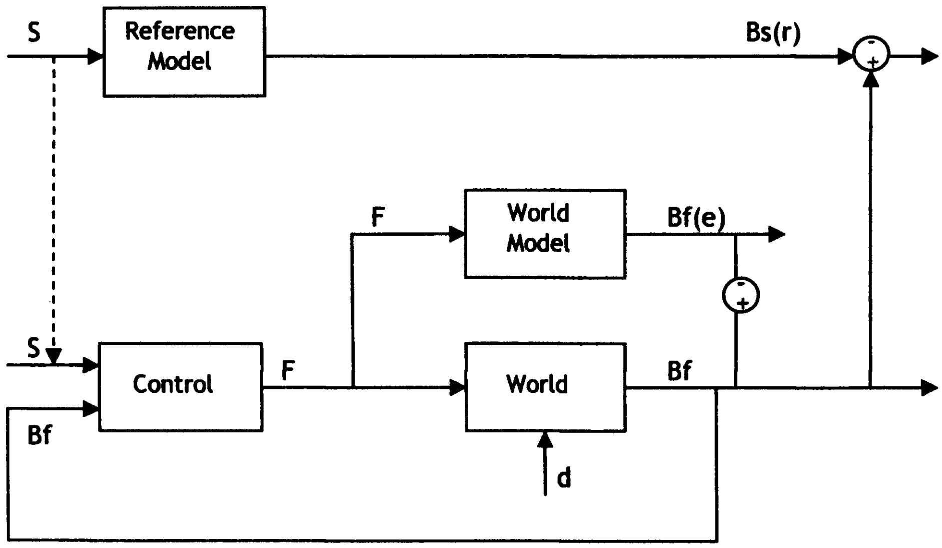


Figure 40 Analysis-Synthesis of F (Formulation-Reformulation).

The first part concerns the World, which represents the space of design changes produced by human and artificial agents. In the computational simulations described in this thesis, human and artificial agents taken together act as a distributed generative system that, based on a set of rules, laws, or premises, and given a particular input, infers and actualises particular changes. More analytically, given a set of observations (structures and functions) about the current state of the design objects, and a set of 'rules' (associating FBS variables), the overall system deduces structural and functional behaviours. In logic terms, this constitutes a form of deductive reasoning and can be represented as follows:

Deductive functions:

S	F	
S → Bs	F → Bf	
Bs	Bf	*production of design changes*

Note that the resulting configuration corresponds to the actual world, a real effect in the design space. In deduction, the FBS variables are given, the causes or initial conditions (F and S) are given, and the rules associating FBS are given. Most of the computational generative systems used in design studies, like shape grammars and other productive systems, correspond to such a deductive process (see Gips and Stiny, 1980). This process is indispensable as it is responsible for exploring and developing the design space. However, as such it does not contribute to any increase or decrease of complexity (it does not have an effect on the descriptive capacity of the system) and hence it cannot lead (on its own) to emergence. In a sense, the complexity of the space is already entailed in the rules that guide the process and the initial conditions. As an example, assume that the world W is generated by a set of shape rules (a shape grammar). Emergence is typically defined when a new pattern or shape is observed in W that cannot be deduced by the rules of the grammar (see for example Knight, 2003). The grammar therefore only determines a description; neither a change in the description, nor a new observational category (as the observation is done by an external to the system observer, the designer). This is consistent with Epstein's argument we briefly outlined above.

The second functional part of this system covers the learning components, which in general take the resulting configurations (observations of the actual world) and produce a generalised rule which associates initial conditions with expected results. This corresponds to an inductive process - a process of law extraction from examples - and can be represented as follows:

Inductive functions:

S	F	
Bs	Bf	
$S \rightarrow Bs$	$F \rightarrow Bf$	*creation of world and reference models*

Neural networks, particularly as used for pattern recognition and classification, are inductive systems par excellence. Connectionist models (Coyne, 1990; Coyne and Newton, 1990; Coyne, 1991; Newton and Coyne, 1991) have been used in design studies because of their ability to form hypotheses about the laws governing the world/design space, which can then be utilised for the generation of design alternatives. The appeal of course is that these hypotheses, patterns, or laws, emerge as a result of the networks' ability to reconfigure themselves based on previously acquired knowledge and in response to the appearance of new examples. In the system that we investigate here, the neural network learning components are used in order to generate and improve knowledge about the world. By generating new laws, induction affects the descriptive capacity of the system. If the gained knowledge enlarges the amount of information available to the system, then it effectively leads to an increase of the relative complexity; if it helps produce more succinct descriptions of the world then it leads to a decrease of relative complexity (and hence emergence).

The third part of the system concerns the controller functions. The controlling actions, which are derived through learning, are launched as possible causes of the desired consequence (a goal, an intended change in the configuration). The process of acquiring the control action is close to explanatory abduction. Abduction affects the descriptive capacity of the system as it defines the number of causes or premises that lead to the desired descriptions. Note that the control action itself aims to reduce complexity by restricting or conditioning the space of possible configurations.

Abductive functions:

$S \rightarrow Bs$	$F \rightarrow Bf$	
Bs	Bf	
S	F	*inference of the control action*

In reality, the picture becomes a bit more complicated if we look at the synergy between the two learning and control components. Combining knowledge about structural and functional change and about the relation between structures and functions (see inductive functions above), the system is also able to create new hypotheses. This function corresponds to innoduction and can be represented as follows:

Innoductive functions:

$$\frac{Bs}{S,\ S \rightarrow Bs} \qquad \frac{Bf}{F,\ F \rightarrow Bf} \qquad \textit{production of new hypotheses}$$

In other words, each agent generates new observables which are used to guide the development of new knowledge about FBS interdependencies and new descriptions. Innoduction therefore may increase or decrease the relative complexity of the system by adding or reducing existing FBS variables.

We have already discussed the difficulties of implementing this model of coordination computationally. The computational model does not exploit the dynamics of open interaction between human and artificial agents which is crucial in order to genuinely produce new descriptive and observational categories (recall the discussion about the computability of emergence). Yet, what this model does achieve is to make clear that coordination in design is not a process which can be easily realised by simple bottom-up modelling of agents with local rules of interaction. It is a process that tries to achieve a balance between increase and decrease of complexity and crucially involves learning both as construction of new categories and as re-construction, re-organisation of knowledge.

The definitions of emergence presented earlier in the chapter, for example those suggested by Cariani, Casti, and Bonabeau and Dessalles, are particularly useful as a basis for understanding the relational character of emergence and its dependence on the existence of an observer. Observation mechanisms are guaranteed in the coordination model through the incorporation of learning and control functions. The present analysis offered an insight into the causal relationships involved in multi-agent design by reference to different patterns of reasoning. What is missing from this model is a clear sense of how micro and macro levels are linked together. What are the effects of macro order in the micro level and vice versa? We will attempt to shed light on this question by building on a definition of multiscale variety presented by Bar-Yam (2004a).

7.2.2 Beyond distributed learning control: Towards an organisational definition of multi-agent design as coordination

The complexity of a system can be defined in many different ways (for a broad review of measurements of complexity see Edmonds, 1999). One commonly used indicator of complexity is variety: the number of independent actions or degrees of freedom available to a system. The Law of Requisite Variety proposed by Ashby (1956) as part of his theory of cybernetics, states that in order for a system to be able to control its environment it must have at least an equal number of control actions (variety) as the environment has of perturbations/disturbances.

Bar-Yam (2004a) developed a version of this law ('multiscale variety'), which additionally takes into account the role of the organisation of individual subsystems of the controlling system. Importantly, a core concept in this is coordination, which he considers to represent any scheme by which individual parts of a system may work together. The gist of the idea behind multiscale variety is that a system 'must be able to coordinate the right number of components to serve each task, while allowing the independence of other sets of components to perform their respective tasks' (ibid: 41).

7.2.2.1 Multiscale variety and organisation

In defining the requisite variety of a system, Bar-Yam distinguishes between the number of possible actions, or behaviours, of the system from the number of its components. Let us consider that a system has N components (or subsystems) and each component has a number of m available actions (or behaviours) - Figure 41.

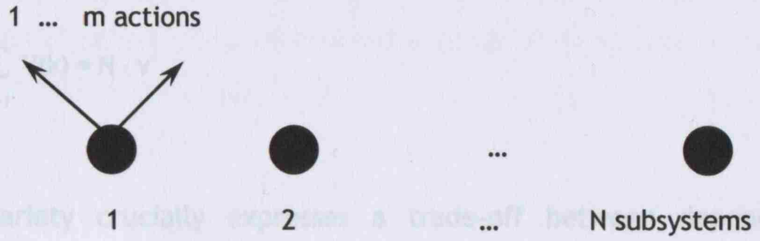


Figure 41 The variety of a system depends from the number of components N and the number of possible actions m of the system.

Then the total number of actions available to the system is m^N. The total variety of the system can be given by $V = \log(m^N) = N \cdot v$, where N is the number of subsystems and v is the variety of each individual component. Bar-Yam further introduces the concept of *scale* (k) to denote the number of dependent components within a subsystem. That is, a k number of different components may be organised together in a n(k) number of different groupings, each with a number of different actions available to them (Figure 42). So k express the number of dependent components in a grouping and n(k) the number of independent groupings.

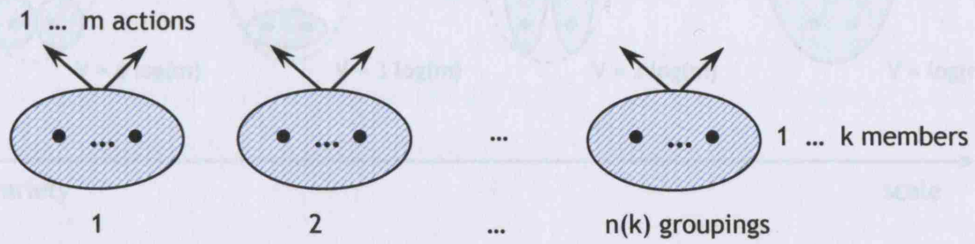


Figure 42 The variety of a system at a particular scale depends from the number of independent groupings n(k) between components and the number of possible actions.

In this case, the variety of a system at scale k can be given by $D(k) = n(k) \cdot v$. Hence, the total variety of the system is equal to the sum of variety at the different scales $V = \sum D(k)$. Bar-Yam shows an alternative definition of multiscale variety as follows:

$$\sum_{k=1}^{N} V(k) = N \cdot v \qquad \textit{Equation 5}$$

Multiscale variety crucially expresses a trade-off between dependence and independence, or scaling and variety. We can illustrate the concept using Figure (Figure 43). The different diagrams represent different scales or different groupings between components. The same number of components N (here N=6) can perform a task at scale N with variety equal to that of one component, or a task at scale 1 with variety N times as great. The finer the scale (from right to left) the greater the variety, and the less coordinated the components of the system are (more independent). On the contrary as the scale becomes coarser (from left to right), the coordination between component increases and the variety of the system is reduced.

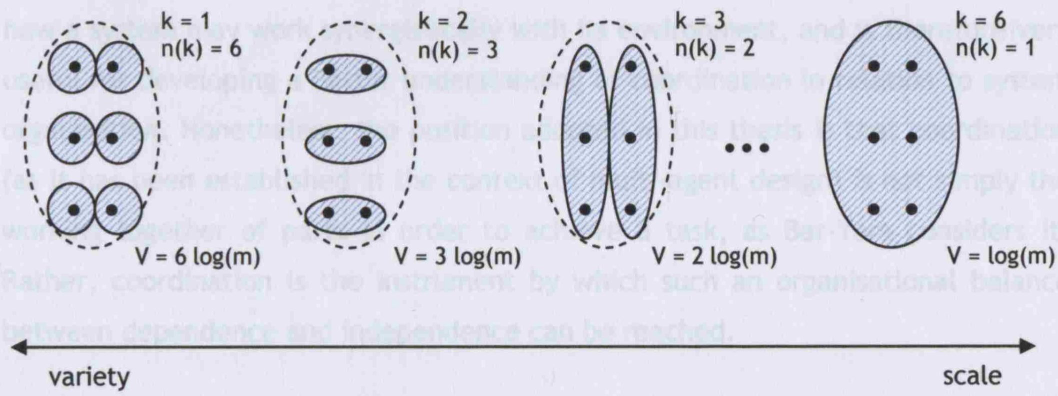


Figure 43 An illustration of the interplay between scale and variety in defining the system organisation. Each diagram represents a different scheme of organisation. k denotes scale, n(k) denotes number of groupings, V denotes variety and m is the number of possible actions.

In the example above, the components are organised in groupings of equal size (i.e. all components form groupings at the same scale). Complex systems will typically have components organised in groupings of different sizes. For example, in Figure 44, components are organised in three different groupings of different scales: a group of one, a group of two and a group of three (i.e. for k=1, n(1)=1, for k=2, n(2)=1 and for k=3, n(3)=1).

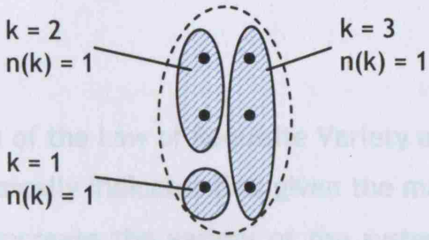


Figure 44 An example of a possible scheme of organisation where components are organised in groups of different sizes.

Multiscale variety therefore is a key characterisation of system organisation. The idea of a trade-off between scaling and variety, is at the heart of understanding how a system may work synergistically with its environment, and is therefore very useful for developing a better understanding of coordination in relation to system organisation. Nonetheless, the position adopted in this thesis is that coordination (as it has been established in the context of multi-agent design) is not simply the working together of parts in order to achieve a task, as Bar-Yam considers it. Rather, coordination is the instrument by which such an organisational balance between dependence and independence can be reached.

7.2.2.2 Multiscale variety and emergence

In a later paper Bar-Yam (2004b) also uses the concept of multiscale variety to develop an understanding of strong emergence. Strong emergence occurs when the properties of a system as a whole may determine (constrain) the parts. Moreover, Bar-Yam suggests that strong emergence also indicates a situation where a global constraint acts on a system as a whole and not (directly) on its components. In brief, the explanation for this is that because of (fluctuating) multiscale variety, the global constraint is not applied uniformly across the system. It follows that we cannot possibly grasp this global property from observations of the properties (and constraints) of the individual subsystems in isolation. We will not get into the details of the mathematical formulation here, but suffice it to say that it gives a useful account of emergence as a macroscopic property which, while it impacts on individual components, it cannot be deduced from them.

The multiscale version of the Law of Requisite Variety as specified by the equation above (Equation 5) basically indicates that given the maximum variety denoted by $N \cdot v$, any action to increase the variety of the system at one scale necessarily impacts on its variety at other scales. So, the number of components of a system imposes **a bound** on the possible schemes of organisation the system can achieve. The behaviour of a system and its capacity to exercise control over its environment is therefore a function of the organisation of the system in response to this environment.

