A conceptual architecture for social deliberation in multi-agent organizations

Frank Dignum[a,*], Virginia Dignum[b], Rui Prada[c] and Catholijn M. Jonker[b]

[a]*Utrecht University, Utrecht, Netherlands*
[b]*Delft University of Technology, Delft, Netherlands*
[c]*INESC-ID and Instituto Superior Tecnico, Universidade de Lisboa, Lisboa, Portugal*

Abstract. Socially interconnected systems should be able to exhibit, understand, and reason about social behaviour, in order to support lasting interconnections that show realistic and desirable properties. We claim that a fundamental new approach based on social concepts is needed to build agents functioning in these socially interconnected systems. In this paper, we explore how the concepts of social practices and social identities can be used to structure deliberations about interactions. We then show the consequences for the architecture and reasoning capabilities of these systems.

Keywords: Social deliberation, multi-agent organisations, agent architecture, social practices

1. Introduction

When trying to build socially interconnected systems it becomes quickly obvious that the agents in these systems should be endowed with forms of social intelligence that allows for them to deliberate about socially realistic interactions and be able to operate in increasingly complex environments and situations. The potential for artificial intelligent systems to interact and collaborate not only with each other but also with human users is no longer science fiction. Healthcare robots, intelligent vehicles, virtual coaches and serious games are currently being developed that exhibit social behaviour to facilitate social interactions, to enhance decision making, to improve learning and skill training, to facilitate negotiations and to generate insights about a domain. Agents in these systems should be endowed with forms of social intelligence that allows for them to deliberate about socially realistic interactions [16].

We informally define social intelligent agents as systems whose behaviour can be interpreted by others as that of perceiving, thinking, moral, intentional, and behaving individuals; i.e. as individuals that can consider the intentional or rational meaning of expressions of others, and that can form expectations about the acts and actions of others [29].

In the last years, many systems have been developed which possess some of these characteristics. In particular, work on Intelligent Virtual Agents and on Social Robotics has delivered many promising results. However, we are still lacking theories, tools and methodologies to guide and ground these developments. That is, current approaches often result in pragmatic, but limited and unstructured solutions that are hard to generalize or integrate with each other.

---

*Corresponding author: Frank Dignum, Utrecht University, Utrecht, Netherlands. E-mail: F.P.M.Dignum@uu.nl.

In this paper, we propose an architecture to endow agents with a social deliberation mechanism that has been engrained in the core of their reasoning, in a similar way as has been evolved in humans [9]. We start from the premise that social interaction (and reality) is not given, but socially constructed [7,30]. The consequence of this stance is that the social interactions are not seen as rigid, but that they arise from frequent execution of similar patterns of interaction which can adapt to the changing environment whenever necessary. Given that we assume that the interaction patterns themselves are flexible we need to use more stable underlying social goals and purposes of these interactions to design the software operating in these environments. Current systems still have very limited understanding of their context, and of their social role. They are therefore not able to reason about their identity and goals in a social context, and therefore cannot be expected to function outside the situations they have been designed for and actually create the social interactions as required.

Currently, sociability is engineered into the system, in a situational and context dependent way. Social signals are not appraised as such but implicitly built into their functionality. Therefore, the behaviour of the system is not conceived as social outside that particular context, and the system is not able to adapt to significant changes. This implies that reuse in different social contexts (such as interacting with different types of suppliers) or cultures (connecting with organizations abroad) often requires a complete re-engineering of the system. A next step forward in AI, is the ability to perceive, reason about and exhibit *socially* intelligent behaviour. This will require a framework containing explicit social principles that can be described, represented and manipulated in a symbolic way.

We argue that deliberative, social, and physical principles must be considered first class components of a computational theory of social intelligence. Being socially intelligent requires a keen understanding of the principles of social reality, and the ability to link social interpretations with individual goals into plans and vice versa. The interrelationship between social and physical contexts is such that the social context defines the possible social interpretations of the physical reality and limits the set of admissible actions; and the physical context determines and constrains the possible social contexts. For example, a raised hand can mean many things: in a class room: a question, in an auction hall: a bid, on the street: a greeting or threat.

It is clear that splitting the context in a social and physical one adds quite some complexity. If an agent plans for a social goal (such as gaining acceptance in a group of companies) it needs to plan physical actions to reach such a goal. Thus, any plan has both social as well as physical consequences. The agent has to deliberate about all these elements in order to decide upon the best cause of action. It might be clear that agents need some structures in order to limit and facilitate this process. In this paper we describe the first steps towards agents which can be called socially intelligent.

As a running scenario of application, in this paper we consider a situation where a company A is deliberating on their response to a merging bid from company P. The companies are competitors on the pharmaceutical market. However, the bidding company P is a large, profit-driven, company. Company A is much smaller and geared towards research and development. Accepting the bid might give A many new possibilities to market their products. However, A might also loose flexibility to do blue sky research and might loose jobs that way. Members of company A act according to four core values: innovation, disregard for money (profit), independence and intellectual superiority. During the negotiations, the differences in values, culture and social identity became clear and may cause difficulties in reaching agreement.

