The social cognitive actor

Helmhout, M.

CHAPTER 1

Introduction

MULTI-AGENT SYSTEMS (MAS) is a field of research that is concerned with the design, analysis and implementation of computational decision and simulation systems—a collection of agents that work in conjunction with each other—to enhance one's understanding of how people cooperate and coordinate their actions to solve problems. Upon its foundation in the mid-nineties, MAS propagated in the fields of cognitive science and social science, and was adopted, more specifically, in cognitive modelling and modelling of multi-agent interaction as in, for instance, social simulation (Conte & Gilbert, 1995; Troitzsch, 1997). However, there has not been enough interaction between these two fields, and tools that try to connect both fields have not been sufficiently developed (Sun, 2006b). In our research, we will address this problem and make an attempt to bring both fields—cognitive science and social science—closer together.

This chapter is structured as follows. Section 1.1 discusses the motivation for selecting organisational, social and cognitive theories and the need for the development of a cognitive agent-based computational social simulation model. Section 1.2 presents the research background of the dissertation. In section 1.3, we will elaborate on the methodology applied in this research. Next, the research questions and objectives are outlined in section 1.4 and we end this chapter with section 1.5 that provides an outline of the dissertation.

1.1 Motivation

Social sciences have adopted MAS for the creation of social simulations in order to shed light on organisational and sociological phenomena. However, as Castelfranchi (2001) points out, a significant theoretical problem exists in the field of social sciences; there is a lack of understanding or explanation of unconscious, unplanned forms of cooperation among intentional agents. Moreover, Moss (2006) states that science that eschews observation in favour of formalism is rife in social science. His claim is that agent-based models can capture independent validation by "...ensuring that the specification is well verified with respect to formal models from cognitive science understanding that those formal models are themselves well validated experimentally and observationally" (ibid., p. 398).

According to Castelfranchi (2001), Artificial Intelligence (AI) and, in particular, MAS—the cognitive modelling of agents and its learning aspects—can contribute to the simulation of *artificial societies*[1]. On the other hand, cognitive science needs Multi-Agent Systems, social simulation, and social sciences in general (Sun, 2006b). There is a real need for new complex cognitive models that take into account social factors in cognition. In other words, there is a shortage of good ideas and theories that address socio-cultural concepts/signs/symbols within social structures from a cognitive standpoint (Sun, 2001). Cognitive agent-based social simulations have adopted MAS to give insights into social-level phenomena based on the individual agent's actions, i.e. the cognitive level.

The aim of this dissertation is to design and implement a cognitive agent-based computational social simulation model based on a selection of social and cognitive theories in an attempt to satisfy the need for a complex cognitive-social model. In comparison to models emanating from traditional social (and economic) sciences, cognitive-social models enable one to gather, organise, and interpret observations and data more directly and test various factors more thoroughly (Sun, 2006b). Furthermore, such models are validated by incorporating behavioural descriptions and previously validated concepts from cognitive science and social psychology at the micro level, and by capturing the statistical output of (longitudinal) studies in the field of social science at the macro level (Moss, 2006).

The motivation for this research is twofold: (1) to reveal the constituents of MAS and the social cognitive agent (cf. Conte & Castelfranchi, 1995a) based on theoretical considerations to (2) construct a simulation model that plausibly explains the interactive social behaviour of agents in a physical and socially situated environment (cf. Gilbert & Troitzsch, 1999). The aim is to relate the behaviour of individuals (micro level) that form a group or an organisation to the behaviour and performance of a group or organisation as a whole (Alexander & Giesen, 1987; Van den Broek, 2001). This rests on the 'methodological individualism' assumption that supports MAS methodology:

> "*all* description and explanation of social phenomena should ultimately be in terms of individuals, their properties, and their interrelations in terms of these properties (Franssen, 1997)" (Van den Broek, 2001, p. 26).

Similarly, it is established in social constructivism (Mead, 1934) that society and all its related issues, such as money and marriage, are products of (symbolic) social interaction between individuals. These interactions are caused by

[1]Artificial societies is used here to refer to social simulations using MAS (Minsky, 1985; Sawyer, 2003).

actions of individuals as part of a group. Accordingly, the individual and its interactions with others are the objects of study in this dissertation. Semiotics, as a study of signs and sign systems, is also addressed in this dissertation and is concerned with the *'symbolic interaction'* of individuals and the (shared) meaning that they assign to constructs (i.e. semiotic codes) which they create through semiosis[2].

In MAS, it is common to define varying *levels of description*; for instance from the biological, cognitive, rational to the social level of description (Newell, 1990). Social and economic sciences often operate at the social level. This is acceptable so far as the properties of actors at lower levels are assumed to be constant (Gilbert, 2006). However, when the analysis of lower levels constrains that at higher levels, or when there is a need to model the social and cognitive levels, then a 'mixed-level' analysis is necessary (Sun, 2006b). Under that circumstance, descriptions at one level can be applied to other levels (Gilbert, 2006).

