Control over Emergence

Martin Helmhout[1] and **Henk Gazendam**[2] and **René Jorna**[3]

Abstract. This paper explains and demonstrates emergence of organisational behaviour as a social cognitive mechanism, i.e. ones own behaviour at the cognitive level is influenced by interaction with others at the social level.

Besides the importance of understanding how behaviour evolves, it is probably more crucial to control emergence or enforce desired behaviour. In our research we demonstrate this by implementing social constructs as regulators or stimuli of behaviour.

The paper discusses a social cognitive architecture ACT-RBot or in short RBot[4] which is based on ACT-R. RBot inherits the cognitive architecture (production system) of ACT-R but provides also a mechanism of social constructs as meta-productions that operates as a social (control) layer. The architecture is implemented in software agents who 'live' in a discrete event simulation environment allowing them to interact and exchange signs.

The combination of RBot and a simulation environment provides observation of behaviour between agents (at the social level), but also introspection of the experiences of the individual agent stored in memory of its cognitive architecture.

We describe two simulation experiments that demonstrate the working of the social cognitive architecture. The first experiment shows that emergence is present at the cognitive (individual) and social level (interaction). The second experiment adds social constructs and authority that allows for (more) control over emergence.

1 INTRODUCTION

The problem with emergence is that it is often difficult to find out what caused it to appear. Secondly, when finding out, it is probably even more difficult to control, coordinate (or prevent) emergence within predefined boundaries.

For the social scientist, emergence is often described at the social level or level of interaction. The cognitive scientist on the other hand tends to focus on the way patterns in the mind of the individual emerge. According to Castelfranchi [11], 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. Cognitive science–in particular Artificial Intelligence–can contribute to the explanation of social phenomena. Likewise, cognitive science needs social science in order to incorporate social factors, i.e. there is a shortage of good ideas and theories that address socio-cultural concepts/signs/symbols with social structures from a cognitive standpoint [44].

In this paper we describe how to connect both sciences (micro-macro link [1]) and in particular explain cognitive emergence caused by social emergence and the other way around. We will also make a distinction (see [24]) between uncontrolled (first order) or spontaneous social emergence and controlled (social) emergence. During the start of the former type of emergence agents are first self-aware and strive to the outcome of their own actions in their best self-interest, but in the later state of emergence, conditioning (unconscious learning [3, p. 95]) takes place and the agents behaviour becomes more automatic. In the case of controlled social emergence, an agent requires social capabilities as well (e.g. social contructs, language,...), i.e. the agent needs to perceive (receive) and produce social constructs or activate them in the presence of relevant stimuli. The social construct serves as a moderator that influences the behaviour of the agent indirectly.

The structure of the paper is as follows. In section 2 we elaborate about social constructivism and social constructs. Section 3 discusses the social cognitive architecture RBot and section 4 describes experiments that demonstrate emergence as well as control over emergence. Finally, section 5 ends the paper with a short discussion.

2 SOCIAL CONSTRUCTIVISM

We argue that first order emergence, which does not require social skills, requires cognitive skills and some form of self-awareness[5]. Although such (economic) agents can be very successful, they are unable to comprehend the needs of others and therefore often strive to local optima instead of the optimal solution for a group as a whole. Economic agents equipped with sophisticated utility functions do not describe any actual economic or any other behaviour of any individual or group of individuals. "...economic agents are not socially embedded in the sense that the behaviour of no individual is influenced by interaction with any other individual" [32, p. 394].

Socially embedded agents are agents who are capable of expressing social behaviour and are aware that they are not alone, but are also part of a group or society. In other words, the agent does not only live in his own world (internal representation) but also builds up relations with the outside world; with physical, social and cultural objects, agents and groups of agents.

Social constructivism as a social psychological theory attempts to explain the relation between society and the individual who is part of that society. Mead [31] analysed the human being as a complex individual who socially constructs the world, itself and other human individuals; social construction of the world is created by a process of interaction. Hacking defines social construction as follows:

> X[, as a social construct,] need not have existed, or need not be at all as it is. X, or X as it is at present, is not determined by the nature of things; it is not inevitable.... X was brought

[1] ACIS, Guildford, United Kingdom, email: martin@acis.nl
[2] University of Groningen, The Netherlands
[3] University of Groningen, The Netherlands
[4] http://act-rbot.sourceforge.net

[5] We assume that the agent is cognitive plausible, has memory and therefore is able to reflect on its own past actions.

> into existence or shaped by social events, forces, history, all of which could well have been different. [26, pp. 6-7]

Hence, a social construction or social construct can be seen as an invention or artefact (cf. [25, 39]) constructed by interaction between members of a social group or interaction between groups. Products of social construction, such as institutions, gender and emotions are social constructs, created, disseminated, and agreed upon by social groups [40, p. 522].

2.1 Affordances, signs & social constructs

Organisational semiotics [41, 42] suggests to combine *affordances* and *signs* to bridge the gap between the social and the individual level of the agent.

> Affordances stress the interaction between a human agent and its environment based on behaviour patterns that have evolved over time in a community. Signs stress the social construction of knowledge expressed in sign structures... Stamper sees affordances as repertoires of behaviours and distinguishes physical affordances and social affordances. [22, pp. 7-8]

A (physical) *affordance* is a set of properties of the environment that makes possible or inhibits activity [23]. After many encounters with the environment, this can result in a *habit of action*, which is a commitment to act with a connected action program that governs the actual acting [19, 36]. From a semiotic point of view, one could say that a physical affordance becomes a *social affordance* as well, the moment the physical affordance is shared between agents in a community. The experience of the object (shared with others) is built up in the mind of the agent; the agent is socially situated through interaction and perception, which is a process of social construction of signs in the agent's mind. The resulting signs are organised as units of knowledge consisting of a representation of an affordance and its associated habit of action.