In Section 2, we present the background from Social Sciences and we identify the key elements of a framework to build socially intelligent systems. We describe how motives can be used to give an agent both social as well as physical direction. Then we describe how social identity, norms, values

and social practices can be used as means to preserve both consistency (at individual and society level) as well as simplicity. We propose our own abstract architecture for social reasoning in Section 3. This architecture is, of course, built on top of the concepts introduced in Section 2. In Section 4, we sketch how this architecture can be used in order to determine a course of action in the example scenario in an efficient way. In Section 5 we discuss the social practice as a process and how this process determines its effectiveness. Finally, before we draw some conclusions and indicate future work in Section 7, we discuss related work in Section 6. Although related work is often discussed before the presentation of own work in this case we believe that it will be easier to understand what we see as related work after we have introduced our approach and resulting agent architecture and deliberation processes.

2. Motivation and background

In order to create truly socially intelligent systems, we need to start from an understanding of the motives for behaviour [14]. Human behaviour results from a need to balance between novelty and control. I.e. in the one hand, seeking out new situations, while, on the other hand, trying to avoid uncertainty and, thus, strive to predict and control the environment. The balance between the two forces is different between persons, but always present and prominent in the way people deal with social interactions. However, it does not readily indicate how people *choose* concrete actions. Thus, we look at theories of human motivations [26] in order to get more concrete handles on *drivers* of behaviour. McClelland argues that there are a number of basic natural incentives that give rise to some motives. Besides the biological (homeostatic) motives such as hunger and need for sleep (which are, in fact, not very salient in most of the social situations), McClelland distinguishes four motives: (1) *achievement*, (2) *power*, (3) *affiliation* and (4) *avoidance*.

Achievement is about achieving goal states and drives people to try to achieve different things, thus, fostering explorative behaviour that satisfies the need for novelty.

Power is about trying to have an impact on the world, and includes both the need to control its own and other's actions.

Affiliation drives people to seek the company of others and to establish and maintain positive interactions with others.

Avoidance leads to self preservation, seeking certainty, and emotional regulation, and fosters the categorization and simplification of behaviour so that it becomes more standardized (and thus predictable).

The achievement motive clearly serves to satisfy the need for novelty as the agent tries to achieve a different (new) situation. Of course the degree of novelty in a situation depends on how much an agent can predict the outcome of its actions and the consequences of reaching a certain state. Driven by the motive for achievement whenever people enter a new situation, they immediately start exploring the environment. This exploration can take long or can be done in a split second depending on circumstances. The exploration enables a person to react quickly to events as the appropriate social practices in relation to the possible social identities have already received a high probability of execution. One could argue that this motive also can be used to drive more social goals, such as, to try to achieve a certain position or reputation in a social group (e.g. to be the most popular or most helpful). In other words, to achieve goals in the state of the social world. However these type of social goals can only be achieved using plans that also have physical consequences. Thus the planning process for these goals should include different (new) aspects from those used in current planning frameworks.

The power motive is about trying to have an impact on the world and reach a sense of control. Thus, this motive clearly serves the need for control. For people, this leads to behaviour that tries to change the physical world just in order to see that one has the capability to do that, but is also used to impact the social context in which people live. Training skills (such as negotiating) can be driven by the power motive when the mastery of the skill gives a person more control over his environment. If an agent masters a skill it can predict the outcome of its actions better and thus control the amount of uncertainty. However, it also leads to attempts to influence other people and engaging in status and power manoeuvres with others in order to control the people in one's environment. Again, if one can control other people it becomes easier to predict their behaviour and thus control the amount of uncertainty.

The affiliation motive drives people to seek the company of others. However, it is not just the company of other people that is needed, but rather to establish and maintain positive interactions (relations) with those people. Therefore, one wants positive interactions that give emotional rewards to all parties involved and lead to further interactions. So, both the quality as well as the quantity of the interactions influence the satisfaction of the need for affiliation. This motive serves a need for social novelty. Contacts are more interesting if one has to add something to the feelings and knowledge of the other. Thus, one constantly tries to variate the interactions slightly in order to keep interactions going. Whether one interacts with the same persons or different persons depends per individual. If one needs more control it is safer to stick to people one knows, but if one needs more novelty meeting new persons is positive. It also indicates why we do not in general like to affront other people. This goes against the affiliation motive. Picking fights, insulting etc. only takes place if alternative persons are around to be affiliated to (e.g. insulting someone from an out-group might be good as it increases the status and power within the group).