In this dissertation, a mixed-level analysis is performed, in which the following levels are discerned in descending order:

1. The *social level* is involved with social laws and overall behaviour of the social system. The advantage of such a level "is that it enables the overall system's behaviour to be studied without having to delve into the implementation of the individual [actors]" (Jennings & Campos, 1997, p. 3), e.g. it can be explained with the help of population statistics or social network characteristics. However, we argue, as this dissertation will make clear, that the individual and other levels of description are important for a better explanation, understanding and prediction of social behaviour (methodological individualism).

2. The *semiotic level*[3], which describes the use of language and signs in communication, interaction and negotiation in order to agree on social constructs (e.g. common plans, or contracts) (Gazendam, 2004; Helmhout et al., 2004; Helmhout, Gazendam, & Jorna, 2005b).

3. The *intentional level*, which ascribes beliefs, desires and intentions to actors. Intentions of others, inferred from knowledge about others' beliefs and desires, enable the examination of others' actions (Dennett, 1987).

4. The *functional level* describes learning and cognitive mechanisms of the actor and is grounded in an empirically validated theory (cf. Anderson & Lebiere, 1998).

5. The *physical/physiological level* is a level described by appealing to physics, biology and chemistry. It predicts or explains behaviour in terms of physical laws or physiological properties (Dennett, 1987).

[2]Semiosis encompasses the triadic notion of meaning processes, i.e. the interaction between representamen, the object and the interpretant (Peirce, 1931).

[3]Actually, semiotics can be assigned to the other levels as well. For instance in cognitive science, the symbol system hypothesis is used for descriptions at the functional level. The semiotic level is distinguished, because organisational semiotics focuses on this level in which social constructs, speech acts, interaction scenarios, shared knowledge units, and so on can be defined (Helmhout, Gazendam, & Jorna, 2004).

A complete design and implementation of a MAS has to incorporate the basic aspects of all the levels listed above. As signs and symbols are common to all levels, MAS and its agents should support the processing of signs and symbols. The implication for the overall system is that:

1. The actor should be able to process and create signs and symbols, i.e. a cognitive plausible agent is a requirement.
2. The environment should support:
   (a) Physical objects that represent signs and symbols.
   (b) A communication and social environment that provides an infrastructure for the transfer of signs and symbols and allows actors to interact in a coordinated manner.
3. The actor must have the capabilities to (socially) construct the world and use a common medium, or language, to exchange, negotiate and reach agreements about (social) constructs with other actors.

A simulation with a MAS can provide (1) valuable insights into and description of organisational behaviour, cognitive and social phenomena (for instance, emergence of organisations), (2) locate errors and gaps in verbal theories, (3) demonstrate which theoretical propositions are logically consistent, (4) predict the state of an agent or system in the future, and (5) discover, formalise and test (new) theories based on the simulation outcomes (Carley, Prietula, & Lin, 1998; Gilbert & Troitzsch, 1999).

In this dissertation, we will provide a selection of theories that are applicable for the design of a MAS. We will construct a model of a MAS, named MRS (Multi-RBot System) and a cognitive and social actor, named RBot, by drawing on these theories. An overview of these theories is provided in the following section.

1.2 Research background

This research draws on many theories or research fields such as organisational behaviour and structures, Computational & Mathematical Organization Theory (CMOT) (Carley & Prietula, 1994a; Carley & Gasser, 1999), organisational semiotics (Gazendam & Liu, 2005), cognitive science/artificial intelligence, social constructivist theory and MAS.

CMOT enables the modelling of business organisations, its organisational and authority structures, organisational skills and procedures depending on the position the agent takes in the organisation (Sawyer, 2003). With the help of CMOT, researchers create a formal model to make predictions about an organisation as an adaptive system composed of agents who themselves are adapting (Carley et al., 1998).

Organisational semiotics tries to understand organisations based on the use of signs, texts, documents, sign-based artefacts and communication, drawing on the basic disciplines like psychology, economics, and information systems science (Gazendam, 2004). It develops this perspective using the established discipline of semiotics, the theory of signs. The strength of semiotics lies in the theory of significant symbol processing (sense making, interpretation, knowledge, information, and culture) and provides a (systematic) way of looking at issues of meaning, interpretation and knowledge—a theory of human culture (Van Heusden & Jorna, 2001). The adjustment to one another of the acts of different human individuals within the human social process (semiosis) takes place through communication by gestures and significant symbols (Mead, 1934, p. 75). This symbol, clearly present in organisational semiotics, contributes to the explanation of the social and organisational processes and facilitates a deeper understanding of specific aspects related to organisational representation and behaviour.