Social constructs are social affordances [28, 41] and can be seen as representations of cooperation and coordination, based on intertwined habits and mutual commitments that are often expressed in sign structures such as agreements, contracts and plans. A social construct [20, 28] is a relatively persistent socially shared unit of knowledge, reinforced in its existence by its frequent use. In organisations, social constructs take the form of, for instance shared stories, shared institutions (behaviour rule systems), shared designs, shared plans, and shared artefacts. These social constructs support habits of action aimed at cooperation and coordinated behaviour.

In order to use social constructs in formal simulation models, we have defined a (not limitied) set of properties of a social construct [27].

- *Attached norms or rules*: social constructs can contain (a collection of) norms or rules that guide action and prescribe appropriate behaviour in a certain context. Our daily encounters with social norms (and law) are evident, for instance, when we are driving with our car on the right side of the street, or being polite for letting the elderly sit in the bus, etc.
- *Written/unwritten (coded/sensory)*: a social construct can be formed and communicated by writing the attached rules and norms down on paper, or they are internalised in agents and with the help of interaction transferred (language or gestures) to others [29].
- *Life span*: every social construct has a starting time, an evolution, a monitoring and controlling period and a finishing time [28]. The life span of every social construct, be it a norm, an emotion or an organisation varies and depends on other properties connected to the social construct, e.g. referred objects or facts in the social construct change over time, a lack of reinforcement of the social construct or changes in enforcement costs (people enforce others to obey the norms attached to the social construct).
- *Roles and identification*: the agent is given a role or identification, e.g. employer, employee, to make clear the authority, control and rules applied to that role [28].
- *Authority, responsibility and control*: according to Fayol [16], authority can be seen as 'the right to give orders' and the expectation that they are followed. Control and power can assure that agents behave responsible; they can be part of a directly involved authoritarian party or an independent third party. Assigning authority to someone creates a responsibility for that person to give orders and control whether other agents take their responsibility in following the 'rules of the game'.
- *Inheritance or prerequisite of other social constructs*: a social construct can become part of a complex network of connections with other constructs (that are often the result of earlier agreements). For example, when preparing a sales contract, sales men refer to their conditions that are registered at the institution of commerce and the registered conditions inherit conditions from public law
- *Scenario*: there can be a more or less standardised process (scenario / script [38]) for establishing a social construct between agents. Scenarios are often put on paper in which a specific order of social constructs over time is written down. In communities, scenarios are often informal and expressed in rituals and transferred from generation to generation.
- *Context*: context as property is a debatable issue, however there are two interpretations of context. Context can be situated *outside* the agent, and—possible at the same time—situated *inside* the agent, i.e. context is represented as symbols in the 'real' world stored externally from the mind and also as symbols stored in the mind of the agent. The external context contains certain elements—so-called affordances—perceived by the agent to which it is sensitive or is triggered by. In contrast to Gibson [23], and similar to Stamper [41, 43], Vera and Simon [48, p. 41] state that affordances "are carefully and simply encoded internal representations of complex configurations of external objects, the encodings capturing the functional significance of the objects". We assume that there has to be sufficient coherence between an internally represented social construct and an element in the external environment in order to activate or trigger the social construct. According to Gazendam, Jorna, and Helmhout [21]:

  > The recognition of a situation in which a [social construct] must become active must not only depend on the recognition of physical affordances like other agents, objects and situations, but also on an ongoing monitoring in the agent's mind of the state of the social context in terms of invisible entities like positions of rights and obligations of himself and of other agents, agreements and appointments that must be held, and so on. (p. 2)

Social constructs with all their properties are internalised and established by a process of socialisation and communication. During their life-span, they are monitored by enacting agents through observation, mentoring, practise, and training [29]. Secondly, they create standards of appropriate behaviour and stabilisation, i.e. they create shared expectations of behaviour and when exercised, they are evaluated by others as well. Thirdly, when they are widely known and ac-

cepted as legitimate, they are often self-enforcing, and its associated psychological cost of rejection will rise, e.g. the individual or group involved can feel psychological discomfort whether or not others detect the rejection [30]. And fourthly, because social constructs are connected to roles and authority, they can create formal and informal social structures. Such a social structure exists out of a network or collection of social constructs. When a social structure is legitimate, legally acknowledged and officially registered, it is often referred to as an organisation or institution.

2.1.1 Emergence of social constructs

The general processes that underlie social construction are the creation of new social constructs, the evolution, the control and ending of those constructs over time. Emergence of social constructs is mainly concerned with creation and evolution and with a lesser extent with control and ending of emergence. For a first order of emergence this is not a problem because it concerns mainly creation and evolution. The agents lack a sense of social affordance; they are insensitive to social control and sometimes even worse, they cannot be stopped. On the other hand, second order emergence cannot take place without social construction. We argue that agents need to be self and socially aware in order to control or even stop emergence.

We distinguish three phases during the life span of a social construct: creation, evolution and ending of the social construct. The control function is a separate process that can be applied during the entire life-span, i.e. during creation, evolution and ending of a social construct.