Finally, the avoidance motive drives people to avoid conflicting and/or 'bad' situations. That is, if interactions with another person are not pleasant, e.g. leading to high levels of anxiety and discomfort, one will withdraw and avoid future interactions with that person. The motive also is active in a broader sense that it tries to avoid situations in which there is a large difference between the perceived and expected situation. That is, situations with a large cognitive dissonance. It leads to self preservation, seeking certainty, and emotional regulation, which fosters the categorization and simplification of behavior so that it becomes more standardized (and thus predictable). So, it is very clearly serving the need for control. It is a main motivator for trying to categorize the environment in which a person lives. Categorizing and standardizing means that real-time reactions can be generated in most situations and thus more (and new) situations can be sought out. This is of prime importance in a more and more complex and dynamic world. We will discuss this issue in more detail in the next subsection.

Taking human motives as starting point has a number of important consequences for the design of socially intelligent systems. Firstly, motives provide a balance between approach and avoidance mechanisms. E.g. where the power motive can lead one to seek dominance over other people, the affiliation motive makes sure that this is not done at all costs, and not always, but is kept within "socially acceptable" bounds. In the same way, the achievement motive leads people to explore new ways to achieve goals, but the avoidance motive takes care that we avoid too much deviation from known situations. Thus, having a system of basic motives like this supports a flexible and situational guidance of agents in a dynamic social context. Secondly, and most importantly for the aim of this paper, these basic human motives show that motivations are not only achievement directed but are mostly social. In particular, social motives lead to different types of plans and goals in all settings where humans (and agents) work together. E.g. The need for getting status is a *social* need that might not lead directly to any physical goal. From the literature [4] it is clear that teams need two different types of leadership (possibly combined in the same person), one taking care of the achievement goal of the team and one taking care of

the affiliation needs of the team. I.e., skills and capabilities to ensure that the team actually performs its task in an efficient way are different from those needed to ensure that team members are happy doing that. Standard BDI models cannot distinguish these different types of goals, nor the consequences of this difference.

As said above, people strive to keep a balance between novelty and control. When we consider environments with many people and social interactions between people (requiring both physical and social planning) it is clear that the need for novelty is easily satisfied by the environment. The agent does not need any special mechanisms for creating novel situations besides the achievement and affiliation drives that it has. Rather, there is a clear need to simplify the context of a social agent, satisfying the need for control, in order to be able to deliberate and effectively decide on a course of action in a complex and dynamic environment in which decisions have to be made real-time.

In the next section we discuss a number of mechanisms studied in the Social Sciences, that play an important role in human deliberation and that can be used for socially intelligent agents as well. These mechanisms represent concrete ways to fulfil motives at a very high level of abstraction. We describe three constructs that are used in the Social Sciences to ensure consistency of behaviour over time: Social Identity, Norms and Routines/Habits. Finally, we discuss the concept of *Social Practices* as a way to describe the combination of social and physical context with respect to a (standard) course of action. We will briefly discuss how the other three concepts could be defined in terms of social practices. I.e. all these concepts are related and merely distinguished because they emphasize a particular aspect of practices and it is for many purposes handy to focus on only a part of social practices.

In the remainder of this section, we provide a brief description of several Social Science theories of behaviour that are used as basis for this work.

2.1. Social identity and values

A social intelligent system must be able to perceive itself and (its position in) the social world. People position themselves, and others, in terms of membership of, possibly many, social groups (i.e. reference groups) and social goals are often based on comparison with others [33,34]. For example, if you want to be a good CEO this means that you identify yourself (at least partly) as a "CEO" and you need to know the position and activities of some (prototypical/ideal) good CEOs such that you can ascertain what kind of action is needed to become respected in that group. Some reference groups are quite stable, such as family and profession while others are more volatile, such as the group of people in a shop or at a meeting. People have different emotional attachments to each of the social groups, which elicits social goals to maintain and pursue certain identities.

What constitutes a "good CEO" relates to the set of values and their priorities associated with a reference group. Many definitions of values exist and many research communities use them in different ways. We see them as criteria with which pairs of situations can be ordered. E.g., the value "environmental friendly" can be used to compare two situations on the basis of how well the nature is preserved in each of them. It can very well be that another value, such as, "comfort" will sort the two situations exactly the other way around. Values are used to reconcile the different reference groups the person belongs to, such that her behaviour is consistent (and expected).

From the above it follows that once a situation triggers a social identity in the agent (or if it chooses to deliberately portray it – e.g. for strategic reasons), the identity comes with a set of values and, thus, prescribes a certain type of behaviour. For example, when a person has the social identity of a research manager she will support researchers in performing their experiments at any time of the day. There is no

deliberation about whether she would prefer to go shopping or go on a date at that moment. Identities are also social, because they give people the sense of belonging to the reference group. The identity is visible for others and expectations can be formed on the basis of knowledge about the reference group. Thus, the identity provides consistency of behavior on an individual level (because an individual with an identity will behave according to that identity) as well as on social level (because all individuals with a certain identity will behave similarly in the same context).