Apart from the organisational aspect, cognitive science and artificial intelligence and, more specifically, cognitive architectures—the design and organisation of the mind—have been studied. Theories of cognitive architectures strive to provide a unified/generalised theory of cognitive systems, a description of the functions and its capacities and a set of principles for constructing a set of models, rather than a set of hypotheses to be empirically tested (Sloman, 1999). The two streams of architectures that can be distinguished are the symbolic and the connectionist stream of cognitive modelling. The first architecture is based on a production system[4], e.g. SOAR (Lehman, Laird, & Rosenbloom, 2006; Newell, 1990), and the second is based on (neural) networks, in which processing is highly distributed. In contrast to the symbolic architecture, there are no task related modules, discrete symbols, or explicit rules present that govern operations (Rumelhart, McClelland, & PDP Research Group, 1986).

According to Sloman (1996, 1999), hybrid architectures, e.g. ACT-R (Anderson, 1983; Anderson & Lebiere, 1998) and CLARION (Sun, 2003) that combine elements of both streams, are necessary, because symbolism tends to be capable of manipulating variables in a way that matches human competence. On the other hand, connectionism tends to be better in detecting similarities and contiguity[5] as well as people's ability to deal with and integrate many pieces of information simultaneously.

In this dissertation, a hybrid architecture is chosen for modelling the actor in order to create a more realistic basis for understanding the impact of individual behaviour in artificial societies. Apart from this division in cognitive architectures, there is a division in streams within AI. One is the classical approach in which the agent holds a symbol system or internal representation of the environment to which the agent applies rules, frames, semantic nets or logics, to reason about itself and its (social) position in the environment. The other approach is the situated, or behaviour-oriented approach (Brooks, 1991b; Steels, 1994). The situated or behaviour-oriented approach defines intelligence in the form of observed behaviour and without explicit representation of symbols and without abstract reasoning such as symbolic AI. The systems often have a set

[4]Production systems are computer languages that are widely employed for representing the processes that operate in models of cognitive systems (Newell & Simon, 1972; Newell, 1973).

[5]Contiguity refers to pattern recognition, i.e. the sequential occurrence or proximity of stimulus and response, causing their association in the mind (Pearsall, 2002).

of stimulus-response reactions, i.e. depending on the situation a proper action repertoire is chosen. Situated AI is closely connected to the environment and therefore is often referred to as embodied cognition[6]. However, the type of situatedness of embodied cognition is often *physical* and *not* social-cultural. The social constructivist theory fills in the void of limited social-cultural situatedness.

The social constructivist theory (cf. Blumer, 1969; Mead, 1934; Vygotsky, 1962, 1978) can provide the theoretical support for bridging the gap between the behavioural and the classical approach, and introduce social aspects to AI. The social constructivist theory explains how the individual, the mind and the self become part of the society through interaction with others and how the society becomes part of the self[7] (Mead, 1934). There is a growing interest in social situatedness and the social construction of reality (e.g. Brooks & Stein, 1994; Dautenhahn, 1995; Dautenhahn, Ogden, & Quick, 2002; Edmonds, 1998; Lindblom & Ziemke, 2003). Social factors in cognition are being introduced to models to advance their cognitive basis in relation to their social environment, e.g. CLARION (Sun, 2002, 2003).

In this dissertation we also apply theories of Multi-Agent Systems (MAS), which originates from the field of Distributed AI (DAI) (Bond & Gasser, 1988; O'Hare & Jennings, 1996). In DAI, most systems have a centralised control system, but with the birth of Internet and multi-processor platforms, researchers have started to experiment with decentralised control structures, in which every node has its own autonomy[8]. This notion of autonomy marked the start of the current commonly-used term agent or actor; an agent is situated in an environment and is capable of autonomous action in that environment (Wooldridge, 1999).

Multi-Agent Systems deal with autonomous agents that have control over their internal state and the way their actions influence the environment. In this dissertation, we are interested in the (socially constructed) interactions between agents and the organisational aspects as regards coordination and cooperation. The field of MAS also clarifies how models for investigating the behaviour of societies and organisations can be designed and built.

But before building a model, first an appropriate methodology needs to be selected. In the next section, we discuss the methodology that is applied in this dissertation.

1.3 Methodology

As mentioned before, the aim of this dissertation is to develop an instrument or simulation model. The development of our model requires a methodologi-

[6]Embodied Cognition is a growing research program in cognitive science that emphasises the formative role that the environment plays in the development of cognitive processes (Cowart, 2006).

[7]In the interaction with society, the individual internalizes not only the character of the individual but also the roles of all others with whom he has interaction, the so called 'generalized other' (Mead, 1934, p. 154).