The process of *creating* a social construct can occur in three ways [17]:

1. Based on adaptive emerging social behaviour; when two agents have a conflicting or mutual goal / interest. For instance people going out of the elevator while others only can go in when it is empty, or a group meeting in which people after a couple of times play their (habitual) role.
   (tacit agreement in a community)
2. The other is communicative action or an authoritative ritual in which with help of communication, e.g. speech acts, a social construct is formed; it starts with the agent that wants to propose, inform, or is requested to inform or propose a social construct to another agent.
   (agreement in a community)
3. The third formation process is external norm formation and a clear specification of roles, in which certain individuals are authorised to prescribe and enforce norms, thereby regulating the individuals they supervise [49].
   (creation on authority)

The first creation process is an implicit social exchange of signs and agents play their (habitual) role without being aware of creating order and a new social construct. The last two creation processes are explicit. The agents communicate or bargain openly about what they want. During the *evolution* of a social construct, agents practise the social construct and learn by interaction with the physical, social and cultural environment: a social construct is adopted, reinforced, adapted or rejected. Normally, the *ending* of a social construct is reached when the end of the (agreed) life span or situation is reached. However, during the process of evolving, the social construct is subject to changes; agents re-negotiate, renew the social construct or simply end their commitments. After the ending of a social construct, (cognitive) agents do not forget the success or failure or contents of a social construct; the agent can reuse the experience with a social construct as prior knowledge in a new negotiation process. Ending of a social construct can also happen when it is not frequently used, i.e. the agent does forget how to apply the social construct in a specific context and therefore cannot commit herself anymore. Therefore, a large amount of social constructs are written down and exist in documents (e.g. law). For those reasons, many people make use of experienced agents (brokers, attorney) who have enough skills concerning social constructs and domain-specific knowledge.

Thus, the processes of creation, evolution and ending depend on the support for the social construct [7], such as the dynamics of the environment (e.g. entering and exiting agents), regime change and the control process that monitors these processes.

2.1.2 Control over emergence

Control processes determine the actual creation, evolution and ending of a social construct. During creation there is somehow a form of control, be it formal or informal, e.g. there is a formal set of rules actively enforced by an institution, an informal ritual process is known by the agents or agents are influenced by social constructs from their personal social and cultural historical background. Control is necessary in order to prevent agents from breaching agreements, become corrupt or even become a threat to society.

Agents feel the need to compare and find out if others are still aware of the conditions and associated norms of the social construct to which they all mutually committed. Conte and Castelfranchi [13] state that in a group of addressees to which a norm applies, the norm should be respected by all its addressees; it is a necessary consequence of the *normative equity principle* defining that agents want their 'normative costs' to be no higher than those of other agents subject to the same norms. Agents stick to a norm as long as the advantages of following up a norm outweigh the advantages of not following the norm, i.e. the agent has to comply with the norm and if necessary defend the norm or give up the norm (when the costs of obedience are too high compared to punishment).

The properties of social constructs are determined during the process of creation and evolution (emergence of social constructs). During that process agents desire to reach an outcome that is beneficial for the agent itself and/or for the community as a whole. Therefore, agents collect information about other's preferences concerning social constructs; especially when dealing with unexplored territory. Berger and Calabrese [9] state: "... [] when strangers meet, their primary concern is one of uncertainty reduction or increasing predictability about the behaviour of both themselves and others in the interaction (p. 100)". Hence, the desire of every agent is to get some form of control or get information about who is in control (authority) and what the 'rules of the game' are. Therefore, the agent (if he desires) has to spend time and energy (cognitive load) in finding out the required knowledge about how other agents behave and interact with each other (code of conduct). Often many struggles concerning power, authority and trust relationships first need to evolve before (tacit) agreements have been reached.[6]

Social constructs that have evolved and are considered stable, still require a certain amount of investment to be maintained in a community; control of emergence does not only concern the creation but the maintenance of the desired social construct as well.

[6] In this paper we do not elaborate about power and trust issues, but they certainly require attention because they are closely attached to (control over) emergence of social constructs

Sometimes it becomes rather difficult to change certain systems. Consider the following radical change of a social construct; starting now, everyone in the world should drive on the left side of the road which increases standardisation in car production in favour of a 'greener' environment. Such a change concerned today is impossible due to the huge amount of switching costs (and sunk cost). Besides that, politics, power and the feeling of uncertainty will keep the original social construct in place. Strong habits will make sure that there will not emerge a new system replacing the old one.

Summarised, control over emergence is desired when social constructs are created the first time introduced and when social constructs need to be maintained over a certain life-span. In section 4, we will demonstrate control over emergence in a simulation experiment.

3 Social Cognitive Architecture

The previous section elaborated that social constructivism enables the agent to become aware of the world around him and how this world can be represented in the mind of the agent. We argue that a social agent should have a cognitive architecture, i.e. be cognitive plausible in order to handle *representations* such as signs, language, relations and social constructs. A cognitive plausible agent (its architecture) is not only based on a physical symbol system [33], cognitive mechanisms, goal-directed behaviour and learning capabilities, but is empirically tested as well [34, 2, 47]:

> [A *cognitive plausible agent*] is a goal-directed decision-maker who perceives, learns, communicates, and takes action in pursuit of its goals, based upon [a physical symbol system] that implements theories of cognition [and is supported by empirical evidence]. [47, p. 88]

In our research we have adopted existing theories of cognition about cognitive architectures and compared the two dominant architectures–SOAR [34] and ACT-R [5]–and decided to adopt the theory ACT-R.[7] ACT-R can be described by three or four levels of description (see [14, 15]) to predict the behaviour of the overall system, i.e. the implementation level as an approximation of the physical level (the sub-symbolic level), the functional level–the functionality of procedures interacting with declarative chunks, and the intentional level–the beliefs, goals, desires and intentions.

Most cognitive architectures focus on the individual and more specifically on the cognitive band ($10^{-1} - 10^{1}$ sec) and the rational band ($10^{2} - 10^{4}$ sec) but *not* much on the social band ($10^{5} - 10^{7}$ sec) [34, p. 122]. ACT-R, as thoroughly empirically tested architecture, also does not pay much attention to (social) interaction between individuals.