2.2. Norms

Norms are the second construct that can be used to categorize and classify behaviours. Norms specify behaviours that promote values. We will not get into all the different kinds of norms and description of their properties at this place (we refer to [3] for a recent overview on norms in multi-agent systems). For the purpose of this paper the function of norms as behavior regulation is the most important. Regulative norms can be described by seven elements: the activation- and termination condition, the normative direction (obligation, permission or prohibition), the action, the violation condition, the punishment and the repair. Norms will indicate whether in a certain context (when the activation condition is true) an action is obliged, permitted or forbidden. E.g. "finishing an experiment goes before attending a meeting". The activation condition of the norm is that a meeting is scheduled while the experiment is not finished. The termination condition for the norm is when the experiment is finished. A norm like "perform experiments according to scientific standards" is always active and thus the activation and termination conditions are "true". The violation condition indicates how one can detect the violation of a norm. Thus, what counts as an action to signal the beginning of a meeting? Does that start with a timer? Or should a phone-call of the chair be involved? The punishment and repair are the actions needed to get from a violation situation to a normal situation again. E.g. the punishment for manipulating the results of an experiment might be firing the employee, while repairing is to run the experiment another time. Although the punishment and repair are important for how norms function in society, the mere fact that the norm exists and is known is enough for most people to follow the norm in normal situations, without deliberating for alternatives.

Norms also have an individual and social side just like identities. When an individual has accepted a norm it means that that individual will act according to that norm (in the appropriate context where the norm is active). Thus, the norm ensures individual consistency of behaviour. The norm also has a social side, because norms indicate what is socially acceptable behaviour. Thus, they ensure consistency of behaviour, not just over one individual, but over all persons for which the norm is active. Often the activation conditions of norms refer to reference groups that are also used for identities. It, therefore, seems logical to include the norms pertaining to particular reference groups with the description of that identity. E.g. the "researcher" identity will come with some norms on how to behave at work and meetings. It should be kept in mind though that we do not assume a particular order of importance between norms and identities. The particular context will determine the most salient aspect and based on that the other aspects can be connected and become salient as well.

2.3. Habits

According to [27] habits are psychological dispositions to repeat past behaviour. Our daily life is full of habits, as ways to simplify decision making. When a certain behaviour has become a habit we do not deliberate about that behaviour anymore, but just repeat it whenever it is triggered by the context. As we will see in the next part on social practices habits are close to practices. However, we see habits as

something that can in principle also be completely material driven, while social practices also have a social dimension.

According to the psychology literature habits are related to goal directed behavior and their origin can often be traced to a deliberate action with a particular goal. However, when the action is repeated often enough in the same condition we can become "conditioned" and perform it without thinking about the goal anymore. E.g. many people have the habit to get up when the alarm sounds. After that they perform a sequence of actions including getting dressed and eating breakfast. All these actions are often performed half-conscious and without deliberation about a "best" course of action. Creating habits frees up our mind to deliberate on more important things and leaving the reaction to many daily recurring situations to our habits.

As many behaviours are performed out of habit, habits can also be used to predict behaviour. This is explicitly done in the Consumat agent model that is used for consumer simulations [20]. It makes the agents simple and the focus can be on the cases where habits are broken. Habits also lead to consistent behaviour. Exactly because they are repetition of the same behaviour in similar conditions.

2.4. Social practices

In the previous sections, we looked at several mechanisms to simplify deliberations about actions while keeping consistency. We now turn to the issue on how to manage the three elements of social intelligence mentioned in the introduction: deliberation, social and physical reality. We need to look for structured theories that govern the rules with which these three elements are related. In the social science in recent years the concept of *social practice* seems to target exactly the same problem. Social practices are patterns of activity meaningful to people as parts of their everyday life, and the way these are typically and habitually performed in (much of) a society, e.g. going to work, cooking, or showering. In social science, social practices are formulated by some researchers as taking the centre stage to the extent that people, and things, occupy secondary roles as the carriers of practice [31]. However, in this paper, we look at the other side of the coin, where social practices are used by persons, or agents, as methods for deliberation. I.e. agents will identify suitable social practices given a situation and take those as heuristics for appropriate behaviour in that situation. Researchers in social science [19,28] have identified three broad categories of elements of social practices:

- Material: Covers all physical aspects of the performance of a practice, including the human body and the actions that can be performed as part of the social practice. This relates to our physical aspects of a situation.
- Meaning: Refers to the issues which are considered to be (socially) relevant with respect to that material, i.e. understandings, beliefs and emotions. This relates to our social aspects of a situation.
- Competence: Refers to skills and knowledge which are required to perform the practice. This relates to our notion of deliberation about a situation.

One could see a social practice as an elaborate condition-action rule. First, a situation is assessed in order to check whether a social practice is relevant. When it is relevant all material elements get a social interpretation (partly) determined by the social practice. Thus, a group of people can be given meaning as being "my friends", "a rival management team" or "managers", depending on the social practice that is activated. Within a social practice several behaviours are possible. Depending on the competence that a person has, she will disregard some. Then an action is chosen for which a kind of standard social effect is also expected within this social practice. Of course all elements from the previous sections play a role in social practices as well. People in a social practice are given social meanings through their social

identities. The actions available in social practices are often (social) norms and when a social practice is often followed the actions within it become habits. In this way all the concepts of the social framework are directly related.