[8]Autonomy is an agent's active use of its capabilities to pursue some goal without intervention by another agent in the decision making process that could possibly influence how that goal should be pursued (Barber & Martin, 1999).

cal approach that supports the development of an (simulation) instrument. In general, we can distinguish two approaches to conduct research: the *empirical cycle*—explaining phenomena by developing theory, and the *regulative cycle*—"solving real, existing problems in a well-defined, step-by-step manner using theoretical frameworks and models to derive a solution to the problem under scrutiny"(Homburg, 1999, pp. 22–23)[9].

The following distinction makes the difference and the resemblance between those two approaches clear (see Gilbert & Troitzsch, 1999). Statistical modelling conforms to the empirical cycle in which the researcher develops a model based on theory, a set of hypotheses or equations that results in a measurement model thereby creating an abstraction of the real world. The next step is to find out whether the model generates the predictions by comparing them to data that have been collected in practise (typically assessed by means of tests of statistical hypotheses). In comparison to statistical modelling, the logic of simulation as a method (see figure 1.1) is quite similar. However, in this case, the aim is to develop an instrument[10] whose methodology is similar to that of the regulative cycle.

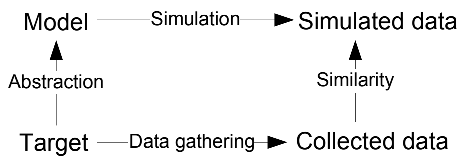

Figure 1.1: The logic of simulation as a method (Gilbert & Troitzsch, 1999).

First, the researcher creates a model of the target (the problem situation), but instead of using statistical equations, the modeller builds a computer program (instrument) that is able to simulate the target (e.g. a social or cognitive process). Next, the computer program generates an amount of data that can be compared to the collected data and thereby validates the model. The main difference between both approaches is that simulation models address the mechanisms that lie behind the processes that determine the final outcome, whereas statistical models "typically aim to explain correlations between variables measured at one single point in time" (Gilbert & Troitzsch, 1999, p. 17).

[9]The empirical cycle contains the following steps: (1) Observation—the exploration of phenomena, (2) Induction—design of an exploratory model, (3) Deduction—deduction of testable hypotheses, (4) Testing—confrontation of hypotheses with empirical data, and (5) Evaluation—rejection, revision or corroboration of the model. The regulative cycle contains the following steps: (1) Problem definition—perceived discrepancy between an actual and normative situation, (2) Diagnosis—formulation of the problem, (3) Plan—design of solutions, (4) Intervention—implementation of the solution proposed, and (5) Evaluation—test if the gap between actual and normative situation has narrowed (cf. Mietus, 1994; Homburg, 1999; Faber, 2006).

[10]Such an instrument can be applied in step 4-Testing-of the empirical cycle. In our case, the instrument serves as a tool that allows for generating data and comparing (testing) the simulation data with empirical data.

1.3.1 The design-based approach

Simulation models require a different approach than statistical models. In statistical models, the techniques applied for validating statistical models are well known by researchers in the field of social sciences and psychology and their corresponding tools are often supplied by software vendors (e.g. SPSS[11]) or open-source foundations (e.g. the R project[12]).

The techniques of computer simulation and more specifically the use of multi-agent simulations is not so widespread among researchers in the fields of psychology and social sciences. The use of programs to simulate outcomes of a model is known from the engineering disciplines, e.g. informatics, system engineering and artificial intelligence.

When creating statistical models, researchers apply off-the-shelf software that provides tools for analysing outcomes, whereas researchers that create simulation models develop their own programs (possible with help of existing tools) that require a different approach, i.e. it requires an engineering/design-based approach. The development of a program is similar to the engineering approach of constructing a building. First an architect creates a design (the design requirements), then an engineer creates the actual design, and finally, construction people build it. During the design process, from architect till construction worker, the necessary tests are carried out for verification and validation of the construction, resulting in iteration, redesign and a possible change of plans and design.

As a methodology to design and implement a Multi-Agent System, we have adopted the *design-based* approach (similar to the regulative cycle) (Beaudoin, 1994; Wright, Sloman, & Beaudoin, 1996), which specifies that:

1. The requirements of the system need to be specified;
2. Proposals of design need to satisfy those requirements;
3. Designs need to be implemented in a computer-simulation or in hardware;
4. The manner and the extent to which the design meets the requirements, and how the simulation embodies the design need to be analysed; and
5. The space of possible designs surrounding the proposed model should be studied and the design, which originates from an iterative process leading to the attainment of all its requirements, should be reconsidered.