Therefore, a new social cognitive architecture (RBot) was designed completely from scratch. RBot follows closely the cognitive theory of ACT-R, but some changes/adjustments were necessary to make the cognitive agent social:

- *Multi-Agent System:* social cognitive agent based simulations require artificial task environments in order to construct various interaction scenarios. We implemented a (environment) server that has the purpose of providing an artificial 'physical' task *and* communication environment. It allows agents to perceive objects and other agents, and to exchange signals and signs. Secondly, an Agent Communication Language (ACL) is implemented that enables ACT-R agents to communicate with each other.
- *Social interaction:* for a cognitive agent to become social, it needs to create and maintain shared knowledge, be able to socially construct its environment and represent social structures (institutions) and habits of action in its mind. In other words, the cognitive agent needs to understand and be able to express itself at the social level (band) as well.
  At the cognitive band, productions fire on average at every 100 ms (see [5]) and the rational analysis takes place at the rational band. Apparently there is a gap ($10^{1} - 10^{6}$ sec) between the social band and rational/cognitive band that needs to be bridged. In order to bridge this gap, we have defined the social construct as a (social) representation in the mind of the agent, or as documents or artefacts in society. The social construct allows us to define social situations in which the agent has to obey or is permitted to follow certain rules that are present in a society or culture. We have extended ACT-R with a special module (see figure 1) containing social constructs (as chunks) that influence certain aspects of behaviour.
  In general, social constructs can be formed in many ways; created by internal explicit or implicit cognition, or perceived from the external environment. Once created, a relatively simple mechanism can describe how social constructs operate. A social construct can have norms attached to it that can respond to changes of social representations in memory that reflect changes in social situations in the environment. The *condition side* of a norm is triggered by a combination of general (social) concepts or more specific instances of concepts. The *action side* has one or more targets (procedures/goals/other norms/etcetera) connected to it. As soon as the norm is triggered, it influences those target chunks by demoting or promoting their activation levels thereby indirectly changing the response functions and the behaviour of the agent.
  A social construct is a social chunk (concept) of cognition that can be linked to other social constructs creating a complex semantic network of constructs and norms. The mechanism or architecture of social constructs resembles the subsumption architecture of Brooks [10]; it is a multi-level system in which the upper (normative) level is a (network of) social construct(s) that influences behaviour at the lower level(s)[8].

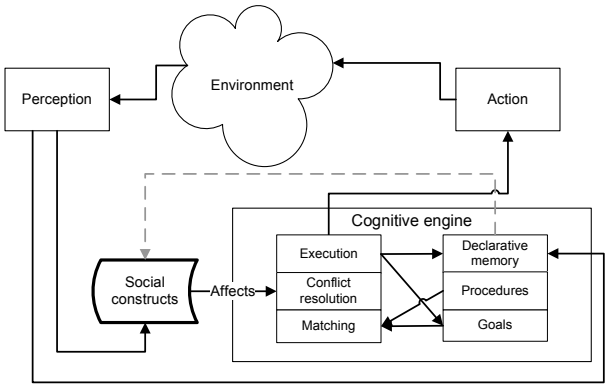

Figure 1. The extension of the cognitive engine with social constructs

- *Adjustment of rational and functional level:* we transformed ACT-R in order to be able to function at the social band, e.g. a slower de-

[7] We will not go into much detail about cognitive architectures here. For in-depth discussion, we refer to chapter 4 of 'The Social Cognitive Actor' [27]

[8] We actually made an attempt to combine ACT-R (hybrid agent: connectionism & symbolism (GOFAI)) with embodied cognition ('New AI')

cay of memory [46, p. 217]. Similar adjustments have been made in changing effort parameters of productions.

Much research still has to be done to make the architecture match its social requirements, however the current proposed mechanism will be a good start in making a cognitive architecture social. We will leave further implementation details behind us and start discuss the necessary parts of the social cognitive architecture that help to clarify the experiments demonstrated in section 4.

3.1 Base-level activation/decay & event discounting

Base level activation is probably *the* most important feature of ACT-R, i.e. it has been used in many environmental experiments [6] and has been the most successfully and frequently used part of the ACT-R theory [4]. Base-level activation B_i is an estimation of the log odds (of all presentations of a chunk in history) that a chunk will be used and is defined as:

$$B_i = \ln\left(\sum_{j=1}^{n} t_j^{-d}\right) + \beta \text{ Base-level Learning Equation}^{9} \quad (1)$$

- t_j represents the time-difference (t_{now} - $t_{presentation}$) that the chunk was represented in memory (created or retrieved),
- n the number of times a chunk is retrieved,
- d the decay rate,
- β the initial activation upon creation of the chunk.

The equation suggests that the more often a chunk is retrieved from memory (high-frequency), the more its base-level activation rises. On the other hand, the activation level of a chunk that is not retrieved at all can drop below an activation threshold level, whereby it becomes almost impossible to retrieve the chunk. Figure 2 exemplifies the retrieval of a chunk at t = 10 and t = 28. It shows a small (average) increase in activation over the specified period. As

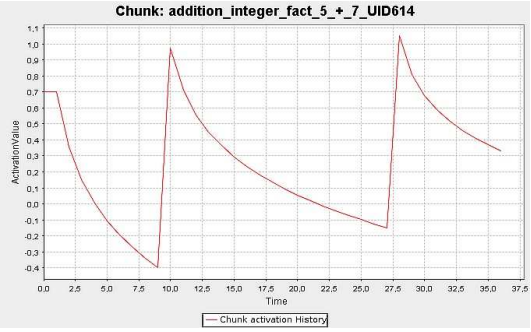

Figure 2. Example chunk, accessed at times 10 and 28.

regards the base-level learning equation, Anderson & Lebiere [5, p. 124] state that "odds of recall... should be power functions of delay, which is the common empirical result known as the Power Law of Forgetting [37]. Although it is less obvious, the Base-Level Learning Equation also predicts the Power Law of Learning [35]". Hence, the base-level learning equation enables the agent to learn and prefer chunks that are often needed for solving problems, and to neglect or forget chunks that are almost never needed.