The components are linked by individuals when carrying out a practice. Each individual embeds and evolves meaning and competence, and adopts material according to his or her motives, identities, capabilities, emotions, and so forth, such that a practice can then be implemented as a composition of components. However, because the social practice is also a shared notion one can expect the other participants in the social practice to share the same interpretation of the situation. There is an (expected) shared situation awareness.

Individuals and societies typically evolve a collection of practices over time that can be adopted in different situations. Social practices are like social norms in that they emerge from individuals, but are not dependent on the individuals any more. They are continuously shaped when they are followed and can differ for individuals with different experiences. E.g. we all share an understanding of the greeting practice, but the exact behaviours and social connotations may differ. Moreover, depending on the situation, the personality and the skills of an individual, carrying out a practice will be a more automatic or a more deliberated process. This corresponds to the different modes of thinking proposed by psychologists:[1] [21]

- System 1, or fast thinking, operates automatically and quickly, with little or no effort and no sense of voluntary control. This includes recognition, perception, and orientation.
- System 2, or slow thinking, allocates attention to activities that demand a high amount of mental effort. Such activities include complex computations, rule following, comparisons, and weighing of options. The operations of System 2 are often associated with the subjective experience of agency, choice, and concentration.

When a social practice is experienced very often the interpretation of the situation can be done in a standard way and leads to a quick decision on an optimal behaviour. E.g. when driving a car we hardly ever think about using the shift or the clutch when changing gear. However, when learning to drive we have to first learn which are the salient elements in the environment that trigger the gear change. Thus, we react slower and need more attention for the driving behaviours themselves. This is an important issue for persons in crisis situations, where decisions have to be taken quick. Experienced persons will very quick distinguish the salient elements in the situation and decide which social practice is most salient and act according to it. In the next section we will go more in-depth into the architecture for individual deliberation based on the use of social practices.

The concepts described in this section (motivations, identities and values, norms, habits and social practices) originate from different social science theories. Therefore they emphasise different aspects of social behaviour. However, they are linked around the patterns of behaviour for social interaction. They all assume a repetition of behaviour that is learned and communicated within social groups or cultures. As such, social identity can be seen as a set of coherent social patterns that defines a group. Habits describe the process of learning and internalising such patterns (ignoring for the most part the social aspects of the behaviour), Norms are abstract specifications of patterns of behaviour that are associated with a moral stance. Finally, social practices provide a contextual scope for patterns of behaviour describing in possibly great detail the situations in which such pattern is applicable, and the results that can be expected from its application.

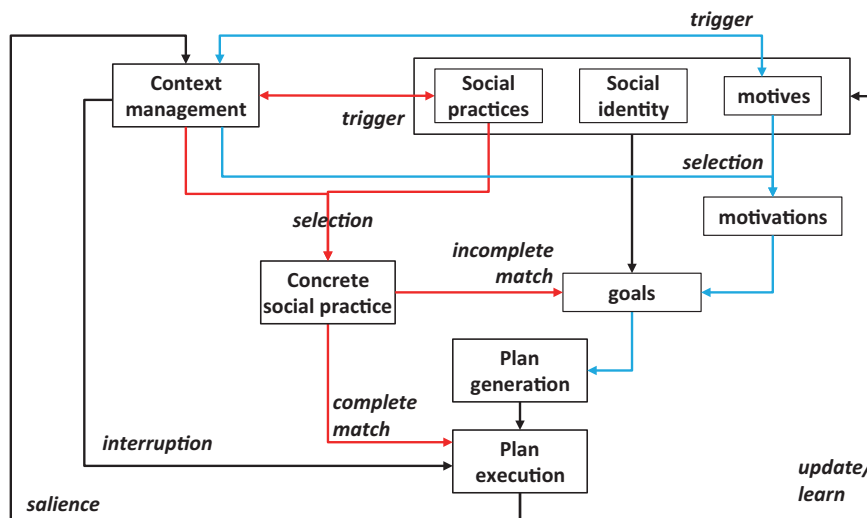[1]The labels of System 1 and System 2 are widely used in psychology.

Fig. 1. Abstract architecture for social reasoning. (Colours are visible in the online version of the article; http://dx.doi.org/10.3233/MGS-150234)

3. Architecture for social deliberation

The social practice model proposed by [19] (cf. Section 6) treat the social practice as the basic building block of the theory, and do not explain how social practices are used in deliberation and how they are influenced by and influence individual agents. The abstract architecture, depicted in Fig. 1, is a first attempt to capture this interaction.