Now for Bar-Yam, the global constraints of a system either arise from evolutionary process or are predetermined by factors external to the (self-)organising process (i.e. by a designer). So for example, within a city the number of buildings and the size of the land available for development, constrain the possible schemes of organisation (the possible configurations) that can be achieved. These global constraints may affect the size, the grouping, and the packing-together of buildings - irrespectively of the individual goals set by the architect or the client of the building. Similarly, in the model with the cuboids elaborated in Chapter 4, the number of cuboids and the size of the space within which their actions are developed, constrain the possible plan configurations. As we have previously discussed however, in multi-agent design, the number of components (the number of buildings, cuboids, agents, or FBS variables) is not given in advance, and defining them is in fact a key subject of the coordination process. Although design is a constrained process (e.g. because of limited resources), extending the design space with new variables or components is necessary for achieving emergence. In this sense, multi-agent design is not simply about finding optimal schemes of organisation that will satisfy certain given constraints, but it requires the elaboration of the very constrains that will bound the space of possible organisational schemes. Coordination as an abstraction of multi-agent design is thought as a process that explains the emergence of a particular organisation.

The same observation about how global constrains exist in Bar-Yam's definition applies to the expressions of relative complexity and emergence provided by Casti, and Bonabeau and Dessales, which we explored earlier in the chapter. Their definitions consider exclusively the interrelationship between system and environment and their relative complexity. The overall complexity of system and environment taken together is subject to external (evolutionary) processes. While evolutionary processes of random variation and mutation are important external, top-down constraints, in design we wish to incorporate also such constraints that are produced by the system itself. This issue points back to the micro-macro debate elaborated in the previous chapter.

7.2.2.3 A preliminary definition of multi-agent design as coordination

We can now attempt a preliminary definition of multi-agent design as coordination from a complex systems perspective. Coordination can be defined as a dynamic generative process towards a (non-permanent) scheme of organisation (a balance between scaling and variety) that explains the emergence of global constraints or purposes. The definition expresses coordination as a process that involves striking the right balance between creating structural dependencies and producing independent functionalities. More importantly, the definition entails the idea that the emergence of global constraints and goals is part of the multi-agent design process; the macro-level boundaries of the design space are causally related to the generative processes that take place at the micro-level of agents. Finally, the definition also implies that global constraints and the choice of organisation to satisfy those constraints co-evolve with each other and are subject to continuous change.

A more formal treatment of the concept is a subject for future elaboration. However, this definition is already useful for understanding coordination as approached in the thesis in relation to the broader context of complex systems science. Combined with the analysis of the distributed learning control model it provides a picture of the causal processes involved in multiagent design and a sense of how organisation impacts and enables those processes. On the other hand, the definition of coordination provided is unique for (multi-agent) design and thus enables us to distinguish design problems from other problems in complexity such as synchronisation, multi-objective optimisation or resource allocation (recall the multiple views of coordination in various fields offered in Chapter 2).

It is also useful to see the advantages of coordination compared to other abstractions used in complex systems science. Self-organisation for example, generally refers to a decrease of complexity (as the system becomes more organised). Coordination (as expressed here) is not linked to a particular effect on the complexity of a system and can be used to refer to the balance of relative complexity between a system and its environment. Furthermore, coordination as an abstraction readily incorporates a meaning of co-ordering among multiple levels. Compared to co-evolution, the concept of coordination is also preferable because it captures a notion of purposeful action, or intentional change, which is absent from co-evolution (the usual notion of co-evolution, not as has been transferred to design by Maher). One could rightly argue that what is meant by coordination is ultimately emergence - provided that we adhere to a specific view of emergence: relative to observer, multi-level, and incorporating bottom-up as well as top-down causation. That said, the problem with using the term emergence is that there is no generally agreed definition of it, and it carries a lot of philosophical baggage which might confuse rather than clarify the core argument. Coordination seems indeed to be the most appropriate abstraction for multi-agent design.

7.3 SUMMARY AND CONCLUSIONS

The aim of this chapter was to unravel a complexity perspective of coordination. For this purpose a review of emergence was offered first in order to distil some important characteristics and conditions to bring to bear in the context of multiagent design. This investigation established the importance of approaching emergence from an organisational perspective, which appreciates the contingent and relative character of complexity and the importance of focussing on causal relationships and interactions. The other insight gained from the review, was that emergence is linked to the creation of new descriptive and observational categories. In the second section of the chapter, Peirce's forms of reasoning were used as a way to analyse the distributed learning control model presented in Chapters 4 and 5 and understand what sort of causal processes take part in coordination, what are the effects on complexity, and how new observables may be created in the process. The main insights were that coordination involves achieving a balance between an increase and decrease of complexity, and that learning and control functions are crucial for incorporating the necessary observation mechanisms. Finally, in the last section of this chapter, the work of Bar-Yam on multiscale variety was used in order to develop a preliminary definition of coordination which may additionally account for the effects of multilevel organisation.

Chapter 8
SUMMARY AND CONCLUSIONS

This chapter reviews the main aims and objectives of the thesis, summarises how these objectives were met, and presents the main results and their implications. It also offers a discussion of possible avenues for future investigation.

8.1 HYPOTHESES AND OBJECTIVES

The overall objective of the thesis, set in place in the introductory chapter, was to develop a coherent framework for understanding multi-agent design decision making with an appreciation for its social and generative, creative, character. The driving hypothesis was that coordination is an appropriate concept for this purpose.

From this general objective a set of more specific objectives was derived:

- To introduce the concept of coordination by drawing from studies in different fields and explicate its relevance for multi-agent design
- To identify the key dimensions of multi-agent design as coordination
- To experiment with computational models and simulations in order to evaluate, and build upon, the established dimensions of coordination

8.2 MEETING THE OBJECTIVES

Let us see how these objectives were satisfied. First of all the concept of coordination was introduced in Chapter 2, in relation (and contrast) to other concepts such as collaboration, cooperation and conflict. This was achieved by a comprehensive review of coordination in design, as well as different fields like decision sciences, organisational science, and distributed artificial intelligence.

Based on this overview, the case was made that coordination is not only an important aspect of design, but indeed a useful abstraction for understanding distributed (or multi-agent) design. Coordination is a concept that gives attention to the social character of design while preserving the idea that the agents who take part in the process have individual knowledge, needs and goals. In this sense, coordination makes it possible to focus on the all-important aspect of distribution, interdependency and social interaction, without supposing the existence of a common goal or a benevolent disposition of agents towards cooperation. Indeed the notion of conflict was considered to be as important an idea as that of equilibrium.

The review also helped identify three key dimensions for understanding and defining multi-agent design through the concept of coordination: the dimensions of learning, decentralised control and co-evolution of problem-solution. Distributed design tasks involve knowledge that is spread among local agents and thus learning is an important instrument not only for enhancing the individual ability of agents to derive design solutions, but also for the creation of a shared knowledge about the overall design task and its constraints. Decentralised control is also derived as a characteristic of distribution: in multi-agent design, goals and requirements are also distributed and the design decisions are taken at a local level without any external central source of control. Finally, design decision making involves the mutual exploration and generation of problem and solution spaces, and achieving a balance between pursuing and reformulating individual goals. These are necessary attributes of coordination, but can also be seen as tools for developing coordinated, creative, design solutions.

In Chapter 4 the hypothesis was further explored that the key dimensions of multi-agent design as coordination can be formalised using the mechanics of distributed learning control. The idea behind distributed learning control is that agents basically perform a control process that generates actions (design decisions) in order to meet time-variant individual targets, despite endogenous uncertainties and exogenous disturbances coming from the actions of others. Learning in this setting corresponds to capturing interdependencies among decision variables in order to improve the controlling ability of agents, as well as to inform the process of goal formation. Coordination is thus conceptualised as a condition that is not explicitly modelled but rather emerges via the distributed actions of agents. These ideas were explored through building computational models and simulations in Chapter 5. One of the main insights gained from this undertaking, was an understanding of the importance of including in the model of coordination an explicit representation of, or linkage to, the external world. This common world provides agents with a source of conflicts, but it also provides them with opportunities for acquiring new knowledge and pursuing the generation of creative design solutions. Another insight gained was the understanding that a clear distinction between goals and expectations is conceptually necessary. Overall, through the process of building computational models and simulations it was possible to better frame the main assumptions about multi-agent design, understand their consequences, and examine the coherence of the general framework of coordination. The experimentation also helped to identify areas for further development and specifically revealed the need to investigate questions of sociality and emergence.

Following this line of thought, Chapter 6 elaborated on what it means to consider design as a social process, specifically by pointing to the issue of the micro-macro link. It was shown that the framework of coordination as distributed learning control can be used to capture the interplay between the micro level of individual goals, needs and decisions, and the macro level where these are expressed, thus giving rise to a complex net of interdependencies, constraints, as well as opportunities for creativity. Finally, Chapter 7 discussed and developed a complex systems view of coordination, which offers a deeper understanding of the relationship between emergence and coordination. Coordination was thus defined within a more general context and related to multi-level organisation.

8.3 CONTRIBUTION OF THE THESIS

It is trite to say that there are many different types of doctoral theses that provide many different types of contributions. For example, particularly in the domains of engineering, computer science or mathematics, a doctoral work may be focussed on developing a new application, deriving a new algorithm, extending a particular application, deriving theorems or proofs etc. The present thesis does not provide technical solutions, but it provides a theoretical framework for understanding multi-agent design based on scientific methodologies and insights. The purpose of this section is to lay out the particular contributions of the thesis.

One of the main contributions of this thesis is that it thoroughly investigates the issue of coordination from different perspectives and provides a comprehensive account of how the concept is used and can be used in design. As discussed in Chapter 3, an important issue that hinders communication between designers from different disciplines is the lack of a common language or common understanding of various terms. The thesis raises awareness about the diversity of perceptions and usages of the terms collaboration, cooperation and coordination within design research, as well as in other domains (such as artificial intelligence, decision sciences and complex systems science), and hence makes a useful step towards establishing a common language and understanding.

Besides this, the very proposition of using the concept of coordination as an abstraction of multi-agent design comes with a number of benefits. As previously explained, at the core of the notion of coordination is the idea that the decisions and actions of agents in a distributed design setting are interdependent. Coordination therefore places emphasis on distribution, interdependence and complexity and facilitates the understanding of multi-agent design as a social process that includes collaboration as well as conflict. This is not only a more realistic perspective of multi-agent design, but it is also a more generic one as it allows capturing design situations where agents do not share a common goal, and collaboration is not a prerequisite for the creation of design solutions. Although there is a growing body of studies interested in social aspects of design, these are usually concerned with descriptions of work practices, or of interactions between designers within a particular setting, and do not provide a general theory of multi-agent design which can be used to observe and interpret reality. The dimensions of multi-agent design identified in the thesis (i.e. the dimensions of learning, decentralised control and co-evolution) are proposed as fundamental features of coordination which can be used for observing, interpreting and even facilitating design activity.

The thesis further contributes to the development of a theory of multi-agent design by grappling with the question of what it is that constitutes design as a social activity. In particular, the thesis argues that for that purpose it is necessary to reconcile the view of design as intentional construction of artefacts (whether these are tangible or not), and the view of design as a product of social activity, that is formed and constrained by social structures. The treatment in Chapter 6, illustrates how the model of coordination as distributed learning control can be used to express the link between individual action and social structure (the link between micro and macro levels of causation). This is valuable for design research, but it is also valuable for social science in general, because it provides a way to understand the function of design in shaping social reality.

It is important to note here that the thesis does not take a ready-made notion of coordination, but it makes a case for a unique conception of coordination that can capture the distinctive character of multi-agent design. The thesis takes the view that there is a special link between design and emergence. Design is not only about the emergence of shapes, unintended behaviours or patterns, but it is also about defining the boundaries and constraints within which design problems and design solutions co-evolve. Using insights from complex systems science, the thesis argues that the relationship between agents, their goals, and the design variables they manipulate, is at the same time a product of the design process, but also a constraint over individual agents. Coordination is then defined as a dynamic process towards a scheme of organisation that entails the emergence of collective design solutions. In other words, coordination is an organising process that explains the emergence of designs. With this definition, the thesis contributes to a characterisation of multi-agent design as a distinct process or 'question', that can be compared and contrasted with other processes and questions, like for instance, problem solving, multi-objective optimisation, synchronisation, or resource allocation. In this way, the thesis makes a useful step towards establishing a common language and understanding of design across different domains. It also helps present design as a unique and noteworthy problem for complexity science. The definition of coordination offered in the last chapter is also a small step towards the development of a formal representation which can be further elaborated and used for analysis.

It was stated at the beginning of the thesis that the current study is part of a larger research programme that ultimately aims to inform the development of design decision support systems. The ideas discussed in the thesis, particularly in the last two chapters, point towards a need to shift our (methodological and epistemological) focus onto organisational characteristics and descriptions of phenomena. The investigation of the issues of sociality and emergence highlighted the need to pay great attention to the interactions and causal influences between system and environment: in the case of computational systems, the environment consists of other computational systems and models, but also individuals and organisations that are engaged in their design and use. Consequently, the tools we develop should aim to enable and exploit the complexity and organisational capacity of those distributed human-computer networks. Recalling the argument in Chapter 7, another important implication is that in order for computers to be able to take part in complex interactions, they should reflect an appropriate level of complexity. Hence, seeking to develop computer models and systems that are capable of higher level functionalities (such as learning, anticipation and reflexivity) is an important target for future design decision support systems.

Finally, another contribution of the thesis is that it tries out an unexplored avenue of investigation which brings together design research and complex systems science. The thesis follows the premise that the two fields share a common epistemological stance. At the core of this stance is the idea that knowledge can only be obtained through a constructive process, rather than through analysis or decomposition. In line with this assumption, the theory of multi-agent design is developed by building computational models and simulations. This is an original methodological approach as knowledge is obtained not by analysis of the produced data and their statistical properties but through designing, building and experimenting with the simulation. This can be seen as a new way of doing 'research by design' which is scientifically sound.

8.4 FUTURE WORK

There are of course many issues deriving from this investigation that are worth further exploration and development. For instance, the issue of how individual and social learning are linked together in multi-agent design is an open question, the elucidation of which will help better understand and support design as a social process. As discussed, an interesting computational experiment to investigate this question, would be to include a representation of the social network of agents (a representation of the way agents are connected) and develop a process such that this network influences the formation of agents' knowledge and goals and (possibly) vice versa. This experimentation would be useful in exploring various hypotheses about successful modes of social learning, successful strategies for goal adoption, or influential parameters for achieving better coordination.

In the computational models presented, the multiple agents interact with each other through a common world. As discussed, communication between agents was established in a non-direct way. Future experimentation could also include direct communication between agents. This would allow us to investigate more traditional issues such as the effects of negotiation or bargaining, but would also potentially give us the opportunity to explore more general questions about the use and evolution of (a shared) language between agents.

Additionally, exploration and experimentation with different forms and techniques of learning would also offer valuable insights. For example, it would be useful to understand when learning is more effective, and whether particular types of learning are more successful or more appropriate for different phases of design. Investigation of these issues may include theoretical, simulation based, as well as empirical research. Although the thesis focussed on the first two types of research, it is recognised that empirical research can help generate insightful hypotheses, and it is also extremely important for testing, substantiating and verifying the theoretical assumptions and results. In truth, all these types of investigation seem to be necessary for further developing the framework of multi-agent design coordination.

Finally, it is important to notice that seeking to develop computer models and systems capable of higher level functionalities in order to enhance the capabilities of future design decision support systems, seems to be a necessary but also extremely difficult pursuit. We have encountered in this thesis many difficulties with attempting to realise such high level functionalities computationally. A potentially useful way forward is to pursue the development of new formal languages able to express the complexity of design problems (see for example Zamenopoulos, 2007). These new formalisms may be proved to be realisable in computer systems, or may require adoption of alternative paradigms that go beyond digital computation. This is a very long-term aim, and can only be achieved through painstaking interdisciplinary work.