[9] [5, p. 124]

For an agent to be socially situated, it needs to be able to learn facts and rules that emerge, but also to forget those when they are not reinforced or are replaced by other rules and facts in a community.

Decay functions are not only applied for chunks but also at the sub-symbolic level of procedures for events (success and failure). ACT-R has a mechanism that takes care of discounting past experiences by implementing an exponentially decaying function that is similar to the base-level learning equation. The equation for discounting successes and failures is:

$$Successes, Failures = \sum_{j=1}^{m,n} t_j^{-d} \text{ Success Discounting Equation}^{10} \quad (2)$$

- m number of successes,
- n number of failures,
- t_j time difference, now - occurrence_time of the success or failure,
- d decay rate.

The effect of the equation on the selection of productions is that there are more exploration possibilities caused by the decay of past successes and failures. For example, it is possible to give different decay rates for successes and failures, i.e. when the decay rate of failures is lower than that of successes, ACT-R tends to forget negative experiences more slowly than positive experiences.

Social constructs are reinforced by their frequent use. Hence, the abstract chunk with its base-level activation is an appropriate container for storing a social construct. The extension with the mechanism of the subsumption architecture combined with ACT-R triggers the social construct to become active when changes are taking place in certain areas of the memory (e.g. perception or communication).

The new model resulted in a social cognitive architecture called (ACT-)RBot which is part of an agent that operates in a Multi-Agent System giving space and time awareness to the agents and the ability to perceive and communicate with each other. In the following section we applied this architecture as the basis for our experiments.

4 Experiments

The purpose of the experiments is to show, as simple as possible, RBot operating in a multi-agent environment. Thereby demonstrating and finding out whether RBot agents can 'live' in a task environment and learn from interaction by cooperation, observation or other means. We have modelled a multi-agent environment of two agents in order to study the emergent behaviour of interaction between two individuals.

The general experiment is a case in which two agents have to pass each other as in a traffic experiment. They have to find a way out to pass each other several times without causing an accident. In the first experiment, there is no fixed traffic rule to drive on the left or the right side. In the second experiment, one of the agents acts as a controller (policeman) by enforcing the other to obey to the traffic rule he was given by higher authority. Hence, the policeman is present to control the way the social construct emerges in the mind of the other agent (given that the other agent respects the authority of the policeman) in case that agent wants to deviate from the common norm.

The basis configuration of both agents is equal; they are given the same parameters, procedures and declarative chunks, equal motivation values to solve goals and equal noise distribution functions. The

[10] [5, p. 141];[5, p. 265]

agents are identical in that sense that the simulation outcome is based on interaction and not on differences at the cognitive level of the agent. The agents learn in the following way; they evade right or left and then see if the evasion was successful or not. Based on this experience, the agent updates its productions and uses this experience in the next encounter.

The experiment showed two different emerging patterns. The first is an immediate lock-in; agents both initially choose the same strategy (e.g. driving on the right side), pass successfully and start to reinforce that particular strategy and leave no opportunity for the other strategy (driving on the left side) to emerge. The second pattern is caused when both agents at first do not select the same strategy, i.e. the agents are on a collision course. The second pattern is the most interesting, because agents have to decide based on experience which strategy is acceptable for both agents. By initially selecting the colliding strategy, the agents show a shifting behaviour from left to right and from right to left. After some collisions, the Boltzmann factor (noise) in the utility function gives the agent freedom of escaping from the hopping behaviour. This gives agents the opportunity to settle into the same successful strategy, both passing either left or right.

The interaction between two agents gives enough material to study phenomena and theories of different fields (e.g. management and organisation, social science, cognitive science, AI, to name a few). For analysis of organisational behaviour, we describe phenomena mainly at the individual level of description as a social construct (the behaviour of the individual agent), but we also describe behaviour at the social level by observing the interaction patterns between agents that emerge during the experiments. Hence, the experiments show behaviour of the individual and its internal cognitive properties and the emergence of behaviour of the collective as well (the behaviour caused by interacting or communicating (social constructs) between individuals).

4.1 Experiment 1: emergence of social constructs

This experiment demonstrates a first order of emergence in which we will explain how a (internalised) social construct emerges and evolves over time. The experiment will highlight that the agents seem to have a 'social agreement', but they actually adapt their own behaviour to the environment including the other agent, i.e. they do not intentionally influence or change each other's behaviour to come to an agreement. Because of its mutual adaptation, a social construct or norm emerges and creates regular stable behaviour in an otherwise chaotic system.

The purpose of the experiment is to show that a social construct can emerge not only by looking at the agents' collective outcome or rational level in the way they socially behave, but also what happens at the cognitive level and especially the sub-symbolic level of each agent.

In order to understand the behaviour of the agents, we want to look for behavioural patterns at two levels, one for the behaviour of the agent that is reflected by its internal history of memory and the other in the role of the external observer that spots the agents having a collision and ending up in an emerged pattern. With help of those patterns, the behaviour of escaping and reinforcement can be explained.

Figure 3 shows the results of the experiment (from an external observers point of view) that started with a collision. First, no particular strategy is chosen; the right move as well as the left move is equally preferred. However, approximately after time step 700, the agents prefer the 'Left' strategy based on interaction and experience built up in memory over time. Due to this interaction, they develop a preference in favour of the left move so strongly that the agents only choose the left strategy, i.e. the utility difference between right and left becomes significantly large that the agents are locked-in into the left passing strategy.