The development of a model is described by Gilbert and Troitzsch (1999) as follows:

> One starts by identifying a 'puzzle', a question whose answer is not known and which it will be the aim of the research to resolve... This leads us to the [specific] *definition* of the target for modelling [stage 1]. [Next,] some *observations* [or collections of data] of

[11]http://www.spss.com
[12]http://www.r-project.org

> the target will be required in order to provide the parameters and initial conditions for our model. One can then make some *assumptions* and design the model [stage 2], probably in the form of a computer program [stage 3]. The simulation itself is performed by executing this program and the output of the simulation is recorded. [Next, w]e need to ensure that the model is correctly implemented and working as intended. This is *verification*—[a step in which the program is debugged—and]. . . *validation*[stage 4], ensuring that the behaviour of the model does correspond to the behaviour of the target. Finally, one needs to know how sensitive the model is to slight changes in the parameters and initial conditions: *sensitivity analysis*. (Gilbert & Troitzsch, 1999, pp. 17–18)

The development of a simulation model is not a strict linear process but rather a cyclical process whereby, after stage 4, stage 5 is entered and the development of the model often starts of again by adjusting the requirements or changing the computer model. This process is called the life cycle of a simulation study (Balci, 1998).

1.3.2 Life cycle of a simulation study

The life cycle of a simulation study exists out of several stages as pointed out in the previous section. This cycle is shown in figure 1.2 (cf. Bosman, 1977; Mitroff, Betz, Pondy, & Sagasti, 1974; Robinson, 2004; Balci, 1998).

The stages or phases in the model are comparable to the design stages discussed in the previous section. The *Real World* is the idea, situation and/or phenomena to be modelled; the stage in which the requirements are specified based on the real world. Then, the *Conceptual Model* is the mathematical/logical/verbal representation of the problem; proposals of the design are created that attempt to satisfy the requirements. The *Computer Model* (e.g. RBot, ACT-R or CLARION) is the implementation of the conceptual model as a program in a computer. The *Experimental Model* or models are specific implementations of the computer model that suits a particular study of the real problem, e.g. after the creation of a computer weather model, the experimental model of the weather in Groningen (Netherlands) is the specific application of the weather model to the situation in Groningen. The next stage is doing simulation runs and generating experimental data, and the interpretation, presentation and understanding of the data produced by the experimental model. The data can be compared to the real situation after which the model (and theory) can be redefined. The developmental cycle should not be interpreted as strictly sequential, i.e. the circle is iterative in nature and reverse transitions are expected (Balci, 1998).

Thus, besides the conceptual model, the following models are distinguished in figure 1.2: the experimental model—a model destined for a specific target and that only works specifically for that target—and the computer model—an (more) abstract model or toolkit that allows for the development of a range of specific experimental models that make use of the basic principles of the abstract

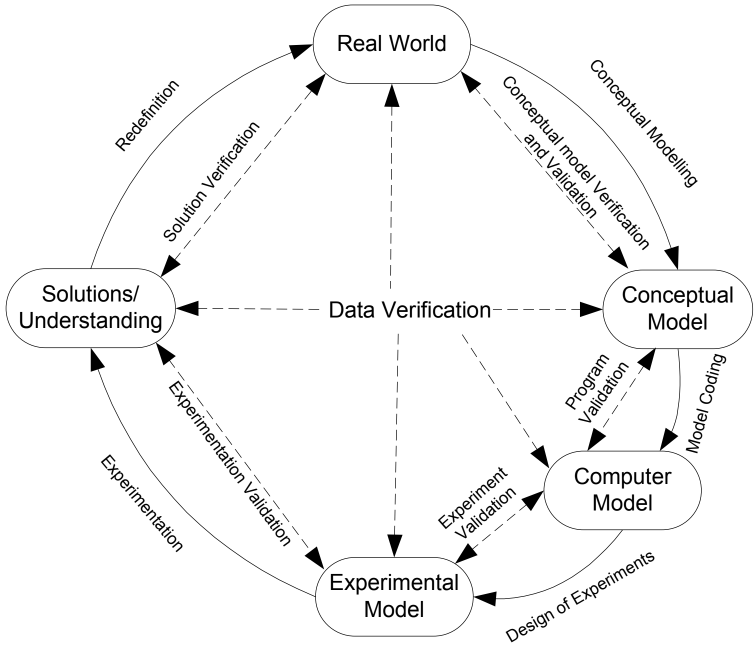

Figure 1.2: Life Cycle of a Simulation Study (adapted from Balci, 1998; Robinson, 2004).

model. In case of an abstract model (in our case RBot/MRS, see chapter 5), we follow a similar path in developing a model. However, the emphasis is not on a perfect fit with specific observed data, but more on the fit of mechanisms and phenomena that can possibly match with some real social phenomena.