APPENDIX 1
MATLAB CODE

1 THE DISTRIBUTED LEARNING CONTROL MODEL - VERSION 2

% distributed learning control model v2. With 3 agents, each consisting of a nn plant model and a nn controller // each agent runs sequentially and there is no other 'external' process to produce the world // in this version all agents start with the same W and then get feedback from the next agent

% initialize inputs to start simulating the controllers [Wej; W]. An initial world W(14, 1) needs also to be provided

```
cIn1=[rand(14, 1); W];

cIn2=[rand(14, 1); W];

cIn3=[rand(14, 1); W];
```

% declarations of variables // numbers refer to agents, e=expected

```
World=W;

Control1=[];

Control2=[];

Control3=[];

World1e=[];

World2e=[];

World3e=[];
```

% add variables W1, W2, W3

```
W1=W;

W2=W;

W3=W;
```

% initialise weights of the neural networks // cannj controller, mannj world model network

```
cann1=init(cann1);
cann2=init(cann2);
cann3=init(cann3);
mann1=init(mann1);
mann2=init(mann2);
mann3=init(mann3);

for t=1:100
```

% simulate CONTROLLER

```
U1=sim(cann1,cIn1); %calculate new first controller signal
U2=sim(cann2,cIn2); %calculate new second controller signal
U3=sim(cann3,cIn3); %calculate new third controller signal

U1=abs(U1);
U2=abs(U2);
U3=abs(U3);
```

% visualize Worlds // the archworld program takes the 14-digit array and produces a coloured image of the archetype

```
archworldU1(U1);
archworldU2(U2);
archworldU3(U3);
```

% save worlds // in order to keep a record of all the worlds/control actions produced by the agents (for time t=100 the size of the final matrix is 14x301)

```
World=[World, U1, U2, U3];
```

```
% prepare inputs for Plant training // here W is different for each agent except from the first cycle
mIn1=[U1; W1];
mIn2=[U2; W2];
mIn3=[U3; W3];

% save control actions
Control1=[Control1, U1];
Control2=[Control2, U2];
Control3=[Control3, U3];

W1_old=W1;
W2_old=W2;
W3_old=W3;

% prepare training W for each agent // each agent takes as world W the control action of the next agent
W1=U2;
W2=U3;
W3=U1;

% train PLANT MODEL of the world // here W is different for each agent
mann1=train(mann1, mIn1, W1); %train first neural network
mann2=train(mann2, mIn2, W2); %train second neural network
mann3=train(mann3, mIn3, W3); %train third neural network

% simulate PLANT MODEL of the world
W1e=sim(mann1, mIn1); %simulate first neural network
W2e=sim(mann2, mIn2); %simulate second neural network
W3e=sim(mann3, mIn3); %simulate third neural network
```

```
W1e=abs(W1e);
W2e=abs(W2e);
W3e=abs(W3e);

% save expected worlds
World1e=[World1e, abs(W1e)];
World2e=[World2e, abs(W2e)];
World3e=[World3e, abs(W3e)];

% prepare inputs to train controller
cIn1T=[W1; W1_old];
cIn2T=[W2; W2_old];
cIn3T=[W3; W3_old];

% train CONTROLLER
cann1=train(cann1, cIn1T, U1); %train first controller neural network
cann2=train(cann2, cIn2T, U2); %train second controller neural network
cann3=train(cann3, cIn3T, U3); %train third controller neural network

% prepare inputs to simulate controller
cIn1=[W1e; W1];
cIn2=[W2e; W2];
cIn3=[W3e; W3];

end
```

2 NEURAL NETWORK DESIGN

Controller

Neural network object

```
>> disp(cann1)

   Neural Network object:

   architecture:

         numInputs: 2
         numLayers: 2
       biasConnect: [1; 1]
      inputConnect: [1 1; 0 0]
      layerConnect: [0 0; 1 0]
     outputConnect: [0 1]
     targetConnect: [0 1]

        numOutputs: 1  (read-only)
        numTargets: 1  (read-only)
    numInputDelays: 0  (read-only)
    numLayerDelays: 0  (read-only)

   subobject structures:

            inputs: {2x1 cell} of inputs
            layers: {2x1 cell} of layers
           outputs: {1x2 cell} containing 1 output
           targets: {1x2 cell} containing 1 target
            biases: {2x1 cell} containing 2 biases
      inputWeights: {2x2 cell} containing 2 input weights
      layerWeights: {2x2 cell} containing 1 layer weight
```

```
  functions:

      adaptFcn: 'trains'
       initFcn: 'initlay'
     performFcn: 'mse'
      trainFcn: 'trainlm'

  parameters:

    adaptParam: .passes
     initParam: (none)
   performParam: (none)
    trainParam: .epochs, .goal, .max_fail, .mem_reduc,
                .min_grad, .mu, .mu_dec, .mu_inc,
                .mu_max, .show, .time

  weight and bias values:

            IW: {2x2 cell} containing 2 input weight matrices
            LW: {2x2 cell} containing 1 layer weight matrix
             b: {2x1 cell} containing 2 bias vectors

  other:

      userdata: (user stuff)
```

Layer 1

```
>> cann1.layers{1}

ans =

    dimensions: 32
   distanceFcn: ''
     distances: []
       initFcn: 'initnw'
   netInputFcn: 'netsum'
```

```
    positions: [0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31]
         size: 32
  topologyFcn: 'hextop'
  transferFcn: 'tansig'
     userdata: [1x1 struct]
```

Layer 2

```
>> cann1.layers{2}

ans =
   dimensions: 14
  distanceFcn: ''
    distances: []
      initFcn: 'initnw'
  netInputFcn: 'netsum'
    positions: [0 1 2 3 4 5 6 7 8 9 10 11 12 13]
         size: 14
  topologyFcn: 'hextop'
  transferFcn: 'logsig'
     userdata: [1x1 struct]
```

Plant model

Neural network object

```
>> disp(mann1)

  Neural Network object:

  architecture:
     numInputs: 2
```

```
    numLayers: 2
  biasConnect: [1; 1]
 inputConnect: [1 1; 0 0]
 layerConnect: [0 0; 1 0]
outputConnect: [0 1]
targetConnect: [0 1]

   numOutputs: 1  (read-only)
   numTargets: 1  (read-only)
numInputDelays: 0  (read-only)
numLayerDelays: 0  (read-only)

subobject structures:
       inputs: {2x1 cell} of inputs
       layers: {2x1 cell} of layers
      outputs: {1x2 cell} containing 1 output
      targets: {1x2 cell} containing 1 target
       biases: {2x1 cell} containing 2 biases
 inputWeights: {2x2 cell} containing 2 input weights
 layerWeights: {2x2 cell} containing 1 layer weight

functions:
     adaptFcn: 'trains'
      initFcn: 'initlay'
   performFcn: 'mse'
     trainFcn: 'trainlm'

parameters:
   adaptParam: .passes
    initParam: (none)
 performParam: (none)
```

```
    trainParam: .epochs, .goal, .max_fail, .mem_reduc,
                .min_grad, .mu, .mu_dec, .mu_inc,
                .mu_max, .show, .time

  weight and bias values:

          IW: {2x2 cell} containing 2 input weight matrices
          LW: {2x2 cell} containing 1 layer weight matrix
           b: {2x1 cell} containing 2 bias vectors

  other:

      userdata: (user stuff)
```

Layer 1

```
>> mann1.layers{1}

ans =

    dimensions: 32
   distanceFcn: ''
     distances: []
       initFcn: 'initnw'
   netInputFcn: 'netsum'
     positions: [0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31]
          size: 32
   topologyFcn: 'hextop'
   transferFcn: 'tansig'
      userdata: [1x1 struct]
```

Layer 2

```
>> mann1.layers{2}

ans =
```

dimensions: 14

distanceFcn: ''

distances: []

initFcn: 'initnw'

netInputFcn: 'netsum'

positions: [0 1 2 3 4 5 6 7 8 9 10 11 12 13]

size: 14

topologyFcn: 'hextop'

transferFcn: 'logsig'

userdata: [1x1 struct]

3 NEURAL NETWORK OBJECT REFERENCE FROM MATLAB

Documentation available online at : http://www.mathworks.com/access/helpdesk/help/toolbox/nnet/

- **Network Properties**

These properties define the basic features of a network. Subobject Properties describes properties that define network details.

Architecture

These properties determine the number of network subobjects (which include inputs, layers, outputs, targets, biases, and weights), and how they are connected.

net.numInputs

This property defines the number of inputs a network receives. It can be set to 0 or a positive integer.

Clarification. The number of network inputs and the size of a network input are *not* the same thing. The number of inputs defines how many sets of vectors the network receives as input. The size of each input (i.e., the number of elements in each input vector) is determined by the input size (net.inputs{i}.size).

Most networks have only one input, whose size is determined by the problem.

Side Effects. Any change to this property results in a change in the size of the matrix defining connections to layers from inputs, (net.inputConnect) and the size of the cell array of input subobjects (net.inputs).

net.numLayers

This property defines the number of layers a network has. It can be set to 0 or a positive integer.

Side Effects. Any change to this property changes the size of each of these Boolean matrices that define connections to and from layers,

net.biasConnect
net.inputConnect
net.layerConnect
net.outputConnect
net.targetConnect

and changes the size of each cell array of subobject structures whose size depends on the number of layers,

net.biases
net.inputWeights
net.layerWeights
net.outputs
net.targets

and also changes the size of each of the network's adjustable parameter's properties.

net.IW
net.LW
net.b

net.biasConnect

This property defines which layers have biases.It can be set to any *N*-by-1 matrix of Boolean values, where N_l is the number of network layers (net.numLayers). The presence (or absence) of a bias to the *i*th layer is indicated by a 1 (or 0) at

net.biasConnect(i)

Side Effects. Any change to this property alters the presence or absence of structures in the cell array of biases (net.biases) and, in the presence or absence of vectors in the cell array, of bias vectors (net.b).

net.inputConnect

This property defines which layers have weights coming from inputs.

It can be set to any N_l x N_i matrix of Boolean values, where N_l is the number of network layers (net.numLayers), and N_i is the number of network inputs (net.numInputs). The presence (or absence) of a weight going to the *i*th layer from the *j*th input is indicated by a 1 (or 0) at net.inputConnect(i,j).

Side Effects. Any change to this property alters the presence or absence of structures in the cell array of input weight subobjects (net.inputWeights) and the presence or absence of matrices in the cell array of input weight matrices (net.IW).

net.layerConnect

This property defines which layers have weights coming from other layers. It can be set to any N_l x N_l matrix of Boolean values, where N_l is the number of network layers (net.numLayers). The presence (or absence) of a weight going to the *i*th layer from the *j*th layer is indicated by a 1 (or 0) at net.layerConnect(i,j).

Side Effects. Any change to this property alters the presence or absence of structures in the cell array of layer weight subobjects (net.layerWeights) and the presence or absence of matrices in the cell array of layer weight matrices (net.LW).

net.outputConnect

This property defines which layers generate network outputs. It can be set to any 1 x N_l matrix of Boolean values, where N_l is the number of network layers (net.numLayers). The presence (or absence) of a network output from the *i*th layer is indicated by a 1 (or 0) at net.outputConnect(i).

Side Effects. Any change to this property alters the number of network outputs (net.numOutputs) and the presence or absence of structures in the cell array of output subobjects (net.outputs).

net.targetConnect

This property defines which layers have associated targets. It can be set to any 1 x N_l matrix of Boolean values, where N_l is the number of network layers (net.numLayers). The presence (or absence) of a target associated with the *i*th layer is indicated by a 1 (or 0) at net.targetConnect(i).

Side Effects. Any change to this property alters the number of network targets (net.numTargets) and the presence or absence of structures in the cell array of target subobjects (net.targets).

net.numOutputs (read-only)

This property indicates how many outputs the network has. It is always equal to the number of 1s in net.outputConnect.

net.numTargets (read-only)

This property indicates how many targets the network has. It is always set to the number of 1's in net.targetConnect.

net.numInputDelays (read-only)

This property indicates the number of time steps of past inputs that must be supplied to simulate the network. It is always set to the maximum delay value associated with any of the network's input weights.

```
numInputDelays = 0;
for i=1:net.numLayers
  for j=1:net.numInputs
    if net.inputConnect(i,j)
      numInputDelays = max( ...
        [numInputDelays net.inputWeights{i,j}.delays]);
    end
  end
end
```

net.numLayerDelays (read-only)

This property indicates the number of time steps of past layer outputs that must be supplied to simulate the network. It is always set to the maximum delay value associated with any of the network's layer weights.

```
numLayerDelays = 0;
for i=1:net.numLayers
  for j=1:net.numLayers
    if net.layerConnect(i,j)
      numLayerDelays = max( ...
       [numLayerDelays net.layerWeights{i,j}.delays]);
    end
  end
end
```

- **Subobject Structures**

These properties consist of cell arrays of structures that define each of the network's inputs, layers, outputs, targets, biases, and weights.

The properties for each kind of subobject are described in <u>Subobject Properties</u>.

net.inputs

This property holds structures of properties for each of the network's inputs. It is always an N_i x 1 cell array of input structures, where N_i is the number of network inputs (net.numInputs).

The structure defining the properties of the *i*th network input is located at net.inputs{i}.

Input Properties. See Inputs for descriptions of input properties.

net.layers

This property holds structures of properties for each of the network's layers. It is always an N_l x 1 cell array of layer structures, where N_l is the number of network layers (net.numLayers).

The structure defining the properties of the *i*th layer is located at net.layers{i}.

Layer Properties. See Layers for descriptions of layer properties.

net.outputs

This property holds structures of properties for each of the network's outputs. It is always a 1 x N_l cell array, where N_l is the number of network outputs (net.numOutputs).

The structure defining the properties of the output from the *i*th layer (or a null matrix []) is located at net.outputs{i} if net.outputConnect(i) is 1 (or 0).

Output Properties. See Outputs for descriptions of output properties.

net.targets

This property holds structures of properties for each of the network's targets. It is always a 1 x N_l cell array, where N_l is the number of network targets (net.numTargets).

The structure defining the properties of the target associated with the *i*th layer (or a null matrix []) is located at net.targets{i} if net.targetConnect(i) is 1 (or 0).

Target Properties. See Targets for descriptions of target properties.

net.biases

This property holds structures of properties for each of the network's biases. It is always an N_l x 1 cell array, where N_l is the number of network layers (net.numLayers).

The structure defining the properties of the bias associated with the *i*th layer (or a null matrix []) is located at net.biases{i} if net.biasConnect(i) is 1 (or 0).

Bias Properties. See Biases for descriptions of bias properties.

net.inputWeights

This property holds structures of properties for each of the network's input weights. It is always an N_l x N_i cell array, where N_l is the number of network layers (net.numLayers), and N_i is the number of network inputs (net.numInputs).

The structure defining the properties of the weight going to the *i*th layer from the *j*th input (or a null matrix []) is located at net.inputWeights{i,j} if net.inputConnect(i,j) is 1 (or 0).