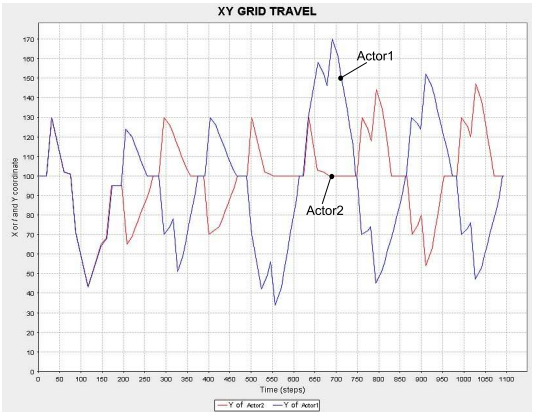

Figure 3. Y coordinates of both actors per time step[11].

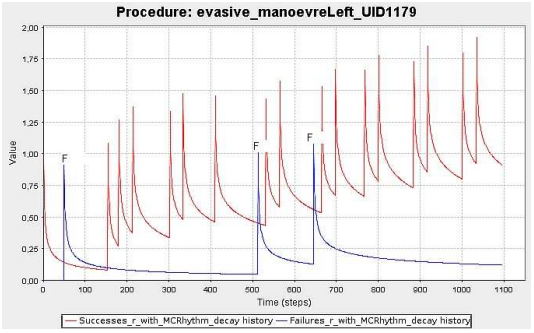

Figure 4. Agent 1: Successes, Failures Left

As an observer, if we only look at the social level and not at the cognitive level of the agents, we could conclude to see some start of "formalised" behaviour. However, after introspection of both agents (see figure 4 for an example of one agent), we observe at the sub-symbolic level of the production (strategy) 'evade left' that the agent reinforces this production, which results in a higher preference for this production than the production 'evade right'. We therefore conclude that the agent is aware of his own successes but does not care about the other agent's successes. In other words, there is not a formal agreement; the emergence exists as an evolved social construct (norm) in the head of both agents. Adaptation to their environment is purely based on their own individual experience (cognitive sub-symbolic learning).

4.2 Experiment 2: control over emergence

The first experiment, first order emergence, was mainly based on the interaction of RBot agents that make decisions based on their individual preferences. A cognitive agent solves a problem by defining a problem space and stating clear goals. Such an agent is however

[11] The graph shows the **y path** of two actors; the application is able to plot also the *x path*, but the y path is sufficient. The moment the paths are overlapping, the actors choose a different strategy and have conflicts because they follow the same y-path. When the paths are each others opposite, or mirrored, then the actors choose the same strategy and no conflicts arise.

not (socially) embedded in the environment, it responds not instantaneously to changes that occur in the (social) environment.

In this experiment we want to demonstrate the impact of social constructs, and particular how a social construct can be part of a coordination mechanism that allows an agent to get control over others' emerging behaviour(s). To get control over agents (car-drivers) and emergent behaviour, a police-officer can be assigned the task of controller (socially empowered by society) to correct behaviour of agents that do not behave according the society's norms or rules.

The experiment has the same initialisation as the previous experiment with the exception of the policeman. The policeman is given two social constructs: (SC1) The social construct of evading to the right side as stated by the government of, for instance, the Netherlands, and (SC2) communicate the social construct, SC1, of the preferred strategy to the other agent when he does not behave according to the norm (tries to collide by staying to long in front of the policeman). Hence, when a chance of collision starts to occur, the other agent needs to receive the social construct with the preferred strategy and be able to store it in memory and act accordingly.

The experiment is kept simple and only demonstrates the impact of a social norm on the behaviour of interacting agents. Therefore, the following assumptions have been made:

1. The agents are defined as being aware of their role, (1) Agent2 as policeman and (2) Agent1 obeying the policeman.
2. No negotiation takes place over the announcement of roles; the roles and relations are predefined by society.
3. Punishment is not included; however, it can be argued that by making the other agent aware of a certain social construct or norm is a correction of the behaviour. This can be regarded as a light form of punishment.

First we will have a look at the social or interaction level, see figure 5. We observe that when a collision is about to take place, Agent1 is immediately corrected in its behaviour and is forced by the policeman to choose the 'rightlanedriving' procedure. The control is so strong that after one encounter, the behaviour settles into a stable pattern. At the individual level of Agent1, see figure 6, we see a huge

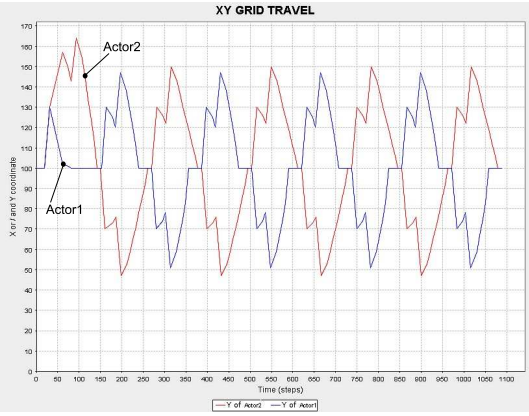

Figure 5. Y coordinates of both actors per timestep.

increase in utility in favour of driving on the right side. The correction being made is approximately a factor 10 and shows that Agent1 has become aware of the social construct 'rightlanedriving'. Besides that, the application software allows us to inspect the memory where we also observed that the social construct has been stored in memory of Agent1.