Gilbert and Troitzsch (1999, p. 19) state that there is a continuum from detailed to abstract modelling and that abstract modelling[13] is research on 'artificial societies', i.e. simulation without reference to any specific 'real world' target. Or as stated by Epstein and Axtell (1996, p. 4):

> We view artificial societies as *laboratories*, where we attempt to 'grow' certain social structures in the computer—or *in silico*—the aim being to discover fundamental local or micro mechanisms that are sufficient to generate the macroscopic social structures and collective behaviours of interest.

In our demonstrative experiments (see chapter 6; experiment 2 and 3), we also create an artificial society and study social behaviour based on interaction between cognitive plausible actors (RBot), i.e. the actor's cognitive mechanisms are based on ACT-R, which is validated with help of empirical data. However,

[13]Abstract modelling in the sense that there is no concrete world to match the model with, but of course there are still concrete implementations of experimental models necessary to demonstrate the working of the abstract model.

the social phenomena that evolve from interactions between those actors are without reference to the real world.

The aim of our dissertation is to develop an abstract model or toolkit, whose purpose is threefold:

1. It can be applied for modelling and understanding of social and cognitive mechanisms that occur between actors and within actors, i.e. the simulation toolkit should allow for easy modelling (create experimental models) of interactive situations between actors delivering results that explain behaviour at the following levels: the social and interactive level between individuals, the level of the individual and the intra-individual (cognitive) level.

2. The second obvious step (and beyond the scope of this dissertation) is the modelling of specific situations that occur in reality, i.e. the gathering of data of interactions between organisations or people in the real world and comparing those with the simulation outcomes.

3. The third option is the application of the model in the virtual world of for instance e-business, gaming industry and many more situations. In these situations, the simulation program becomes an (pro) active agent that takes part in a virtual organisation or artificial society of interacting agents.

The design of our model aims to fulfil the first purpose and thereby supports other specific research as mentioned in point two and three, i.e. the design (RBot/MRS explained in chapter 5) is an abstract model or tool (purpose 1) that allows for the creation of many models that are suited to specific situations (purpose 2 and 3) to which those models should apply. The requirements of the abstract model are that they allow for the design of Multi-Agent Systems in order to explain interactive social behaviour between and based on *cognitive plausible actors*. The abstract model should give freedom for the creation of a range of diverse models—simple as well as complex models.

1.4 Research questions and objectives

Social (and organisational) behaviour is an outcome of the interactions between individuals. These interactions can lead to (temporarily) stable patterns of (organisational) behaviour. An *organisation* can be seen as a group of people that has habits of action aimed at cooperation and coordination of work. An organisation is not a physical, tangible object like an apple or a computer keyboard. Its observation, demarcation, and existence are dependent on the existence of human habits and human-produced signs; the organisation is a product of interactive social behaviour (Helmhout et al., 2004). An organisation can be traced back to representations (in the mind of actors) that structure the interaction among people, thereby demarcating a group of people who are members of an organisation (Van Heusden & Jorna, 2001).

As mentioned before, the aim of the dissertation is to design and implement a cognitive agent-based computational social simulation model to explain interactive (social) behaviour based on a selection of theories. This leads to the following research questions that are divided into (1) theoretical research questions and (2) design and implementation questions.

1 What are the aspects of actors that plausibly explain interactive (social) behaviour?

To answer this question, we need theories, models and requirements that address the individual as well as the interaction between individuals. The following three sub-questions will elaborate this:

1.1 What type of a model can explain interactive (social) behaviour?

Methodological individualism and social constructivism argue that social behaviour should be explained in terms of individuals, their properties and their interrelations in terms of these properties. Multi-Agent System as a model and methodology (see chapter 2) supports this view and can serve to explain interactive social behaviour; MAS is concerned with the study of behaviour, and the modelling of a collection of agents that interact with each other and their environment (Sycara, 1998).

1.2 What is required for an actor to exhibit (stable) social behaviour?

For an actor to exhibit social behaviour, it requires to create, interpret and exchange signs and meaning with other actors. Although theories of MAS address coordination, cooperation and negotiation issues, they focus only on social structures and overlook processes that describe how these structures emerge, stabilise and disappear. Whereas social constructivism and organisational semiotics (see chapter 3) consider that (stable) social behaviour requires (1) the creation of shared knowledge and the social construction of reality (cf. Berger & Luckmann, 1967), and (2) semiotic resources such as signs/symbols/social constructs that refer to social structures, institutions and habits of action.

1.3 What kind of an actor can plausibly handle signs, relations and social constructs?

An actor that handles signs, relations and social constructs needs to have a system that supports representations and mechanisms that manipulate these representations and exhibit intelligent action. More strongly, a general intelligent actor (system), whatever additional structures and processes it may have, will contain a physical symbol system (Newell & Simon, 1976; Newell, 1980). Therefore, the actor should be equipped with a cognitive architecture (see chapter 4). Such a cognitive plausible actor is able to generate and process symbols, create and 'understand' social constructs and build relations with other actors.