The architecture shows the prominent place of social practices as an early input into the deliberation process. One of the modules is the context management module where the first interpretation of the situation takes place. However, this process is not only fed by external perceptions but also by the motives and the social practices. E.g. when entering a meeting with many people a person with a high affiliation motive starts looking for (possibly known) persons to interact with, while a person with a low affiliation motive might scan the room for a place to stand quietly. In addition, as soon as social practices are selected as potentially fitting with the current situation they will drive the search for salient features in the environment that fit that social practice as well and might be used for determining (further) actions. E.g. if a person is getting to a meeting it will look for extended hands. This does not mean that this is the only feature that is searched for. Several patterns can be searched in parallel. However, their number is limited to patterns that can be *expected* within a current social practice. The process of sensing in the context of a social practice leads to what we call a concrete social practice. I.e. a social practice for which many parameters are filled in based on the current context. Another way social practices drive the context management is when a social practice is active. This social practice indicates important states that can be expected and are subsequently searched. E.g. when a meeting room is entered for a meeting a person will check when the boss is sitting down and where and use that as trigger to sit at an appropriate place as well.

Thus, we see that the parallel tracks of pro-active and reactive behavior already start with the sensing behavior. We have emphasized these tracks by using red arrows for the quick (reactive) deliberation and blue arrows for the deliberation involving the motives and other cognitive concepts leading to the slower deliberation. Whether the motivation actually leads to setting a new goal is influenced by the

social practices again. If the person is entering the room when a meeting is about to start, he might find a quiet place and people will be happy that he does not interfere with the meeting procedure. However, when there are after meeting drinks the social practice might dictate that the person mingles and talks with others.

If the concrete social practice that is detected to be relevant is specific enough to leave only one course of action open, this course of action is directly executed. E.g. when a person recognizes that a handshake is used to greet another person she will immediately extend a hand to start the handshake, without deliberation about possible other actions. However, in the meeting between hostile companies the persons might not be willing to shake hands with their opponents. In that case a greeting should still be performed but the course of action is not directly clear from the situation and more deliberation takes place. In the figure this is illustrated by the fact that the concrete social practice is used as input for the goal module that will determine possible goals within the context of this social practice. This deliberation can contain a complex process itself, such as, the ones used in FAtiMA [11] or BRIDGE [13] but here we have limited ourselves for simplicity to goal deliberation. This can lead to quite complex deliberation already. E.g. suppose a CEO is strongly achievement driven. During a management meeting he will try to use a social practice that serves that motivation and maybe directs all discussion towards quick and efficient decision making. However, in order to achieve the ultimate goal of the decisions he needs the strong commitment of his staff. Thus, in order to get quick decisions with a high amount of commitment he will give opportunity for all persons to give their opinions before making a decision that reflects the inputs. He will balance different aspects in choosing a course of action within a social practice.

When discussing the architecture on the basis of the above scenarios the social practices take a leading role in organizing possible courses of action. Note that our architecture does not depend on a fixed set of plans per goal nor that it needs a large set of plans to be searched through. The social practices combine material and social aspects in such a way that one can start from either side and check the appropriateness of the other aspect for the current situation. This avoids having to reason separately about both aspects and combining them afterwards. Having the social practices can also instantiate elements in the deliberation even if they are not totally clear from the initial interpretation of the context, such as the roles and expected goals.

The final aspect that we included in the architecture is the learning that takes place after the action has been executed. After each action the system should not just check whether the action succeeded or failed, but also whether it can use the result as feedback on the choices it made during the deliberation and whether it should refine or adjust its library of social practices. E.g. it might notice that it expects a handshake in the greeting practice while not everyone is shaking its hand. Thus, it can extend this social practice with some alternative ways of greeting like bowing or hand waving. However, it might also learn that it successfully completed the social practice of decision making and update the priority of the plans it executed for this social practice. In this way it can update its memory even without explicitly storing every interaction. Finally, we should remark that where physical effects of actions can usually be measured with sensors, the social effects are often not visible and have to be derived from consequent actions of the partners. Thus, more subtle sensing and interpretation is needed to learn the most efficient social interaction patterns.

4. Social practices in practice

The previous section illustrates how our proposed architecture for social intelligent systems could function. This section describes how the elements of social identity, norm and social practices can be

---

Algorithm 1 Social Deliberation Cycle

1: *salId ← SalientIdentity* ()
2: *socPracList ← AssociatedSocPrac (salId)*
3: **for** (*sl ∈= socPracList*) **do**
4:       **if** *applicable* (*sl, context*) **then**
5:             *isl ← Instantiated (sl, context)*
6:             **while** *action ← NextAction (isl)* **do**
7:                   *action*

---

used in the scenario sketched in the introduction in order to provide efficient and consistent behavior, while still reacting to events in the environment.