Input Weight Properties. See Input Weights for descriptions of input weight properties.

net.layerWeights

This property holds structures of properties for each of the network's layer weights. It is always an N_l x N_l cell array, where N_l is the number of network layers (net.numLayers).

The structure defining the properties of the weight going to the *i*th layer from the *j*th layer (or a null matrix []) is located at net.layerWeights{i,j} if net.layerConnect(i,j) is 1 (or 0).

Layer Weight Properties. See Layer Weights for descriptions of layer weight properties.

Functions

These properties define the algorithms to use when a network is to adapt, is to be initialized, is to have its performance measured, or is to be trained.

net.adaptFcn

This property defines the function to be used when the network adapts. It can be set to the name of any network adapt function. The network adapt function is used to perform adaption whenever adapt is called.

[net,Y,E,Pf,Af] = adapt(NET,P,T,Pi,Ai)

Side Effects. Whenever this property is altered, the network's adaption parameters (net.adaptParam) are set to contain the parameters and default values of the new function.

net.gradientFc

This property defines the function used to calculate the relationship between the network's weights and biases and performance either as a gradient or Jacobian. The gradient function is used by many training functions.

Side Effects. Whenever this property is altered, the network's gradient parameters (net.gradientParam) are set to contain the parameters and default values of the new function.

net.initFcn

This property defines the function used to initialize the network's weight matrices and bias vectors. You can set it to the name of the network initialization function. The initialization function is used to initialize the network whenever init is called.

net = init(net)

Side Effects. Whenever this property is altered, the network's initialization parameters (net.initParam) are set to contain the parameters and default values of the new function.

net.performFcn

This property defines the function used to measure the network's performance. You can set it to the name of any of the performance functions. The performance function is used to calculate network performance during training whenever train is called.

[net,tr] = train(NET,P,T,Pi,Ai)

Side Effects. Whenever this property is altered, the network's performance parameters (net.performParam) are set to contain the parameters and default values of the new function.

net.trainFcn

This property defines the function used to train the network.You can set it to the name of any of the training function. The training function is used to train the network whenever train is called.

[net,tr] = train(NET,P,T,Pi,Ai)

Side Effects. Whenever this property is altered, the network's training parameters (net.trainParam) are set to contain the parameters and default values of the new function.

Parameters

net.adaptParam

This property defines the parameters and values of the current adapt function. Call help on the current adapt function to get a description of what each field means.

help(net.adaptFcn)

net.gradientParam

This property defines the parameters and values of the current gradient function. Call help on the current initialization function to get a description of what each field means.

help(net.gradientFcn)

net.initParam

This property defines the parameters and values of the current initialization function. Call help on the current initialization function to get a description of what each field means.

help(net.initFcn)

net.performParam

This property defines the parameters and values of the current performance function. Call help on the current performance function to get a description of what each field means.

help(net.performFcn)

net.trainParam

This property defines the parameters and values of the current training function. Call help on the current training function to get a description of what each field means.

help(net.trainFcn)

Weight and Bias Values

These properties define the network's adjustable parameters: its weight matrices and bias vectors.

net.IW

This property defines the weight matrices of weights going to layers from network inputs. It is always an N_l x N_i cell array, where N_l is the number of network layers (net.numLayers), and N_i is the number of network inputs (net.numInputs).

The weight matrix for the weight going to the *i*th layer from the *j*th input (or a null matrix []) is located at net.IW{i,j} if net.inputConnect(i,j) is 1 (or 0).

The weight matrix has as many rows as the size of the layer it goes to (net.layers{i}.size). It has as many columns as the product of the input size with the number of delays associated with the weight.

net.inputs{j}.size * length(net.inputWeights{i,j}.delays)

These dimensions can also be obtained from the input weight properties.

net.inputWeights{i,j}.size

net.LW

This property defines the weight matrices of weights going to layers from other layers. It is always an N_l x N_l cell array, where N_l is the number of network layers (net.numLayers).

The weight matrix for the weight going to the *i*th layer from the *j*th layer (or a null matrix []) is located at net.LW{i,j} if net.layerConnect(i,j) is 1 (or 0).

The weight matrix has as many rows as the size of the layer it goes to (net.layers{i}.size). It has as many columns as the product of the size of the layer it comes from with the number of delays associated with the weight.

net.layers{j}.size * length(net.layerWeights{i,j}.delays)

These dimensions can also be obtained from the layer weight properties.

net.layerWeights{i,j}.size

net.b

This property defines the bias vectors for each layer with a bias. It is always an N_l x 1 cell array, where N_l is the number of network layers (net.numLayers).

The bias vector for the *i*th layer (or a null matrix []) is located at net.b{i} if net.biasConnect(i) is 1 (or 0).

The number of elements in the bias vector is always equal to the size of the layer it is associated with (net.layers{i}.size).

This dimension can also be obtained from the bias properties.

net.biases{i}.size

Other

The only other property is a user data property.

Userdata

This property provides a place for users to add custom information to a network object. Only one field is predefined. It contains a *secret* message to all Neural Network Toolbox users.

net.userdata.note

APPENDIX 2

PUBLICATIONS DERIVING FROM THIS RESEARCH

Alexiou, K. and T. Zamenopoulos (2001). A connectionist paradigm in the coordination and control of multiple self-interested agents. Proceedings of the 2nd National Conference Input 2001: Information Technology and Spatial Planning, Democracy and Technologies, Tremiti Islands, Italy, Politecnico di Bari, Dipartimento di Architettura e Urbanistica.

Alexiou, K. and T. Zamenopoulos (2002a). Artificial design and planning support: interactive plan generation and coordination in distributed decision making. In Proceedings of the 6th International Conference in Design and Decision Support Systems in Urban Planning, Eindhoven University of Technology.

Alexiou, K. and T. Zamenopoulos. (2002b). 'Designing plans: A control based coordination model, CASA Working Paper 48.' from http://www.casa.ucl.ac.uk/working_papers/paper48.pdf.

Alexiou, K. and T. Zamenopoulos (2008). 'Design as a social process: a complex systems perspective.' Futures **forthcoming.**

Zamenopoulos, T. and K. Alexiou (2002). Learning to be creative and the creative memory: A discussion motivated by a control based coordination model. Learning and Creativity Workshop notes, AID '02 Conference, Cambridge, UK.

Zamenopoulos, T. and K. Alexiou (2003a). Computer-aided creativity and learning in distributed cooperative human-machine networks. In CAAD Futures 2003, Dordrecht, The Netherlands, Kluwer Academic Publishers.

Zamenopoulos, T. and K. Alexiou (2003b). Structuring the plan design process as a coordination problem: the paradigm of distributed learning control coordination. Advanced Spatial Analysis: The CASA book of GIS. P. Longley and M. Batty. Redlands, California, ESRI Press: 407-426.

Zamenopoulos, T. and K. Alexiou. (2004). 'Design and anticipation: towards an organisational view of design systems, CASA Working Paper 76.' from http://www.casa.ucl.ac.uk/working_papers/paper76.pdf

Zamenopoulos, T. and K. Alexiou (2005). Linking design and complexity: a review. Proceedings of the ECCS 2005 Satellite Workshop Embracing Complexity in Design. J. Johnson, T. Zamenopoulos and K. Alexiou. Paris 17 November 2005, The Open University: 91-102.

Zamenopoulos, T. and K. Alexiou (2007a). 'Towards an anticipatory view of design.' Design Studies **28**(4): 411-436.

Zamenopoulos, T. and K. Alexiou (2007b). 'Rethinking the cybernetic basis of design: the concepts of control and organization.' Kybernetes **forthcoming.**

Bibliography

Achten, H. (2002). An approach for agent-systems in design support. Intelligent Agents in Design Workshop notes, AID '02 Conference, Cambridge, UK.

Ackoff, R. L. (1974). Redesigning the Future: A systems approach to societal problems. New York, John Wiley.

Ackoff, R. L. and F. Emery (1972). On Purposeful Systems. Chicago, Aldine-Atherton.

Agre, P. E. and D. Chapman (1987). Pengi: An implementation of a theory of activity. In Proceedings of the Sixth National Conference on Artificial Intelligence (AAAI-87), Seattle, Washington, AAAI.

Agre, P. E. and S. J. Rosenchein, Eds. (1996). Computational Theories of Interaction and Agency. Cambridge, Massachusetts, MIT Press.

Akin, Ö. (1986). Psychology of Architectural Design. London, Pion Limited.

Alexander, C. (1964). Notes on the synthesis of the form. Cambridge, MA, Harvard University Press.

Alexander, E. R. (1992). Approaches to Planning: Introducing current planning theories, concepts and issues. Philadelphia, Gordon and Breach Science Publishers.

Alexander, J. C. and B. Giesen (1987). From Reduction to Linkage: The Long View of the Micro-Macro Link. The Micro-Macro Link. J. C. Alexander, B. Giesen, R. Münch and N. J. Smelser. Berkeley, California, University of California Press: 1-42.

Alexander, J. C., B. Giesen, et al., Eds. (1987). The Micro-Macro Link. Berkeley, California, University of California Press.

Alexiou, K. and T. Zamenopoulos (2001). A connectionist paradigm in the coordination and control of multiple self-interested agents. Proceedings of the 2nd National Conference Input 2001: Information Technology and Spatial Planning, Democracy and Technologies, Tremiti Islands, Italy, Politecnico di Bari, Dipartimento di Architettura e Urbanistica.

Alexiou, K. and T. Zamenopoulos (2002a). Artificial design and planning support: interactive plan generation and coordination in distributed decision making. In Proceedings of the 6th International Conference in Design and Decision Support Systems in Urban Planning, Eindhoven University of Technology.

Alexiou, K. and T. Zamenopoulos. (2002b). 'Designing plans: A control based coordination model, CASA Working Paper 48.' from http://www.casa.ucl.ac.uk/working_papers/paper48.pdf.

Alexiou, K. and T. Zamenopoulos (2008). 'Design as a social process: a complex systems perspective.' Futures **forthcoming**.

Allen, T. F. H. and T. B. Starr (1982). Hierarchy. Perspectives for Ecological Complexity. Chicago and London, The University of Chicago Press.

Archer, B. L. (1965). Systematic Method for Designers. London, The Design Council.

Archer, B. L. (1970). An overview of the structure of the design process. Emerging Methods in Environmental Design and Planning. G. T. Moore. Cambridge, Massachusetts, The MIT Press: 285-307.

Arentze, T. and H. Timmermans (2002). A multi-agent model of negotiation processes between multiple actors in urban developments: A framework and results of numerical experiments, Eindhoven University of Technology [Unpublished paper].

Argyris, C. and D. Schön (1978). Organizational Learning: A theory of action perspective. Reading, Massachusetts, Addison Wesley.

Arias, E., H. Eden, et al. (2000). 'Transcending the individual human mind - creating shared understanding through collaborative design.' ACM Transactions on Computer-Human Interaction 7(1): 84-113.

Arrow, K. J. (1951). Social Choice and Individual Values. New York, John Wiley & Sons.

Ashby, R. W. (1956). An Introduction to Cybernetics. London, Chapman & Hall.

Ashby, R. W. (1962). Principles of the self-organizing system. Principles of Self-Organization. H. von Foerster and G. W. Zopf Jr. London, UK, Pergamon Press: 255-278.

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

Austin, S., A. Newton, et al. (2002). 'Modelling and managing project complexity.' International Journal of Project Management 20(3): 191-198.

Axelrod, R. (1984). The Evolution of Cooperation. New York, Basic Books.

Axelrod, R. (1997). 'Advancing the art of simulation in the social sciences.' Complexity 3(2): 16-22.

Axelrod, R. (1997). The Complexity of Cooperation: Agent-based models of competition and collaboration. Princeton, New Jersey, Princeton University Press.

Axelrod, R. (1997). 'The dissemination of culture: A model with local convergence and global polarization.' Journal of Conflict Resolution 41(2): 203-226.

Baas, N. A. (1994). Emergence, Hierarchies, and Hyperstructures. Artificial Life III. C. G. Langton. Reading, Massachusetts, Addison Wesley. XVII: 515-537.

Bak, P. (1996). How Nature Works: the science of self-organized criticality. New York, Copernicus.

Ball, L. J. and T. C. Ormerod (2000). 'Applying ethnography in the analysis and support of expertise in engineering design.' Design Studies 21(4): 403-421.

Bandura, A. (1977). Social Learning Theory. Englewood Cliffs, NJ, Prentice-Hall.

Bandura, A. (1989). 'Human agency in social cognitive theory.' American Psychologist 44(9): 1175-1184.

Bar-Yam, Y. (2004a). 'Multiscale variety in complex systems.' Complexity 9(4): 37-45.

Bar-Yam, Y. (2004b). 'A mathematical theory of strong emergence using multiscale variety.' Complexity 9(6): 15-24.

Batty, M. (1979). On Planning Processes. Resources and Planning. B. Goodall and A. Kirby. Oxford, Pergamon Press: 17-45.

Batty, M. (1980). 'Limits to prediction in science and design science.' Design Studies 1(3): 153-159.

Batty, M., J. Barros, et al. (2004). 'Cities: Continuity, Transformation, and Emergence, CASA Working Paper 72.' from http://www.casa.ucl.ac.uk/working_papers/paper72.pdf.

Batty, M., J. Desyllas, et al. (2002). 'The discrete dynamics of small-scale spatial events: Agent-based models of mobility in carnivals and street parades, CASA Working Paper 56.' from http://www.casa.ucl.ac.uk/working_papers/Paper56.pdf.

Batty, M., B. Jiang, et al. (1998). 'Local movement: Agent-based models of pedestrian flows, CASA Working Paper 4.' from http://www.casa.ucl.ac.uk/local_movement.doc.

Batty, S. E. (1977). 'Game-theoretic approaches to urban planning and design.' Environment and Planning B 4(2): 211-239.

Batty, S. E. (1996). Urban Land Use Planning in Western Europe: Rules, Games and Groups. Department of Political Science. Buffalo, The State University of New York. **PhD Thesis.**

Bazjanac, V. (1974). Architectural design theory: Models of the design process. Basic Questions of Design Theory. W. R. Spillers. Amsterdam, North-Holland: 3-20.

Beckers, R., O. E. Holland, et al. (1994). From local actions to global tasks: Stigmergy and collective robotics. In Artificial Life IV, Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems, Cambridge, Massachusetts, MIT Press.

Bedau, M. A. (1997). Weak emergence. Philosophical Perspectives: Mind, Causation, and the World Vol 11. J. Tomberlin. Oxford, Blackwell Publishers.

Beer, S. (1974). Designing Freedom. London, John Wiley & Sons.

Benenson, I. (1998). 'Multi-agent simulations of residential dynamics in the city.' Computers, Environment and Urban Systems 22(1): 25-42.

Bennett, C. H. (1985). Dissipation, Information, computational complexity and the definition of organization Emerging Syntheses in Science. D. Pines. Redwood, CA, Addison-Wesley: 215-233.

Berker, I. and D. C. Brown (1996). 'Conflicts and negotiation in single function agent based design systems.' Concurrent Engineering: Research and Applications, Special Issue on Multi-Agent Systems in Concurrent Engineering, Brown, D.C., Lander, S., Petrie, C. (eds) 4(1): 17-33.

Bertalanffy, L. v. (1968). General System Theory. Foundations, Development, Applications. New York, George Braziller.