The theoretical research questions (1.1 until 1.3) have addressed the three main requirements necessary for the implementation of a cognitive agent-based social simulation tool:

- A simulated environment that provides a physical, communication and social environment for actors to live in.
- Social constructs, which are signs that serve as holders for sharing and reinforcing social norms, structures and institutions.
- A plausible cognitive agent that consists of a symbol system and cognitive/learning mechanisms that support the creation, manipulation, transfer of symbols/signs and is able to exhibit intelligent action.

Against this background, we can design and implement a multi-actor system. This task can be expressed in the form of the following research question.

2 How can (cognitive and social) plausible actors be implemented in a Multi-Agent System?

First, the cognitive actor, which is the main component of a MAS, is constructed. A production system is chosen from a selection of different approaches that are possible within AI (cf. Ferber, 1999, p. 125). It is the most elaborate choice that matches the necessary requirements for an actor to be a rational problem solver. A realistic computerised model of a cognitive plausible actor could be created with two cognitive architectures, either SOAR (Lehman et al., 2006; Newell, 1990) or ACT-R (Anderson & Lebiere, 1998). ACT-R was chosen here as the architecture to model the individual actor (see chapter 4 for a discussion). In its most recent release, ACT-R is an architecture of a single agent for which no multi-agent architecture version is yet available. Therefore, the ACT-R architecture is transformed and extended into a new architecture, named RBot. Together with a newly created MAS environment (Multi-RBot System or MRS), it is possible to incorporate multiple actors in a task environment.

In order to test the validity of a cognitive actor (RBot) as part of the new model (MRS), a comparison[14] with ACT-R is made in this dissertation. Consequently, the accompanying question is:

2.1 Is RBot comparable to ACT-R as a valid cognitive plausible actor?

After modelling the single actor and the platform on which to run multiple actors, a test is performed to see whether social constructs can emerge from the interaction of actors. The accompanying task environment is based on a well known experiment of *traffic laws* (Shoham & Tennenholtz, 1992a, 1995). In our task environment, actors face each other and need to pass each other to make progress. They have two choices. They can either drive left or drive right. Collision detection allows actors to react to each other's behaviour and, if necessary, to correct their behaviour.

Computational & Mathematical Organization Theory serves as an enlightening example of how organisational and social aspects (e.g. the traffic laws) can

[14] Multiple experiments were conducted to compare RBot with ACT-R. One of these experiments is demonstrated in this dissertation (see experiment 1, chapter 6).

be modelled with the help of MAS. CMOT is a multi-level system that is descriptive at the individual level—cognitive plausible actors as boundedly rational individuals (March & Simon, 1958)—and normative at the second level, i.e. the rules of the game such as defining different roles and authority structures. However, models of CMOT such as Radar-Soar (Carley et al., 1998) have the drawback that the macro-features of the model are pre-programmed (off-line) and do not emerge during interactions. The approach in this dissertation is to see whether behaviours, i.e. social constructs, can emerge from the interactions of agents, without the need to pre-program behaviour. An experiment is designed based on the proposed task environment to demonstrate the emergence of social constructs. The question that the experiment tries to answer is:

2.2 Is it possible that social constructs and organisational behaviour can emerge from social interaction?

In many situations, the interaction between agents is guided by social constructs that are formed as a habit between actors. The society is filled with rules (March, Schulz, & Zhou, 2000) that emerge from interactions, take on an implicit or explicit form and survive until they are no longer exercised. Coordination and cooperation is only possible when agents that interact with each other are aware of each other's social needs. In other words, the actor should have not only a representation of itself and the environment, but also that of (i.e. the needs of) the others; the society as a whole. The model of the cognitive agent is enhanced by adding a normative level of social constructs as representations to the actor rather than to the overall system. The question that the last demonstrative experiment tries to answer is the following:

2.3 Is a social construct a coordination mechanism that can influence the behaviour of interacting related actors towards a certain desired behaviour?

The objective of the design and implementation questions (2.1 until 2.3) is to address the mechanisms inside the actor that are influenced by interactions with other actors and thereby socially constructed. The underlying assumption is that the actor is socially connected to others on a continual basis. The second objective is that these questions are simultaneously design questions. First, the requirements are drawn up. Second, the design or model is constructed. This is followed by an implementation of the system, and fourth, demonstrative experiments show its working. The design questions guide the construction of a model that attempts to bridge the gap between cognitive and social sciences (cf. Sun, 2006a).

1.5 Dissertation outline

The dissertation is structured in the following way (see figure 1.3). The introduction, i.e. the current chapter, explains the theoretical disciplines that are used and how these disciplines are applied to the research, discuss the methodology, followed by the research questions, objectives and the dissertation outline.