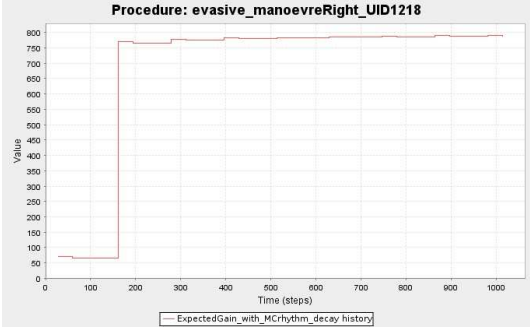

Figure 6. Agent 1: Utility Right

The experiment demonstrates clearly that the policeman is able to control the start of the emergence and keep control over the emergence by sending the social construct with the preferred strategy, be it by speech or by signalling his hand.

5 DISCUSSION

Emergence is around us in all kinds of varieties. However, to note that there are some properties that causes a system to emerge into a certain direction is not good enough. One has to look for a general explanation of (social) phenomena by studying not only the social or group level but the cognitive or individual level as well. Especially when studying second order of emergence in which individuals become aware of the emergence and the implications for others and themselves, then one *has* to look at the cognitive level of the agent and wonder if a differential equation still is powerful enough to give sufficient explanation.

We argue that especially in social situations an agent has to be able to built up representations in a cognitive plausible way, i.e. an agent should be able to express itself in a language and exchange representations with other agents. We have adopted the cognitive architecture ACT-R because it is empirically grounded and as a hybrid architecture allows for storing representations (symbolism) but connectionsm-like properties (activation) as well.

In the case of second order emergence, social constructs enable us to model social phenomena as a separate module (modularity of the mind [18]) of representations while leaving the underlying cognitive model of ACT-R as much as possible the same. The social construct mechanism described in this paper is conform the mechanism proposed by Bargh [8, p. 115] who also states that there are mental links between representations of motives and goals in memory and the representations of the social situations / constructs in which those motives have been frequently pursued in the past.

In cognitive psychology [45] and social psychology [12] empirical research has shown that there is a "dual process model", i.e. there is a distinction between implicit and explicit processes respectively automatic (reflexive) and controlled (reflective) processes. In this paper we did not put the emphasis on duality, but we can assume that social constructs or procedures, after they are first explicitly (or consciously) processed, become more of a habit or implicit/automatic process after emergence has taken place. The strengthening of the habit can be caused by the implicit association between situation and habit or by explicit reasoning (control) about the situation and the habit.

In case of control over emergence, it can be compared with government regulations that reduces uncertainty and enforces a system (society) to emerge in a way that is desired by society. However, many control mechanisms are based on semi-successful traditional

methods and are not based on research. For instance, we still do not understand how riots start to emerge; we can model group behaviour and take factors like weather condition, in-out group, relations, physical distance and so on to estimate how a riot emerges, but have no idea about the individual motives of taking part in a riot let alone how to model such an individual.

We present RBot as a general cognitive *and* social model that sheds light on the (intra)individual and the group level but not at the cost of too much complexity, lack of understanding and computational power. Besides that, its general application in the domains of social science and simulation, organisation studies, agent systems and cognitive science can hopefully revive the emergence of an interdisciplinary field and more cross-fertilisation in those areas.

REFERENCES

[1] J. C. Alexander and B. Giesen, 'From Reduction to Linkage: the Long View of the Micro-Macro Link', in *The Micro-Macro Link*, eds., J. C. Alexander, B. Giesen, R. Munch, and N. J. Smelser, 1–44, University of California Press, Berkeley, (1987).

[2] J. R. Anderson, *The adaptive character of thought*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1990.

[3] J. R. Anderson, *How Can the Human Mind Occur in the Physical Universe?*, Oxford University Press, 2007.

[4] J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere, and Y. Qin, 'An Integrated Theory of Mind', *Psychological Review*, **111**(4), 1036–1060, (2004).

[5] J. R. Anderson and C. Lebiere, *The atomic components of thought*, Erlbaum, Mahwah, NJ, 1998.

[6] J. R. Anderson and L. J. Schooler, 'Reflections of the environment in memory', *Psychological Review*, **2**, 396–408, (1991).

[7] R. Axelrod, 'An Evolutionary Approach to Norms', *American Political Science Review*, **80**(4), 1095–1111, (1986).

[8] J. A. Bargh, *Handbook of motivation and cognition*, volume 2, chapter Auto-motives: Preconscious determinants of social interaction, 93–130, New York: Guilford, 1990.

[9] R. C. Berger and R. J. Calabrese, 'Some explorations in initial interaction and beyond: Toward a developmental theory of interpersonal communication', *Human Communication Research*, **1**, 99–112, (1975).

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

[11] C. Castelfranchi, 'The theory of social functions: challenges for computational social science and multi-agent learning', *Cognitive Systems Research*, **2**, 5–38, (2001).

[12] *Dual-process theories in social psychology*, eds., S. Chaiken and Y. Trope, New York: Guilford Press, 1999.

[13] R. Conte and C. Castelfranchi, 'Understanding the functions of norms in social groups through simulation', in *Artificial Societies: The Computer Simulation of Social life*, eds., N. Gilbert and R. Conte, 252–267, UCL Press, London, (1995).

[14] D. C. Dennett, *Brainstorms: Philosophical essays on mind and psychology*, Harvester Press, Hassocks, Sussex, 1978.

[15] D. C. Dennett, *The Intentional Stance*, Bradford Books / MIT Press, Cambridge, MA, 1987.

[16] H. Fayol, 'Administration industrielle et generale', Technical report, Bulletin de la Sociéte de l'Industrie Minerale Dunod, Paris, (1916/1918).