Besussi, E. and A. Cecchini, Eds. (1996). Artificial Worlds and Urban Studies. Venice, DAEST, IUAV.

Bishop, I. D. and H. R. Gimblett (2000). 'Management of recreational areas: GIS, autonomous agents, and virtual reality.' Environment and Planning B: Planning and Design 27(3): 423-435.

Boden, M. A., Ed. (1996). Artificial Intelligence. San Diego, California, Academic Press.

Bonabeau, E. and J.-L. Dessalles (1997). 'Detection and emergence.' Intellectica **25**(2): 85-94.

Bonabeau, E., J.-L. Dessalles, et al. (1995a). 'Characterizing emergent phenomena (1): a critical review.' Revue Internationale de Systémique **9**(3): 327-346.

Bonabeau, E., G. Theraulaz, et al. (1994). The building behaviour of lattice swarms. In Artificial Life IV, Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems, Cambridge, Massachusetts, MIT Press.

Bond, A. H. and L. Gasser, Eds. (1988). Readings in Distributed Artificial Intelligence. San Mateo, CA, Morgan Kaufmann.

Bourdieu, P. (1977). Outline of a Theory of Practice. London, Cambridge University Press.

Bourdieu, P. (1984). Distinction: a social critique of the judgement of taste. London, Routledge & Kegan Paul.

Bourdieu, P. and L. J. D. Wacquant (1992). An Invitation to Reflexive Sociology. Chicago, University of Chicago Press.

Boyd, R. and P. J. Richerson (1988). Culture and the Evolutionary Process. Chicago, University of Chicago Press.

Braha, D., A. Minai, et al., Eds. (2006). Complex Engineered Systems: Science Meets Technology. Berlin, Springer.

Branki, N. E. (1994). An AI-based framework for collaborative design. Papers from the AAAI Workshop: AI and Collaborative Design, AAAI Press.

Branki, N. E., E. A. Edmonds, et al. (1993). 'A study of socially shared cognition in design.' Environment and Planning B: Planning and Design **20**(3): 295-306.

Brazier, F. M. T., C. M. Jonker, et al. (1996). Modelling project coordination in a multi-agent framework. In Proceedings of the Fifth Workshop on Enabling Technologies: Infrastructure and Collaborative Enterprises (WET ICE '96), Los Alamitos, IEEE.

Brazier, F. M. T., C. M. Jonker, et al. (2000). 'Compositional design and reuse of a generic agent model.' Applied Artificial Intelligence **14**(5): 491-538.

Brazier, F. M. T., L. V. Moshkina, et al. (2001). 'Knowledge level model of an individual designer as an agent in collaborative distributed design.' Artificial Intelligence in Engineering **15**(2): 137-152.

Brill, E. D., J. M. Flash, et al. (1990). 'MGA: A Decision Support System for complex, incompletely defined problems.' IEEE Transactions on Systems, Man and Cybernetics **20**(4): 745-757.

Brooks, R. A. (1986). 'A robust layered control system for a mobile robot.' IEEE Journal of Robotics and Automation **2**(1): 14-23.

Brown, D. C., B. Dunskus, et al. (1994). Using single function agents to investigate negotiation. AAAI Workshop: Models of Conflict Management in Cooperative Problem Solving, AAAI Press.

Brueckner, S. A., G. Serugendo, et al., Eds. (2006). Engineering Self-Organising Systems. Proceedings of the Third International Workshop ESOA 2005, Lecture Notes in Artificial Intelligence 3910. Berlin, Heidelberg, Springer.

Bucciarelli, L. L. (1988). 'An ethnographic perspective on engineering design.' Design Studies 9(3): 159-168.

Buller, D. J., Ed. (1999). Function, Selection and Design. New York, State University of New York.

Byrne, D. (2005). 'Complexity, configurations and cases.' Theory, Culture & Society 22(5): 95-111.

Calinescu, A., J. Efstathiou, et al. (1998). 'Applying and assessing two methods for measuring complexity in manufacturing.' The Journal of the Operational Research Society 49(7): 723-733.

Camerer, C. and T. H. Ho (1999). 'Experience weighted attraction learning in normal-form games.' Econometrica 67: 827-874.

Camerer, C. and M. Knez (1997). Coordination in organizations: A game theoretic perspective. Organizational Decision Making. Z. Shapira. New York, Cambridge University Press: 158-188.

Campbell, D. (1974). Downward causation in hierarchically organized biological systems. Studies in the Philosophy of Biology. F. Ayala and T. Dobzhansky. Berkeley, University of California Press: 179-186.

Campbell, M. I., J. Cagan, et al. (1998). A-Design: Theory and implementation of an adaptive, agent-based method for conceptual design. Artificial Intelligence in Design '98 (AID '98). J. S. Gero and F. Sudweeks. Dordrecht, Kluwer Academic Publishers: 579-598.

Campbell, M. I., J. Cagan, et al. (1999). 'A-Design: An agent-based approach to conceptual design in a dynamic environment.' Research in Engineering Design 11(3): 172-192.

Campbell, R. J. and M. H. Bickhard. (2001). 'Physicalism, emergence and downward causation.' from http://eprints.anu.edu.au/archive/00000029 Accessed on Jan 2006.

Cariani, P. (1989). On the Design of Devices with Emergent Semantic Functions. Department of Systems Science. New York, State University of New York at Binghamton. PhD Thesis.

Cariani, P. (1991). Emergence and Artificial Life. Artificial Life II. C. Langton, C. Taylor, J. D. Farmer and S. Rasmussen. Redwood City, California, Addison Wesley. X: 775-797.

Carley, K. and A. Newell (1994). 'The nature of the social agent.' Journal of Mathematical Sociology 19(4): 221-262.

Carley, K. M. and M. J. Prietula, Eds. (1994). Computational Organization Theory. Hillsdale, NJ, Lawrence Erlbaum Associates.

Castelfranchi, C. (1998). 'Modelling social action for AI agents.' Artificial Intelligence 103(1-2): 157-182.

Castelfranchi, C. (2001). 'The theory of social functions: challenges for computational social science and multi-agent learning

' Journal of Cognitive Systems Research 2(1): 5-38.

Castelfranchi, C. and R. Conte (1996). Distributed artificial intelligence and social science: Critical issues. Foundations of Distributed Artificial Intelligence. G. M. P. O' Hare and N. R. Jennings, John Wiley & Sons: 527-542.

Castelfranchi, C. and R. Conte (1998). 'Limits of economic and strategic rationality for agents and MA systems.' Robotics and Autonomous Systems 24(3-4): 127-139.

Casti, J. L. (1986). On System Complexity: Identification, Measurement, and Management. Complexity, Language, and Life: Mathematical Approaches. J. L. Casti and A. Karlqvist. Berlin, Springer-Verlag. 16: 146-173.

Chadwick, G. (1978). A Systems View of Planning. Towards a Theory of the Urban and Regional Planning Process. Oxford, Pergamon Press.

Chaitin, G. J. (1974). 'Information-theoretic computational complexity.' IEEE Transactions on Information Theory 20(1): 10-15.

Chakrabarti, A. and A. T. M. Blessing (1996). 'Representing functionality in design.' Artificial Intelligence for Engineering Design Analysis and Manufacturing (AIEDAM) 10(4): 251-253.

Chakrabarti, A. and A. T. M. Blessing, (eds) (1996). 'Special Issue in Representing Functionality in Design.' Artificial Intelligence for Engineering Design Analysis and Manufacturing (AIEDAM) 10(4).

Chiu, M.-L. (2002). 'An organizational view of design communication in design collaboration.' Design Studies 23(2): 187-210.

Chu, D. and W. K. Ho (2006). 'A category theoretical argument against the possibility of artificial life: Robert Rosen's central proof revisited.' Artificial Life 12(1): 117-134.

Churchman, C. W. (1967). 'Wicked problems.' Management Science 14(4): 141-142.

Clancey, W. J. (1997). Situated Cognition. Cambridge, Cambridge University Press.

Coates, G., R. I. Whitfield, et al. (2000). 'Coordination approaches and systems - Part II: An operational perspective.' Research in Engineering Design 12(2): 73-89.

Coates, P. and R. Thum (2000). Parallel systems and architectural form. In Proceedings of Greenwich 2000: Digital Creativity Symposium, Greenwich, University of Greenwich.

Cohen, M. D., J. G. March, et al. (1972). 'A garbage can model of organizational choice.' Administrative Science Quarterly 17(1): 1-25.

Cohen, M. D. and L. S. Sproull, Eds. (1996). Organizational Learning. Thousand Oaks, Sage.

Conte, R. (2003). 'Book review: Simulating organizations: computational models of institutions and groups.' Journal of Artificial Societies and Social Simulation 6(2).

Conte, R. and C. Castelfranchi (1995). Cognitive and Social Action. London, UCL Press.

Conte, R. and C. Dellarocas, Eds. (2001). Social Order in Multiagent Systems. Dordrecht, The Netherlands, Kluwer Academic Publishers.

Conte, R. and M. Paolucci (2001). 'Intelligent social learning.' Journal of Artificial

Societies and Social Simulation 4(1).

Cooper, R. (1999). Coordination Games. Cambridge, Cambridge University Press.

Cooper, R. W., D. V. DeJong, et al. (1989). 'Communication in the Battle of the Sexes game: Some experimental results.' Rand Journal of Economics 20(4): 568-587.

Cooper, R. W., D. V. DeJong, et al. (1990). 'Selection criteria in coordination games: Some experimental results.' The American Economic Review 80(1): 218-233.

Corning, P. A. (2002). 'The re-emergence of 'emergence': a venerable concept in search of a theory.' Complexity 7(6): 18-30.

Coyne, R. D. (1988). Logic Models of Design. London, Pitman.

Coyne, R. D. (1990). 'Design reasoning without explanations.' AI magazine 11(4): 72-80.

Coyne, R. D. (1990). 'Modelling the emergence of design descriptions across schemata.' Environment and Planning B: Planning and Design 18(4): 427-458.

Coyne, R. D. and S. Newton (1989). 'Design reasoning by association.' Environment and Planning B: Planning and Design 17(1): 39-56.

Coyne, R. D., F. Sudweeks, et al. (1996). 'Who needs the Internet? Computer-mediated communication in design firms.' Environment and Planning B: Planning and Design 23(6): 749-770.

Craig, D. L. and C. Zimring (2000). 'Supporting collaborative design groups as design communities.' Design Studies 21(2): 187-204.

Crawford, V. P. (1995). 'Adaptive dynamics in coordination games.' Econometrica 63(1): 103-144.

Cross, N., Ed. (1984). Developments in Design Methodology. Chichester, John Wiley & Sons.

Cross, N. (1985). 'Styles of learning, designing and computing.' Design Studies 6(3): 157-162.

Cross, N. (1993). A history of design methodology. Design Methodology and Relationships with Science. M. J. de Vries, N. Cross and D. P. Grant. Dordrecht, The Netherlands, Kluwer Academic Publishers: 15-27.

Cross, N. and A. C. Cross (1995). 'Observations of teamwork and social processes in design.' Design Studies 16(2): 143-170.

Cumming, M. (2002a). A bottom up approach to design coordination. In Proceedings of the 6th International Conference in Design and Decision Support Systems in Architecture, Eindhoven University of Technology.

Cumming, M. (2002b). Flexible and distributed coordination models for collaborative design. In Proceedings of the 20th eCAADe Conference - Connecting the Real and the Virtual [design e-ducation], Warsaw, Poland, Faculty of Architecture, Warsaw University of Technology.

Damper, R. I. (2000). 'Emergence and levels of description.' International Journal of Systems Science 31(7): 811-818.

Darley, V. (1994). Emergent Phenomena and Complexity. Artificial Life IV. R. A. Brooks and P. Maes. Cambridge, Mass, The MIT Press: 411-416.

Dautenhahn, K. (1998). 'The art of designing socially intelligent agents - science, fiction and the human in the loop.' Applied Artificial Intelligence 12(7-8): 573-617.

Dautenhahn, K., Ed. (2000a). Socially Intelligent Agents: the Human in the Loop, Technical Report FS-00-04. Menlo Park, California, AAAI Press.

Dautenhahn, K., Ed. (2000b). Human Cognition and Social Agent Technology. Amsterdam/Philadelphia, John Benjamins.

Dautenhahn, K., A. H. Bond, et al., Eds. (2002). Socially Intelligent Agents: Creating relationships with computers and robots. Dordrecht, The Netherlands, Kluwer Academic Publishers.

Dave, B. (1998). Teamwork constructs in architectural design. In Proceedings of the OZCHI'98 Conference, Los Alamitos, CA, IEEE Press.

Davidson, D. (1970). Mental Events. Experience and Theory. L. Foster and J. W. Swanson. Amherst, University of Massachusetts Press: 79-101.

de Bono, E. (1985). Six Thinking Hats. Boston, Little, Brown and co.

Demirbas, O. O. and H. Demirkan (2003). 'Focus on architectural design process through learning styles.' Design Studies 24(5): 437-456.

Deneubourg, J. L., S. Goss, et al. (1991). The dynamics of collective sorting: Robot-like ant and ant-like robot. In Simulation of Adaptive Behavior: From Animals to Animats, MIT Press/Bradford Books.

Dickey, J. W. and A. P. Vernon (1998). PlanMaster: A new application for interactive planning support. In Proceedings of the 5th International Conference on Computers in Urban Planning and Urban Management (CUPUM'98), Bombay, India, Narosa Publishing House.

Dijkstra, J., J. Jessurun, et al. (2002). A multi-agent cellular automata model of pedestrian movement. Pedestrian and Evacuation Dynamics. M. Schreckenber and S. D. Sharma. Berlin, Springer-Verlag: 173-180.

Dijkstra, J. and H. Timmermans (2002). 'Towards a multi-agent model for visualizing simulated user behavior to support the assessment of design performance.' Automation in Construction 11(2): 135-145.

Domeshek, E. A., J. L. Kolodner, et al. (1994). Using theories to overcome social obstacles in design collaborations. Papers from the AAAI Workshop: AI and Collaborative Design, AAAI Press.

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

Drogoul, A. and J. Ferber (1992). From tom thumb to the dockers: Some experiments with foraging robots. In Proceedings of the Second International Conference on Simulation of Adaptive Behavior: From Animals to Animats 2, Cambridge, Massachusetts, MIT Press/Bradford Books.

Duffy, A. H. B., D. C. Brown, et al. (1998). 'Special Issue: Machine Learning in Design.' Artificial Intelligence for Engineering Design Analysis and Manufacturing (AIEDAM) **12**(2).

Duffy, A. H. B., D. C. Brown, et al. (1996). 'Special Issue: Machine Learning in Design.' Artificial Intelligence for Engineering Design Analysis and Manufacturing (AIEDAM) **10**(2).

Dunskus, B. V., D. L. Grecu, et al. (1995). 'Using single function agents to investigate conflicts.' Artificial Intelligence in Engineering Design, Analysis and Manufacturing (AIEDAM), Special issue: Conflict Management in Design, I.Smith (ed) **9**(4): 299-312.

Durfee, E. H., V. R. Lesser, et al. (1989). 'Trends in Cooperative Distributed Problem Solving.' IEEE Transactions on Systems, Man and Cybernetics **1**(1): 63-83.