In the meeting between managers of both companies to discuss the merger in the scenario, their differences became apparent. Company P managers wear a suit and tie, while managers of company A are casually dressed. Moreover, the managers of company A use the meeting as an opportunity to express their "independence" and "intellectual superiority". They might act aggressively, being impolite, interrupting, and insulting members of the other group. As a consequence, negotiations can reach an impasse, and the government decides to intervene. Given the economic and labour consequences of cancelling the merger, the government has a large stake in the process.

However, the presence of the government in the negotiations may lead to different results. On the one hand, both companies may perceive the government officials as meddling or repressing, and united by their common identity as captains of industry, combine together to oppose the government as it limits their independence. On the other hand, the intervention by the government may lead the managers to realize that they have a common responsibility towards the country and their employees, and will increase their efforts to reach a suitable merger agreement.

From the above scenario sketch it becomes clear that people use their identities and motives, and their assessment of the current situation, to determine which action to take and what the meaning of the behavior of someone else is. Roughly, the deliberation cycle goes as follows (illustrated in Fig. 2):

1. *Salient Identity.* The most salient Identity given the situation is determined. Given the context characteristics and possibly the Motives of the agent, one or more Identities became salient. E.g. being at home in the evening, triggers the saliency of the family Identity, while entering the working building triggers the employee Identity.

2. *Abstract Social Practice.* Identify the abstract Social Practices associated with that Identity. There are possibly different Social Practices applicable for an Identity, e.g. For the employee Identity, one can consider the Social Practice of work-meeting, morning-coffee-break and so on.

3. *Match Context.* Determine if those Social Practices match the current context. By analysing the Material and Meaning aspects of a Social Practice, the agent can then determine which of the Social Practices is most relevant for the situation.

4. *Instantiation.* For a Social Practice matching the situation, identify a possible concrete instantiation for a matching Social Practice. Further analysis of the context supports the instantiation of the Social Practice into more concrete versions. E.g. for the meeting Social Practice, the presence of other actors all wearing suits and ties, will determine that one should follow the more concrete Social Practice of board-meeting.

5. *Decide on Action.* For the instantiated Social Practice, decide which action to take from the actions associated with that Social Practice. This choice will also depend on the agent's own goals, past experiences, skills, etc. However, this choice is limited by the actions assumed as normally accepted in the situation, as described by the social practice.
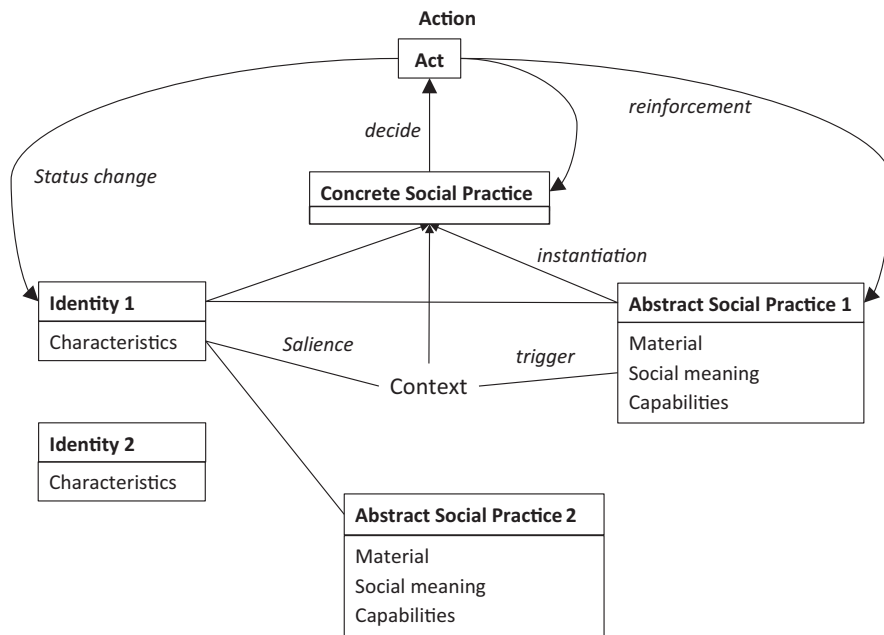
Fig. 2. Abstract architecture for social reasoning.

6. *Take action.* The planning and execution of actions is performed by the agent according to its architecture as for instance the architecture for social-aware agents described in Fig. 1. Note, that the agent expects other agents in the Social Practice to take action according to their roles in the Social Practice.

7. *Evaluation.* Evaluate the result of actions as feedback to Identity (change of status and commitment) and to Social Practice (reinforcement). Through this evaluation process, agents will learn and adapt Social Practices according to the success of their application of the Social Practice to the given situation.