Chapter 2 commences with a general introduction to Multi-Agent Systems. The chapter explains the notion of agency regarding different types of agents and how these agents interact with each other in an environment. It gives the reader an overview of the literature in the area of MAS.

Chapter 3 introduces the social actor and explains several theories including the social constructivist theory and organisational semiotics that bridge the gap between social (behavioural) and cognitive science. The theory describes how actors can be embedded in organisations and society. Secondly, based on these theories, a model of the social actor is presented, i.e. a model that is reactive, pro-active and socially and physically situated in the environment.

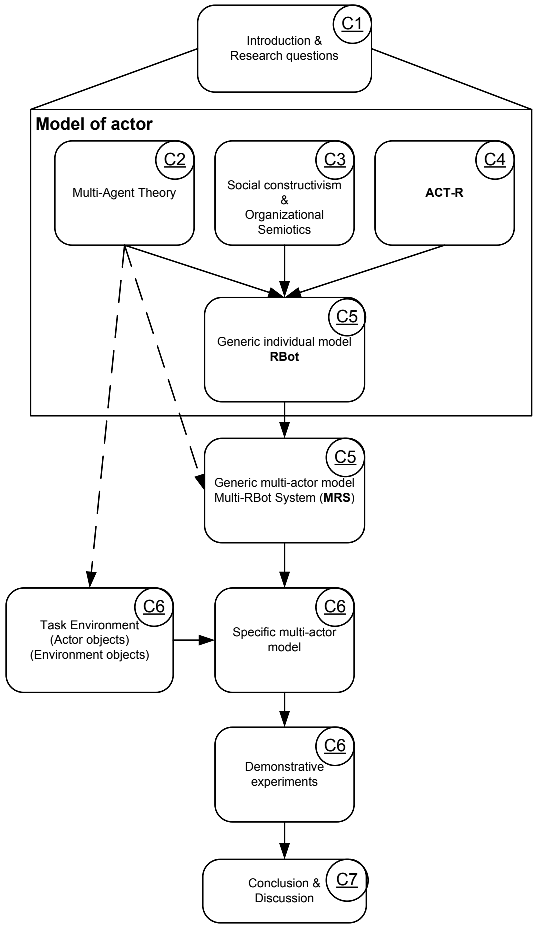

Figure 1.3: Structure of the dissertation.

In chapter 4, the selection of a cognitive architecture is discussed. The cognitive architecture ACT-R is selected as the architecture for constructing a computer model (RBot) of a cognitive plausible actor. ACT-R is discussed in detail, because it is the main contributor to the new RBot model. The subsumption architecture (Brooks, 1986) is drawn upon as a supporting theory to implement social behaviour at a normative level.

In chapter 5, the design chapter, we will discuss (1) the architecture of the social and cognitive actor (RBot), and (2) the multi-agent environment (MRS). First, the actor is modelled in a bottom-up approach starting with the memory model, then the controller and handlers, followed by the implementation of social constructs. Second, the server (as part of MRS) is modelled in which a controller regulates (1) registration of actors, (2) the time synchronisation between actors happens, (3) the communication between actors with the help of a message-based architecture is regulated, and (4) the physical, communication and social environment is maintained.

In chapter 6, the three experiments are discussed. The first experiment focuses on the functioning of the single actor. The experiment shows that RBot processes 'symbols in the mind' similar to the way ACT-R treats those symbols. This is done by demonstrating a multi-column addition problem in which RBot behaves similar to ACT-R in solving this problem, i.e. with help of a goal stack, a set of procedures and a declarative memory.

The other two experiments demonstrate the working of RBot as well as its benefits in explaining the emergence and influence of social constructs as a coordination mechanism. The first one of these two experiments is a traffic situation in which two actors confront and try to pass each other. Both actors have a similar set of procedures. However, they have a choice of selecting either the left or the right side to pass the oncoming driver/actor. The interaction between the actors and the individual experience of success and failure in passing the other emerge into a stable behaviour over time. The assumption is that both actors/drivers build up a preference, or come to a tacit mutual agreement that is in favour of both actors. The other, and the last, experiment is an extension of the previous traffic situation. However, in this experiment, one of the actors acts as an authority (a policeman assigned by society) to communicate laws or social constructs to other actors. Compared to the previous experiment, the authoritative actor in the last experiment is equipped with a mechanism that supports it in conveying the 'rules of the game' to influence the behaviour of others when rules are not abided by.

Finally, chapter 7 initially presents the contributions of the research and it draws general conclusions bearing in mind the research questions addressed in this chapter. Second, it advocates the importance of cognitive actors in increasing plausibility in explaining social behaviour and the empirical testing of cognitive models to enrich social simulations. Third, new questions are raised and last, the chapter highlights avenues for further research.