Durfee, E. H., V. R. Lesser, et al. (1990). Cooperation through communication in a distributed problem-solving network. Cognition, Computing and Cooperation. S. P. Robertson, W. Zachary and J. B. Black. Norwood, NJ, Ablex: 159-186.

Earl, C., J. Johnson, et al. (2005). Complexity. Design Process Improvement. A review of current practice. J. Clarkson and C. Eckert. London, Springer-Verlag: 174-197.

Eckert, C., J. Clarkson, et al. (2004). 'Change and customization in complex engineering domains.' Research in Engineering Design **15**(1): 1-21.

Eckert, C. and M. Stacey (2000). 'Sources of inspiration: a language of design.' Design Studies **21**(5): 523-538.

Edmonds, B. (1999). Syntactic Measures of Complexity. Manchester, University of Manchester. **PhD Thesis.**

Edmonds, B. (2005). Simulation and Complexity: how they can relate? Virtual Worlds of Precision - computer-based simulations in the sciences and social sciences. V. Feldmann and K. Mühlfeld. Munster, Lit Verlag: 5-32.

Edmonds, E. A., L. Candy, et al. (1994). 'Support for collaborative design: agents and emergence.' Communications of the ACM **37**(7): 41-47.

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. (2002). Can computers be social? Proceedings of the Fifth International Conference on Computing Anticipatory Systems (CASYS 2001). D. M. Dubois, American Institute of Physics: 95-106.

Emmeche, C., S. Køppe, et al. (1997). 'Explaining emergence: towards an ontology of levels.' Journal for General Philosophy of Science **28**(1): 83-119.

Emmeche, C., S. Køppe, et al. (1997). Levels, emergence and three version of downward causation. Donwward Causation. Minds, Bodies and Matter. P. B. Andersen, C. Emmeche, N. O. Finnemann and P. V. Christiansen. Aarhus, Aarhus University Press: 13-34.

Epstein, J. M. (1999). 'Agent-based computational models and generative social science.' Complexity **4**(5): 41-60.

Epstein, J. M. and R. Axtell (1996). Growing Artificial Societies: Social science from the bottom up. Washington, DC, Brookings Institution Press, Cambridge, Massachusetts, MIT Press.

Erev, I. and A. E. Roth (1997). 'Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria.' The American Economic Review **88**(4): 848-881.

Farrell, J. (1987). 'Cheap talk, coordination and entry.' Rand Journal of Economics **18**(1): 34-39.

Felder, R. M. and L. K. Silverman (1988). 'Learning and teaching styles in engineering education.' Engineering Education **78**(8): 674-681.

Ferber, J. (1994). Simulating with reactive agents. Many-Agent Simulation and Artificial Life. E. Hillebrand and J. Stender. Amsterdam, IOS Press: 8-30.

Ferber, J. (1999). Multi-Agent Systems: An introduction to Distributed Artificial Intelligence. London, Addison-Wesley.

Ferrand, N. (1996). Modelling and supporting multi-actor spatial planning using multi-agent systems. In Proceedings of the Third NCGIA Conference on Integrating GIS and Environmental Modelling, Santa Fe, New Mexico.

Festinger, L. (1957). A Theory of Cognitive Dissonance. Stanford, Stanford University Press.

Findlay, R. A. and S. L. Haugen (2000). From individual inquiry and attention to cohorts to a 'collaborative critique': The use of student groups to support individual designers. In Proceedings of the 5th International Conference in Design and Decision Support Systems in Architecture, Eindhoven University of Technology.

Fischer, G. (1999). Social creativity, symmetry of ignorance and Meta-Design. In Proceedings of the Third Conference on Creativity and Cognition, Loughborough, UK, ACM.

Fischer, G. (2000). 'Symmetry of ignorance, social creativity and meta-design.' Knowledge Based Systems **13**(4): 527-537.

Fischer, G. and K. Nakakoji (1997). 'Computational environments supporting creativity in the context of lifelong learning and design.' Knowledge Based Systems **10**(1): 21-28.

Flinn, M. V. (1996). 'Culture and the evolution of social learning.' Evolution and Human Behavior **18**(1): 23-67.

Forester, J. (1998). 'Creating public value in planning and urban design: The three abiding problems of negotiation, participation and deliberation.' Urban Design International **3**(1): 5-12.

Forrest, S. (1990). 'Emergent computation: self-organizing, collective, and cooperative phenomena in natural and artificial computing networks. Introduction to the Proceedings of the Ninth Annual CNLS Conference.' Physica D **42**: 1-11.

Forrester, J. (1969). Urban Dynamics. Waltham, Massachusetts, Pegasus Communications.

French, M. E. (1992). Form, Structure and Mechanism. London, MacMillan.

Frozza, R. and L. O. Alvares (2002). Criteria for the analysis of coordination in multi-agent applications. Coordination 2002, Lecture Notes in Computer Science 2315. F. Arbab and C. Talcott. Berlin Heidelberg, Springer-Verlag: 158-165.

Fudenberg, D. and D. M. Kreps (1993). 'Learning mixed equilibria.' Games and Economic

Behaviour 5(3): 320-367.

Fudenberg, D. and D. K. Levine (1998). The Theory of Learning in Games. Cambridge, Massachusetts, MIT Press.

Genesereth, M. R., M. Ginsber, et al. (1986). Cooperation without communication. In Proceedings of the 5th National Conference on Artificial Intelligence (AAAI-86), Philadelphia, PA, AAAI.

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, S. Rude and G. Grein. Aachen, Shaker Verlag: 47-56.

Gero, J. S. (1999). 'Recent design science research: Constructive memory in design thinking.' Architectural Science Review 42(1): 3-5.

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. and H. Fujii (2000). 'A computational framework for concept formation for a situated design agent.' Knowledge-Based Systems 13(6): 361-368.

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

Gero, J. S. and U. Kannengiesser (2003). A function-behaviour-structure view of social situated agents. CAADRIA 03. A. Choutgrajank, E. Charoenslip, K. Keatruangkamala and W. Nakapan. Bangkok, Rangsit University: 707-715.

Gero, J. S. and U. Kannengiesser (2004). A function-behaviour-structure ontology of processes. Design Computing and Cognition (DCC'06). J. S. Gero. Berlin, Heidelberg, Springer: 407-422.

Gero, J. S. and U. Kannengiesser (2004). 'The situated function-behaviour-structure framework.' Design Studies 25(4): 373-391.

Gero, J. S. and V. Kazakov (2003). On measuring the visual complexity of 3D solid objects. e-Activities in Design and Design Education, Europia Proceedings. B. Tuncer, S. Ozsariyildiz and S. Sariyilidiz. Paris, Europia: 147-156.

Gero, J. S. and M. L. Maher (1993). Artificial Intelligence in Collaborative Design, Technical Report WS-93-07. Menlo Park, California, AAAI Press.

Gero, J. S. and G. Nath (1997). Formalising situated learning in agent-based systems. In CAAD Futures '97, Dordrecht: Kluwer Academic Publishers.

Gero, J. S. and R. Sosa (2002). Creative design situations: Artificial creativity in communities of design agents. CAADRIA 2002, New York, Prentice Hall.

Geyer, F. and J. van der Zouwen, Eds. (1986). Sociocybernetic Paradoxes. Observation, control and evolution of self-steering systems. London, Sage.

Geyer, F. and J. van der Zouwen, Eds. (2001). Sociocybernetics. Complexity, autopoiesis and observation of social systems. Westport, Connecticut, Greenwood Press.

Giddens, A. (1984). The Constitution of Society. Outline of the Theory of Structuration. Cambridge, Polity Press.

Giesen, B. (1987). Beyond Reductionism: Four Models Relating Micro and Macro Levels. The Micro-Macro Link. J. C. Alexander, B. Giesen, R. Münch and N. J. Smelser. Berkeley, California, The California University Press: 337-387.

Gilbert, N. (1995). Emergence in Social Simulation. Artificial Societies: The computer simulation of social life. N. Gilbert and R. Conte. London, UCL Press: 144-156.

Gilbert, N. (2000). Modeling Sociality: The View from Europe. Dynamics in Human and Primate Societies. Agent-based Modeling of Social and Spatial Processes. T. A. Kohler and G. J. Gumerman. New York, Oxford, Oxford University Press: 355-371.

Gilbert, N. (2002). Varieties of emergence. Proceedings of the Agent 2002 Conference on Social Agents: Ecology, Exchange, and Evolution. Argonne, Argonne National Laboratory and University of Chicago: 41-50.

Gilbert, N. and R. Conte, Eds. (1995). Artificial Societies: The computer simulation of social life. London, UCL Press.

Gilbert, N. and R. Conte (1995). Introduction. Computer simulation for social theory. Artificial Societies: The computer simulation of social life. N. Gilbert and R. Conte. London, UCL Press: 1-15.

Gilbert, N., M. den Besten, et al. (2006). 'Emerging artificial societies through learning.' Journal of Artificial Societies and Social Simulation 9(2).

Gilbert, N. and K. G. Troitzsch (1999). Simulation for the Social Scientist. Buckingham, Open University Press.

Gips, J. and G. Stiny (1980). 'Production systems and grammars: a uniform characterization.' Environment and Planning B 7(4): 399-408.

Goel, V. (1988). 'Complicating the 'logic of design'.' Design Studies 9(4): 229-234.

Grecu, D. L. and D. C. Brown (1996). Learning by single function agents during spring design. Artificial Intelligence in Design '96 (AID '96). J. S. Gero and F. Sudweeks. Boston, Kluwer Academic Publishers: 409-428.

Gross, M. D., E. Y.-L. Do, et al. (1998). 'Collaboration and coordination in architectural design: Approaches to computer mediated team work.' Automation in Construction 7(6): 465-473.

Gurney, K. (1997). An Introduction to Neural Networks. London, UCL Press.

Hagan, M. T., H. B. Demuth, et al. (1996). Neural Network Design. Boston, PWS Publishing Company.

Haken, H. (1987). Advanced Synergetics. Berlin, Springer.

Harris, B. (1972). Foreward. Decision-Making in Urban Planning. I. M. Robinson. Beverly Hills, CA, SAGE Publications: 9-14.

Harsanyi, J. C. and R. Selten (1988). A General Theory of Equilibrium Selection in Games, MIT Press.

Helbing, D. and P. Molnár (1995). Self-organization phenomena in pedestrian crowds. Self-Organization of Complex Structures: From Individual to Collective Dynamics. F. Schweitzer. London, Gordon and Breach: 569-577.

Hogg, L. M. J. and N. R. Jennings (2001). 'Socially intelligent reasoning for autonomous agents.' IEEE Transactions on Systems, Man and Cybernetics: Part A: Systems and Humans 31(5): 381-393.

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

Holland, J. H. (1995). Hidden Order. Reading, MA, Addison-Wesley.

Hopkins, L. D. (1981). The decision to plan: planning activities as collective goods. Urban Infrastructure, Location and Housing. W. R. Lierop and P. Nijkamp. Alphen aan den Rijn, Netherlands, Sijthoff and Noordhoff: 237-296.

Hopkins, L. D. (2001). Urban Development: The logic of making plans. Washington DC, Island Press.

Huber, J. (1991). Macro-Micro Linkages in Sociology. Newbury Park, California, Sage.

Hughes, J. A., J. O'Brien, et al. (1997). Designing with ethnography: a presentation framework for design Proceedings of the conference on Designing interactive systems: processes, practices, methods, and techniques Amsterdam, The Netherlands ACM Press: 147-158

Hunt, K. J. and D. Sbarbaro (1995). Studies in artificial neural network based control. Neural Network Applications in Control. G. W. Irwin, K. Warwick and J. Hunt. London, Institution of Electrical Engineers: 109-139.

Hutchins, E. (1991). The social organization of distributed cognition. Perspectives on Socially Shared Cognition. L. Resnick, J. Levine and S. Teasley. Washington, DC, American Psychological Association: 283-307.

Hutchins, E. (1995). Cognition in the Wild. Cambridge, Massachusetts, MIT.

Hutchins, E. and B. Hazlehurst (1995). How to invent a lexicon: The development of shared symbols in interaction. Artificial Societies: The computer simulation of social life. N. Gilbert and R. Conte. London, UCL Press: 157-189.

Innes, J. E. (1996). 'Planning through consensus building: A new view of the comprehensive planning ideal.' Journal of the American Planning Association 62(4): 460-472.

Innes, J. E. and D. E. Booher (2000). 'Indicators for sustainable communities: A strategy building on complexity theory and distributed intelligence.' Planning Theory & Practice 1(2): 173-186.

Jacobs, O. L. R. (1993). Introduction to Control Theory. New York, Oxford Science Publications.

Jennings, N. R. (1993). 'Commitments and conventions: The foundation of coordination in multi-agent systems.' The Knowledge Engineering Review 8(3): 223-250.

Jennings, N. R. (1996). Coordination techniques for distributed artificial intelligence. Foundations of Distributed Artificial Intelligence. G. M. P. O' Hare and N. R. Jennings, John Wiley & Sons: 187-210.

Jennings, N. R. and J. R. Campos (1997). 'Towards a social level characterisation of socially responsible agents.' IEEE Proceedings on Software Engineering 144(1): 11-25.

Johnson, C. W. (2006). 'Editorial. What are emergent properties and how do they affect the engineering of complex systems?' Reliability Engineering and System Safety 91(12): 1475-1481.

Johnson, J. H. (1995). The multidimensional networks of complex systems. Networks in Action. D. Batten, J. Casti and R. Thord. Berlin, Springer-Verlag: 49-79.

Jones, J. C. and D. Thornley, Eds. (1963). Conference in Design Methods. Oxford, Pergamon.

Kalay, Y. E. (1998). 'P3: Computational environment to support design collaboration.' Automation in Construction 8(1): 37-48.

Kannengiesser, U. and J. S. Gero (2002). Situated agent communication for design. Agents in Design 2002. J. S. Gero and F. M. T. Brazier. Sydney, Key Centre for Design Computing and Cognition, University of Sydney: 85-94.

Kauffman, S. (1995). At Home in the Universe: The Search for the Laws of Self-Organization and Complexity. New York; Oxford, Oxford University Press.

Kecman, V. (2001). Learning and Soft Computing. Support Vector Machines, Neural Networks and Fuzzy Logic Models. Cambridge, Massachusetts, The MIT Press.

Keeney, R. L. and H. Raiffa (1976). Decisions with Multiple Objectives: Preferences and Value Tradeoffs. New York, John Wiley & Sons.

Kennedy, J. (1997). The particle swarm: Social adaptation of knowledge. In Proceedings of the 1997 International Conference on Evolutionary Computation, Pistataway, NJ, IEEE.

Kim, J., Ed. (1993). Supervenience and Mind: Selected Philosophical Essays. Cambridge, Cambridge University Press.

Kim, J. (1999). 'Making sense of emergence.' Philosophical Studies 95(1): 3-36.

Klein, M. (1991). 'Supporting conflict resolution in cooperative design systems.' IEEE Transactions on Systems, Man and Cybernetics 21(6): 1379-1390.

Klein, M. and S. C. Y. Lu (1989). 'Conflict resolution in cooperative design.' Artificial Intelligence in Engineering 4(4): 168-180.

Klein, M., H. Sayama, et al. (2002). 'A complex systems perspective on computer-supported collaborative design technology.' Communications of the ACM 45(1): 27-31.