[17] F. Flynn and J. Chatman, 'What's the norm here? Social categorization as a basis for group norm development', in *Research in groups and teams*, eds., J. Polzer, E. Mannix, and M. Neale, 135–160, JAI Press, Elsevier Science, London, (2003).

[18] J. A. Fodor, *The modularity of Mind*, Cambridge MA: MIT Press, 1983.

[19] H. W. M. Gazendam, *Information, Organisation and Technology: Studies in Organisational Semiotics*, chapter 1. Semiotics, Virtual Organizations and Information Systems, 1–48, Kluwer Academic Publishers, Boston, 2001.

[20] H. W. M. Gazendam, 'Models as coherent sign structures', in *Dynamics and change in organizations: Studies in organizational semiotics 3*, eds., H. W. M. Gazendam, R. J. Jorna, and R. S. Cijsouw, 183–213, Kluwer, Boston, (2003).

[21] H. W. M. Gazendam, R. J. Jorna, and J. M. Helmhout, 'Monitoring a social context based on social constructs', in *Interfacing Society, Technology and Organisations*, Campinas, Brasil, (2006). UNICAMP/IC.

[22] H. W. M. Gazendam and K. Liu, 'The evolution of organisational semiotics: A brief review of the contribution of Ronald Stamper', in *Studies in organisational semiotics*, eds., J. Filipe and K. Liu, Dordrecht, (2005). Kluwer Academic Publishers.

[23] J. J. Gibson, *The Ecological Approach to Visual Perception*, Houghton Mifflin, Boston, MA, 1979.

[24] N. Gilbert, 'Varieties of emergence', in *Social Agents: Ecology, Exchange, and Evolution Conference*, Chicago, (2002).

[25] N. Goodman and C. Z. Elgin, *Reconceptions in Philosophy and Other Arts and Sciences*, Routledge, London, 1988.

[26] I. Hacking, *The Social Construction of What?*, Harvard University Press, Cambridge, MA, 1999.

[27] J.M. Helmhout, *The Social Cognitive Actor: A multi-actor simulation of organisations*, Ph.D. dissertation, SOM research school: University of Groningen, 2006.

[28] K. Liu, *Semiotics in information systems engineering*, Cambridge Universisty Press, Cambridge, England, 2000.

[29] J. G. March, M. Schulz, and X. Zhou, *The Dynamics of Rules: Change in written Organizational Codes*, Stanford University Press, Stanford, CA, 2000.

[30] R. H. McAdams, 'The Origin, Development, and Regulation of Norms', *Michigan Law Review*, **96**, 338–433, (1997).

[31] G. H. Mead, *Mind, Self and Society from the Perspective of a Social Behaviorist*, University of Chicago, Chicago, 1934.

[32] S. Moss, *Cognition and Multi-Agent Interaction*, chapter 15. Cognitive Science and Good Social Science, 393–400, MIT Press, Cambridge, MA, 2006.

[33] A. Newell, 'Physical symbol systems', *Cognitive Science*, **4**, 135–183, (1980).

[34] A. Newell, *Unified theories of cognition*, Harvard University Press, Cambridge, MA, 1990.

[35] A. Newell and P. S. Rosenbloom, *Cognitive skills and their acquisition*, chapter 1. Mechanisms of skill acquisition and the law of practice, 1–56, Lawrence Erlbaum Associates, Hillsdale, NJ, 1981.

[36] C. S. Peirce, *Collected Papers*, volume I-VIII, Harvard University Press, Cambridge, MA, 1931.

[37] D. C. Rubin and A. E. Wenzel, 'One hundred years of forgetting: A quantitative description of retention', *Psychological Review*, **103**, 734–760, (1996).

[38] R. Schank and R. Abelson, *Scripts, Plans, Goals and Understanding*, Erlbaum Associates, Hillsdale, 1977.

[39] H. A. Simon, *The sciences of the artificial*, MIT Press, Cambridge, MA, 3rd edn., 1996.

[40] S. Sismondo, 'Some Social Constructions', *Social Studies of Science*, **23**(3), 515–553, (1993).

[41] R. K. Stamper, *Information in Business and Administrative Systems*, John Wiley & Sons, New York, 1973.

[42] R. K. Stamper, 'Organisational semiotics: Informatics without the computer?', in *Information, organisation and technology: Studies in organisational semiotics*, eds., K. Liu, R.J. Clarke, P. B. Andersen, and R. K. Stamper, Kluwer Academic Publishers, Boston, MA, (2001).

[43] R. K. Stamper, K. Liu, M. Hafkamp, and Y. Ades, 'Understanding Roles of Signs and Norms in Organisations', *Journal of Behaviour and Information Technology*, **19**(1), 15–27, (2000).

[44] R. Sun, 'Cognitive science meets multi-agent systems: a prolegomenon', *Cognitive Systems Research*, **14**(1), 5–28, (2001).

[45] R. Sun, *Duality of Mind*, Lawrence Erlbaum Associates, Mahwah, NJ, 2002.

[46] N. A. Taatgen, *Learning without Limits: From problem solving towards a unified theory of learning*, Ph.D. dissertation, University of Groningen: Faculty PPSW, Groningen, NL, 1999.

[47] H. Van den Broek, *On Agent Cooperation: The relevance of cognitive plausibility for multiagent simulation models of organizations*, Ph.D. dissertation, SOM Graduate School / Research Institute : Systems, Organization and Management, 2001.

[48] A. H. Vera and H. A. Simon, 'Situated Action: A Symbolic Interpretation', *Cognitive Science*, **17**(1), 7–48, (1993).

[49] F. A. Von Hayek, *Law, legislation, and liberty: The political order of a free people.*, volume 3, Routledge & Kegan Paul, London, 1979.