Kleindorfer, P. R., H. C. Kunreuther, et al. (1993). Decision Sciences. An integrative perspective. Cambridge, Cambridge University Press.

Klir, G. J. and B. Yuan (1995). Fuzzy Sets and Fuzzy Logic: theories and applications. New Jersey, Prentice-Hall.

Knight, T. (2003). 'Computing with emergence.' Environment and Planning B: Planning and

Design 30(1): 125-155.

Knorr-Cetina, K. (1981). Introduction. The micro-sociological challenge of macro-sociology: towards a reconstruction of social methodology and methodology. Advances in Social Theory and Methodology. Toward an Integration of Micro and Macro Sociologies. K. Knorr-Cetina and A. V. Cicourel. Boston, London, Routledge & Kegan Paul: 1-47.

Knorr-Cetina, K. and A. V. Cicourel, Eds. (1981). Advances in Social Theory and Methodology. Toward an Integration of Micro and Macro Sociologies. Boston, London, Routledge & Kegan Paul.

Koestler, A. (1967). The Ghost in the Machine. London, Hutchinson of London.

Kolb, D. A. (1984). Experiential Learning. Experience as the source of learning and development. Englewood Cliffs, New Jersey, Prentice Hall.

Kolmogorov, A. N. (1965). 'Three approaches to the quantitative definition of Information.' Problems of Information Transmission 1(1): 1-17.

Kosko, B. (1994). Fuzzy Thinking. London, HarperCollins.

Koutamanis, A. (2001). Modeling irregular and complex forms. AVOCAAD - Added Value of Computer Aided Architectural Design Conference Proceedings. N. Koenraad, P. Tom, V. Johan and V. Johan. Brussel, Belgium, Departement Architectuur Sint-Lucas, Hogeschool voor Wetenschap en Kunst

Kubik, A. (2003). 'Toward a formalization of emergence.' Artificial Life 9(1): 41-65.

Küppers, S., S. Gage, et al. (2000). Distributed simulation system for spatial design: parallel physical and virtual interactions between people and objects. In Proceedings of Greenwich 2000: Digital Creativity Symposium, Greenwich, University of Greenwich.

Kvan, T. (1999). Designing Together Apart. Computer supported collaborative design in architecture. Department of Design and Innovation. Milton Keynes, Open University. PhD Thesis.

Kvan, T. (2000). 'Collaborative design: what is it?' Automation in Construction 9(4): 409-415.

Kvan, T. and J. Yunyan (2005). 'Students' learning styles and their correlation with performance in architectural design studio.' Design Studies 26(1): 19-34.

Lai, S. K. (1998). 'From organized anarchy to controlled structure: effects of planning on the garbage-can decision process.' Environment and Planning B: Planning and Design 25(1): 85-102.

Lai, S. K. (2006). 'A spatial garbage-can model.' Environment and Planning B: Planning and Design 33(1): 141-156.

Lai, S. K. and C. I. Tang (1995). Effects of planning in a garbage can decision process: a simulation comparison. In Proceedings of the Fourth International Conference on Computers in Urban Planning and Urban Management (CUPUM 1995), Melbourne, Australia, Department of Geography and Environmental Studies, University of Melbourne.

Landau, Y. D. (1979). Adaptive Control: the model reference approach. New York and Basel, Marcel Dekker.

Lane, D. (2006). Hierarchy, Complexity, Society. Hierarchy in Natural and Social Science. D. Pumain. Dordrecht, The Netherlands, Springer. 3: 81-119.

Lave, J. and E. Wenger (1991). Situated Learning: Legitimate Peripheral Participation. Cambridge, Cambridge University Press.

Lawson, B., M. Bassanino, et al. (2003). 'Intentions, practices and aspirations: understanding learning in design.' Design Studies 24(4): 327-339.

Leclercq, P. and R. Juchmes (2002). 'The absent interface in design engineering.' Artificial Intelligence in Engineering Design, Analysis and Manufacturing (AIEDAM), Special Issue: Human-Computer Interaction in Engineering Contexts, Parmee, I.C., Smith, I.F.C. (eds) 16(3).

Lees, B., C. Branki, et al. (2001). 'A framework for distributed agent-based engineering design support.' Automation in Construction 10(5): 631-637.

Lesser, V. R. (1999). 'Cooperative Multiagent Systems: A personal view of the state of the art.' IEEE Transactions on Knowledge and Data Engineering 11(1): 133-142.

Levesque, H. J., P. R. Cohen, et al. (1990). On acting together. In Proceedings of the 8th National Conference on Artificial Intelligence (AAAI-90), Boston, MA, AAAI.

Liew, P.-S. and J. S. Gero (2002). An implementation model of constructive memory for a design agent. Agents in Design 2002. J. S. Gero and F. M. T. Brazier. Sydney, Key Centre for Design Computing and Cognition, University of Sydney: 257-276.

Ligtenberg, A., A. K. Bregt, et al. (2001). 'Multi-actor-based land use modelling: spatial planning using agents.' Landscape and Urban Planning 56(1): 21-33.

Liu, H., M. Tang, et al. (2001). 'Supporting learning in a shared design environment.' Advances in Engineering Software 32(4): 285-293.

Liu, H., M. Tang, et al. (2002). 'Supporting evolution in a multi-agent cooperative design environment.' Advances in Engineering Software 33(6): 319-328.

Liu, Y.-T. (1998). 'Personal versus cultural cognitive models of design creativity.' International Journal of Technology and Design Education 8(2): 185-195.

Lloyd, P. (2003). 'Special Issue on Designing in Context.' Design Studies 24(3).

Lucas, R. E. J. (1986). Adaptive behavior and economic theory. Rational Choice: The contrast between economics and phychology. J. H. Kagel and A. E. Roth. Chicago, University of Chicago Press: 217-242.

Luce, D. R. and H. Raiffa (1957). Games and Decisions: Introduction and critical survey. New York, John Wiley & Sons.

Luhmann, N. (1986). The Autopoiesis of Social Systems. Sociocybernetic Paradoxes. F. Geyer and J. van der Zouwen. London, Sage: 172-192.

Maes, P. (1992). Behavior-based artificial intelligence. In Proceedings of the Second International Conference on Simulation of Adaptive Behavior: From Animals to Animats 2, Cambridge, Massachusetts, MIT Press/Bradford Books.

Maher, M. L. (1994). 'Creative design using a genetic algorithm.' Computing in Civil Engineering: 2014-2021.

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

Maher, M. L., D. C. Brown, et al. (1994). 'Special Issue: Machine Learning in Design.' Artificial Intelligence for Engineering Design Analysis and Manufacturing (AIEDAM) 8(2).

Maher, M. L. and J. Poon (1995). Co-evolution of the fitness function and design solution for design exploration. Proceedings of the IEEE International Conference on Evolutionary Computing, IEEE. 1: 240-244.

Maimon, O. and D. Braha (1996). 'On the complexity of the design synthesis problem.' IEEE Transactions on Systems, Man and Cybernetics 26(1): 142-151.

Malone, T. W. and K. Crowston (1990). What is coordination theory and how can it help design cooperative work systems? CSCW 90, Los Angeles, CA USA, ACM.

Malone, T. W. and K. Crowston (1994). 'The interdisciplinary study of coordination.' ACM Computing Surveys 26(1): 87-119.

March, J. G. and J. P. Olsen (1975). 'The uncertainty of the past: Organizational learning under ambiguity.' European Journal of Political Research 3: 147-171.

March, L. (1976). The Logic of Design and the Question of Value. The Architecture of Form. L. March. Cambridge Cambridge University Press: 1-40.

Massam, B. H. (1988). 'Multi-criteria Decision Making (MCDM) Techniques in Planning.' Progress in Planning 30(1).

Massam, B. H. (1993). The Right Place: Shared responsibility and the location of public facilities. London, Longman.

Mattessich, P. W. and B. R. Monsey (1992). Collaboration: What makes it work. St. Paul, Minnesota, Amherst H. Wilder Foundation.

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

McCarthy, J. (2002). 'What is Artificial Intelligence?' 2002 edition. Retrieved 10/3/03, 2002, from http://www-formal.stanford.edu/jmc/whatisai/whatisai.html.

McElreath, R., M. Lubell, et al. (2005). 'Applying evolutionary models to the laboratory study of social learning.' Evolution and Human Behavior 26(6): 483-508.

McLoughlin, B. (1977). Control and Urban Planning. London, Faber and Faber.

McLoughlin, B. P. (1997). 'Emergence and Supervenience.' Intellectica 25(1): 25-43.

Morgan, L. (1923). Emergent Evolution. London, Williams & Norgate.

Moschinski, J. and Z. Ogonowski (1995). Advanced Control with Matlab & Simulink. Hemel Hempstead, Ellis Horwood.

Moulin, B. and B. Chaib-Draa (1996). An overview of Distributed Artificial Intelligence. Foundations of Distributed Artificial Intelligence. G. M. P. O'Hare and N. R. Jennings. New York, John Wiley & Sons.

Münch, R. and N. J. Smelser (1987). Relating the Micro and Macro. The Micro-Macro Link. J. C. Alexander, B. Giesen, R. Münch and N. J. Smelser. Berkeley, California, University of California Press: 356-387.

Nardi, B. A. (1995). Studying context: a comparison of activity theory, situated action models, and distributed cognition. Context and Consciousness: activity theory and human-computer interaction. B. A. Nardi. Cambridge, Mass, MIT Press.

Narenda, K. S. (1996). 'Neural networks for control: theory and practice.' Proceedings of the IEEE **84**(10): 1385-1406.

Newell, A. (1968). On the analysis of human problem solving protocols. Calcul et Formalisation dans le Sciences de L'Homme. J. C. Gardin and B. Jaulin. Paris, Centre National de la Recherche Scientifique.

Newell, A. (1990). Unified Theories of Cognition. Cambridge, Massachusetts, Harvard University Press.

Newell, A., J. C. Shaw, et al. (1957). Elements of a Theory of Problem Solving, Rand Corporation Report P-971.

Newton, S. and R. D. Coyne (1991). The impact of connectionist systems on design. Artificial Intelligence in Design '91. J. S. Gero. Oxford, Butterworth- Heinemann.

Nicolis, G. and I. Prigogine (1977). Self-Organization in Non-Equilibrium Systems. New York, Wiley.

Norman, D. A. (1993). Things That Make Us Smart: defending human attributes in the age of the machine. Reading, Massachusetts, Addison-Wesley, Perseus.

Ochs, J. (1995). Coordination problems. Handbook of Experimental Economics. J. H. Kagel and A. E. Roth. Princeton, Princeton University Press: 195-249.

O'Connor, T. and H. Y. Wong (2002). Emergent Properties. The Stanford Encyclopedia of Philosophy. E. N. Zalta.

Osborn, A. F. (1963). Applied Imagination: principles and procedures of creative problem solving New York, NY, Charles Scribner's Sons.

Ossowski, S. (1999). Co-ordination in Artificial Agent Societies. Berlin Heidelberg, Springer-Verlag.

O'Sullivan, D. and M. Haklay (2000). 'Agent-based models and individualism: is the world agent-based?' Environment and Planning A **32**(8): 1409-1425.

Pahl, G. and W. Beitz (1984). Engineering Design. London, Design Council Books.

Parisi, D., F. Cecconi, et al. (1995). Kin-directed altruism and attachment behaviour in an evolving population of neural networks. Artificial Societies: The computer simulation of social life. N. Gilbert and R. Conte. London, UCL Press: 238-251.

Pattee, H. H. (1988). Simulations, Realizations, and Theories of Life. Artificial Life. C. Langton. Reading, Massachusetts, Addison Wesley Publishing Company. VI.

Peacocke, A. (2003). Complexity, Emergence and Divine Creativity. From Complexity to Life. On the Emergence of Life and Meaning. N. H. Gregersen. Oxford, New York, Oxford University Press: 187-205.

Peirce, C. S. (1934). Collected Papers of Charles Sanders Peirce, Vol 5 Pragmatism and Pragmaticism. Cambridge, Massachusetts, Harvard University Press.

Peirce, C. S. (1958). Collected Papers of Charles Sanders Peirce, Vol 7 Science and Philosophy. Cambridge, Massachusetts, Harvard University Press.

Perry, M. (1997). Distributed Cognition and Computer Supported Collaborative Design: The Organisation of Work in Construction Engineering, Brunel University. **PhD Thesis.**

Perry, M. (1999). The application of individually and socially distributed cognition in workspace studies: two peas in a pod? Proceedings of the European Conference on Cognitive Science. S. Bagnara. Siena, Italy: 87-92.

Perry, M. and D. Sanderson (1998). 'Coordinating joint design work: the role of communication and artefacts.' Design Studies 19(3): 273-288.

Poon, J. and M. L. Maher (1997). 'Co-evolution and emergence in design.' Artificial Intelligence in Engineering 11(3): 319-327.

Portugali, J. (2000). Self-Organization and the City. Berlin, Heidelberg, Springer-Verlag.

Prietula, M. J., K. M. Carley, et al., Eds. (1998). Simulating Organizations: Computational models of institutions and groups. Menlo Park, CA, AAAI Press/The MIT Press.

Qin, S. and J. Johnson (2005). 'Special issue on Exploring Complexity in Collaborative Design and Solutions.' CoDesign 1(4).

Rapoport, A., Ed. (1974). Game Theory as a Theory of Conflict Resolution. Dodrecht, D. Reidel Publishing Company.

Rapoport, A., D. A. Seale, et al. (1998). 'Equilibrium play in large group market entry games.' Management Science 44(1): 119-141.

Reffat, R. M. (2000). Computational Situated Learning in Designing. Department of Architectural and Design Science, Faculty of Architecture. Sydney, University of Sydney. **PhD Thesis.**

Reffat, R. M. and J. S. Gero (1999). Situatedness: A new dimension for learning systems in design. Architectural Computing: From Turing to 2000 (eCAADe '99), Liverpool, University of Liverpool.

Resnick, M. (1994). Turtles, Termites, and Traffic Jams : Explorations in massively parallel microworlds. Cambridge, Massachusetts, MIT Press.

Rittel, H. and M. Webber (1973). 'Dilemmas in a general theory of planning.' Policy Sciences 4: 155-169.

Rittel, H. W. J. (1984b). Second-generation design methods. Developments in Design Methodology. N. Cross. New York, John Wiley & Sons: 317-327.

Rittel, H. W. J. and M. M. Webber (1984a). Planning problems are wicked problems. Developments in Design Methodology. N. Cross. New York, John Wiley & Sons: 135-144.

Ritzer, G. (2000). Modern Sociological Theory. Boston, McGraw Hill.

Ronald, E. M. A., M. Sipper, et al. (1999). 'Design, observation, surprise! A test of emergence.' Artificial Life 5(3): 225-239.

Roozenburg, N. F. M. (1993). 'On the pattern of reasoning in innovative design.' Design Studies 14(1): 4-18.

Roozenburg, N. F. M. and J. Eekels (1995). Product Design: Fundamentals and Methods. Chichester, John Wiley & Sons.

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.

Rosenman, M. A. and J. S. Gero (1998). 'Purpose and function in design: from the socio-cultural to the techno-physical.' Design Studies 19(2): 161-186.

Rosenman, M. A. and J. S. Gero (1999). 'Purpose and function in a collaborative CAD environment.' Reliability Engineering and System Safety 64(2): 167-179.

Rosenschein, J. S. and G. Zlotkin (1994). Rules of Encounter: Designing conventions for automated negotiation among computers. Cambridge, Massachusetts, MIT Press.

Ross, D. (2002). 'Game Theory, Stanford Encyclopedia of Philosophy.' 2002 edition. Retrieved 30/1/03, from http://plato.stanford.edu/archives/spr2002/entries/game-theory/.

Rowe, P. G. (1987). Design Thinking. Cambridge, Massachusetts, The MIT Press.

Russell, S. J. and P. Norvig (1995). Artificial Intelligence: A modern approach. Englewood Cliffs, NJ, Prentice Hall.

Saarlos, D. J. M., T. A. Arentze, et al. (2001). Introducing multi-agents in an integrated GIS/VR system for supporting urban planning and design decisions. In Proceedings of the 7th International Conference on Computers in Urban Planning and Urban Management CUPUM 2001 [CD-ROM], Hawaii, Honolulu, University of Hawaii.

Salomon, G., Ed. (1993). Distributed Cognitions: Psychological and Educational Considerations. Cambridge and New York, Cambridge University Press.

Salomon, G. and D. N. Perkins (1998). 'Individual and social aspects of learning.' Review of Research in Education 23(1): 1-24.

Salomon, G., D. N. Perkins, et al. (1991). 'Partners in cognition: extending human intelligence with intelligent technologies.' Educational Researcher 20(3): 2-9.

Sandholm, T. W. (1999). Distributed rational decision making. Multiagent Systems: A modern approach to distributed artificial intelligence. G. Weiss. Cambridge, Massachusetts, MIT Press: 210-258.

Sanoff, H. (2000). Community Participation Methods in Design and Planning New York, John Wiley & Sons.

Sanoff, H. (2007). 'Special Issue on Participatory Design.' Design Studies 28(3): 213-340.

Saunders, R., Gero, J.S. (2002a). How to study artificial creativity. In Proceedings of the 4th International Conference on Creativity & Cognition (C&C'02), Loughborough University, ACM.

Saunders, R. and J. S. Gero (2001). Artificial creativity: A synthetic approach to the study of creative behaviour. Computational and Cognitive Models of Creative Design V, Sydney, Key Centre for Design Computing and Cognition, University of Sydney.

Saunders, R. and J. S. Gero (2002). Curious agents and situated design evaluations. Agents in Design 2002. J. S. Gero and F. M. T. Brazier. Sydney, Key Centre for Design Computing and Cognition, University of Sydney: 133-149.

Sawyer, R. K. (2000). Simulating Emergence and Donward Causation in Small Groups. Multi-Agent-Based Simulation: Second International Workshop, MABS 2000, LNCS 1979. S. Moss and P. Davidsson. Berlin, Heidelberg, Springer Verlag. **1979**: 49-67.

Sawyer, R. K. (2001). 'Emergence in sociology: contemporary philosophy of mind and some implications for sociological theory.' AJS **107**(3): 551-585.

Sawyer, R. K. (2003). 'Artificial societies. Multiagent systems and the micro-macro link in sociological theory.' Sociological Methods & Research **31**(3): 325-363.

Sawyer, R. K. (2004). 'The mechanisms of emergence.' Philosophy of the Social Sciences **34**(2): 260-282.

Sawyer, R. K. (2004). 'Social explanation and computational simulation.' Philosophical Explorations **7**(3): 219-231.

Schelhorn, T., D. O'Sullivan, et al. (1999). 'STREETS: An agent-based pedestrian model, CASA Working Paper 9.' from http://www.casa.ucl.ac.uk/streets.pdf.

Schelling, T. C. (1969). 'Models of segregation.' American Economic Review, Papers and Proceedings **58**(2): 488-493.

Schelling, T. C. (1978). Micromotives and Macrobehavior. New York, W. W. Norton and Company.

Schelling, T. C. (1980). The Strategy of Conflict. Cambridge, Massachusetts, Harvard University Press.

Schillo, M., K. Fischer, et al. (2000). The Micro-Macro Link in DAI and Sociology. Multi-Agent-Based Simulation: Second International Workshop, MABS 2000, LNCS 1979. S. Moss and P. Davidsson. Berlin, Heidelberg, Springer Verlag. **1979**: 133-148.

Schön, D. and D. Wiggins (1992). 'Kinds of seeing and their functions in designing.' Design Studies **13**(2): 135-156.

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

Sen, A. K. (1979). Collective Choice and Social Welfare. Amsterdam, The Netherlands, Elsevier Science Publishers.

Sen, S. and G. Weiss (1999). Learning in multiagent systems. Multiagent Systems: A modern approach to distributed artificial intelligence. G. Weiss. Cambridge, Massachusetts, MIT Press: 259-298.

Senge, P. (1990). The Fifth Discipline. The art and practice of the learning organization. London, Random House.

Shen, W., D. H. Norrie, et al. (2001). Multi-Agent Systems for Concurrent Intelligent Design and Manufacturing. London and New York, Taylor & Francis.

Shih, S.-G., T.-P. Hu, et al. (2006). 'A game theory-based approach to the analysis of cooperative learning in design studios.' Design Studies 27(6): 711-722.

Shoham, Y. and M. Tennenholtz (1992). On the synthesis of useful social laws for artificial agent societies. In Proceedings of the 10th National Conference on Artificial Intelligence (AAAI-92), San Diego, CA, AAAI.

Sim, S. K. and A. H. B. Duffy (2004). 'Evolving a model of learning in design.' Research in Engineering Design 15(1): 40-61.

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

Simon, H. A. (1969,1996). The Sciences of the Artificial. Cambridge, Massachusetts, The MIT Press.

Simon, H. A. (1991). 'Bounded rationality and organizational learning.' Organization Science 2(1): 125-134.

Sloman, A. (2000). Architectural requirements for human-like agents both natural and artificial: what sorts of machines can love? Human Cognition and Social Agent Technology. K. Dautenhahn. Amsterdam/Philadelphia, John Benjamins: 163195.

Smithers, T. (1998). Towards a knowledge level theory of design process. Artificial Intelligence in Design '98. J. S. Gero and F. Sudweeks. Dordrecht, Kluwer Academic Publishers.

Smithers, T. (2002). Synthesis in designing. Artificial Intelligence in Design '02 (AID '02). J. S. Gero. Dordrecht, Kluwer Academic Publishers: 3-24.

Sosa, R. (2005). Computational Explorations of Creativity and Innovation in Design. Sydney, University of Sydney. **PhD Thesis.**

Sosa, R. and J. S. Gero (2002). Creative individuals or creative situations? A computational model of creativity based on social influence. SIGraDi Caracas 2002, Caracas, Ediciones Universidad Central de Venezuela.

Sosa, R. and J. S. Gero (2005). 'A computational study of creativity in design: the role of society.' Artificial Intelligence for Engineering Design Analysis and Manufacturing (AIEDAM) 19(4): 229-244.

Starkey, K. (1996). How Organizations Learn. London, International Thompson Business Press.

Steadman, J. P. (1979). 'The history and science of the artificial.' Design Studies 1(1): 49-58.

Steadman, J. P. (1998). 'Sketch for an archetypal building.' Environment and Planning B: Planning and Design 27(Anniversary Issue): 92-105.

Steadman, J. P. and L. Waddoups (2000). A catalogue of built forms using a binary representation. Proceedings of the 5th International Conference on Design and Decision Support Systems in Architecture (DDSS). The Netherlands, Einhoven University: 353-373.

Steels, L. (1994). 'The artificial life roots of artificial intelligence.' Artificial Life 1(1).

Steels, L. and R. A. Brooks, Eds. (1993). The Artificial Life Route to Artificial Intelligence: Building embodied situated agents. Hillsdale, New Jersey, Lawrence Erlbaum Associates.

Stephan, A. (1999). 'Varieties of emergentism.' Evolution and Cognition 5(1): 49-59.

Stone, P. and M. Veloso (2000). 'Multiagent Systems: A survey from a machine learning perspective.' Autonomous Robots 8(3): 345-383.

Stumpf, S. C. and J. T. McDonnell (2002). 'Talking about team framing: using argumentation to analyse and support experiential learning in early design episodes.' Design Studies 23(1): 5-23.

Suchman, L. A. (1987). Plans and Situated Actions: The problem of human-machine communication. Cambridge, Cambridge University Press.

Suh, N. P. (1999). 'A theory of complexity, periodicity and the design axioms.' Research in Engineering Design 11(2): 116-131.

Sun, R. (2001). 'Cognitive science meets multi-agent systems: a prolegomenon.' Philosophical Psychology 14(1): 5-28.

Sun, R., Ed. (2006). Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation. New York, Cambridge University Press.

Sun, R. e. (2001). 'Special Issue on multidisciplinary studies of multi-agent learning.' Cognitive Systems Research 2(1): 1-96.

Susi, T. and T. Ziemke (2001). 'Social cognition, artefacts and stigmergy: A comparative analysis of theoretical frameworks for the understanding of artefact-mediated collaborative activity.' Journal of Cognitive Systems Research 2(4): 273-290.

Suwa, M., J. S. Gero, et al. (1999). Unexpected discoveries and s-inventions of design requirements. In Computational Models of Creative Design IV, Key Centre of Design Computing and Cognition, University of Sydney, Sydney, Australia.

Sycara, K. P. (1988). Resolving goal conflicts via negotiation. In Proceedings of the AAAI-88, St.Paul, MN, AAAI.

Sycara, K. P. (1998). 'Multiagent Systems.' AI Magazine 19(2): 79-92.

Tessier, C., L. Chaudron, et al., Eds. (2001). Conflicting Agents: Conflict management in Multi-Agent Systems. Dordrecht, The Netherlands, Kluwer Academic Publishers.

Theraulaz, G. and E. Bonabeau (1999). 'A brief history of stigmery.' Artificial Life 5(2): 97-116.

Torrens, P. M. (2001). 'Can Geocomputation save urban simulation? Throw some agents into the mixture, simmer and wait ... , CASA Working Paper 32.' from http://www.casa.ucl.ac.uk/paper32.pdf.

Turoff, M. and H. A. Linstone (1975). The Delphi Method Techniques and Applications. Reading, Mass, Addison-Wesley.

Umeda, Y. and T. Tomiyama (1997). 'Functional reasoning in design.' IEEE Expert 12(2):

42-48.

Urry, J. e. (2005). 'Special Issue on Complexity.' Theory, Culture & Society 22(5).

Van Huyck, J. B., R. C. Battalio, et al. (1990). 'Tacit coordination games, strategic uncertainty and coordination failure.' The American Economic Review 80(1): 234-248.

Van Loon, P. P. (2000). Team design from the individual point of view: A humanistic approach. In Proceedings of the 5th International Conference in Design and Decision Support Systems in Architecture, Eindhoven University of Technology.

Varela, F. (1979). Principles of Biological Autonomy. New York, Elsevier/North-Holland.

Verhagen, H. (2000). Norm Autonomous Agents. Department of Computer and Systems Sciences. Stockholm, The Royal Institute of Technology and Stockholm University. **Licentiate PhD Thesis.**

Von Foerster, H. (1984). Observing Systems. Salinas, California, Intersystems Publications.

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

Voogd, H. (1983). Multicriteria Evaluation for Urban and Regional Planning. London, Pion Limited.

Vygotsky, L. S. (1978). Mind in Society: the Development of Higher Psychological Processes. Cambridge, Mass; London, Harvard University Press.

Wates, N. (2000). The Community Planning Handbook. London, Earthscan.

Weiss, G., Ed. (1999). Multiagent Systems: A modern approach to distributed artificial intelligence. Cambridge, Massachusetts, MIT Press.

Wenger, E. (1998). Communities of Practice. Learning, meaning and identity. New York, Cambridge University Press.

White, R. and G. Engelen (2000). 'High-resolution integrated modelling of the spatial dynamics of urban and regional systems.' Computers, Environment and Urban Systems 24(5): 383-400.

Whitfield, R. I., G. Coates, et al. (2000). 'Coordination approaches and systems - Part I: A strategic perspective.' Research in Engineering Design 12(1): 48-60.

Wiener, N. (1948). Cybernetics, or control and communication in the animal and the machine. Cambridge, Massachusetts, The MIT Press.

Winograd, T. and F. Flores (1986). Understanding Computers and Cognition: A new foundation of design. Norwood, NJ, Ablex Publishing Corporation.

Woltjer, J. (2000). Consensus Planning: The relevance of communicative planning theory in Dutch infrastructure development. Aldershot, Ashgate.

Wooldridge, M. J. (2002). An Introduction to MultiAgent Systems. London, John Wiley & Sons.

Wooldridge, M. J. and N. R. Jennings (1995). 'Intelligent agents: Theory and practice.' Knowledge Engineering Review 10(2): 115-152.

Wu, Z. and A. H. B. Duffy (2004). 'Modeling collective learning in design.' Artificial Intelligence for Engineering Design Analysis and Manufacturing (AIEDAM) **18**(4): 289-313.

Xiang, W.-N. (1993). 'A GIS/MMP-based coordination model and its application to distributed environmental planning.' Environment and Planning B: Planning and Design **20**(2): 195-220.

Zadeh, L. (1965). 'Fuzzy sets.' Information and Control **8**: 338-353.

Zamenopoulos, T. (2007). Design out of Complexity: the mathematical basis for an organisational level theory of design. Bartlett School of Graduate Studies. London, University College London. **PhD Thesis.**

Zamenopoulos, T. and K. Alexiou (2002). Learning to be creative and the creative memory: A discussion motivated by a control based coordination model. Learning and Creativity Workshop notes, AID '02 Conference, Cambridge, UK.

Zamenopoulos, T. and K. Alexiou (2003a). Computer-aided creativity and learning in distributed cooperative human-machine networks. In CAAD Futures 2003, Dordrecht, The Netherlands, Kluwer Academic Publishers.

Zamenopoulos, T. and K. Alexiou (2003b). Structuring the plan design process as a coordination problem: the paradigm of distributed learning control coordination. Advanced Spatial Analysis: The CASA book of GIS. P. Longley and M. Batty. Redlands, California, ESRI Press: 407-426.

Zamenopoulos, T. and K. Alexiou. (2004). 'Design and anticipation: towards an organisational view of design systems, CASA Working Paper 76.' from http://www.casa.ucl.ac.uk/working_papers/paper76.pdf

Zamenopoulos, T. and K. Alexiou (2005). Linking design and complexity: a review. Proceedings of the ECCS 2005 Satellite Workshop Embracing Complexity in Design. J. Johnson, T. Zamenopoulos and K. Alexiou. Paris 17 November 2005, The Open University: 91-102.

Zamenopoulos, T. and K. Alexiou (2007a). 'Towards an anticipatory view of design.' Design Studies **28**(4): 411-436.

Zamenopoulos, T. and K. Alexiou (2007b). 'Rethinking the cybernetic basis of design: the concepts of control and organization.' Kybernetes **forthcoming.**

Zeleny, M. (1982). Multiple Criteria Decision Making. New York, McGraw-Hill.

Zentall, T. and J. B. G. Galef, Eds. (1988). Social Learning. Hillsdale, NJ, Erlbaum.

Zhang, J. and D. A. Norman (1994). 'Representations in distributed cognitive tasks.' Cognitive Science **18**: 87-122.