We show the use of this process by applying it to the deliberation by actors in our scenario. We assume three identities "Manager", "Government" and "Employee (of company A)". All actors have the default "Citizen" identity. Figure 3 shows that an actor may decide that its identity of "employee of A" is the most salient in the situation where two groups of different companies encounter each other. The situation triggers the instantiation of the social practice "Oppose out-group" to the social practice "Be aggressive towards managers P". The actor may then decide to take the action "Intellectually insult managers P". Even if the result of this action is the failure of the meeting, it may still contribute positively towards the agent's status, given that it can be seen as conforming to the social identity of the group, i.e. being a 'hero' within the group. However, overdoing it and therefore resulting in a failure when a positive outcome was possible while maintaining its status in the group, will lead to a re-evaluation of the applicability of the social practice and might eventually lead to abandonment of the social identity.

In order to determine which of the actions associated with the social practice to take, the agent will take into account its goals, norms and habits, as described in Section 3. In fact, the 'decide' activity identified in Fig. 2 is achieved by the deliberation architecture illustrated in Fig. 1. Moreover, a formal language is needed to describe the contents of a social practice in a level of specificity that enables agents to reason about those concepts. In Fig. 2 and its concretisations in the remainder of this section (Figs 3
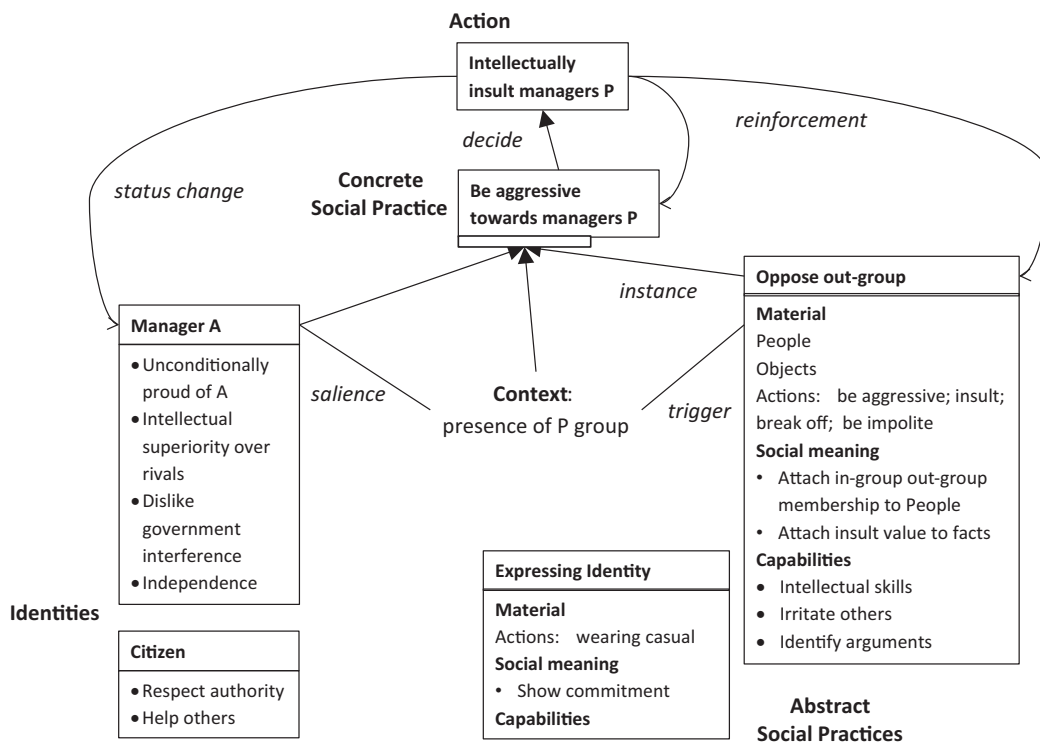
Fig. 3. Social deliberation for a manager.

and 4), we merely assume the structure of social practices defined by [28] and described in Section 2.4. This structure does not provide enough information to be used in agent's deliberation. We are currently developing a formal language for the specification of social practices.

In common situations, e.g., daily activities, agents will resort to habits, which are practices that require little to no deliberation and yield acceptable results (fast thinking). In such a case, there is a complete match between the information sensed from the context and some social practice the agent knows about (c.f. deliberation architecture in Fig. 1). For instance, meetings with managers of a supplier company lead to standard greetings and exchanges of information plus possible new contract.

In other cases, deliberation is less direct and takes into account goals and motivations of the agent. The merger situation happens seldom and a manager might have worked very hard for many years to get his current position. Thus, one manager might include the avoidance and power motive and the goal of preserving salary and position. However, the founder of the company may actively seek a fight to preserve the research values of the company. The reasoning here is that intellectual superiority over rivals is a social norm in the group and by upholding that norm and showing skill (capability) in that activity improves the view others have on him.

A similar deliberation mechanism for a government official is illustrated in Fig. 4. Getting called into the negotiations between the companies he automatically will look for consensus and tries to preserve the jobs of laborers as much as possible. That is, the match between sensed input, identity (of government official) and social practice is direct and complete, and little explicit deliberation is needed. As described in the scenario, this action may cause (unexpected) reactions from the managers, from both A and P. In this case the managers see the government official as a person limiting their independence, and unite
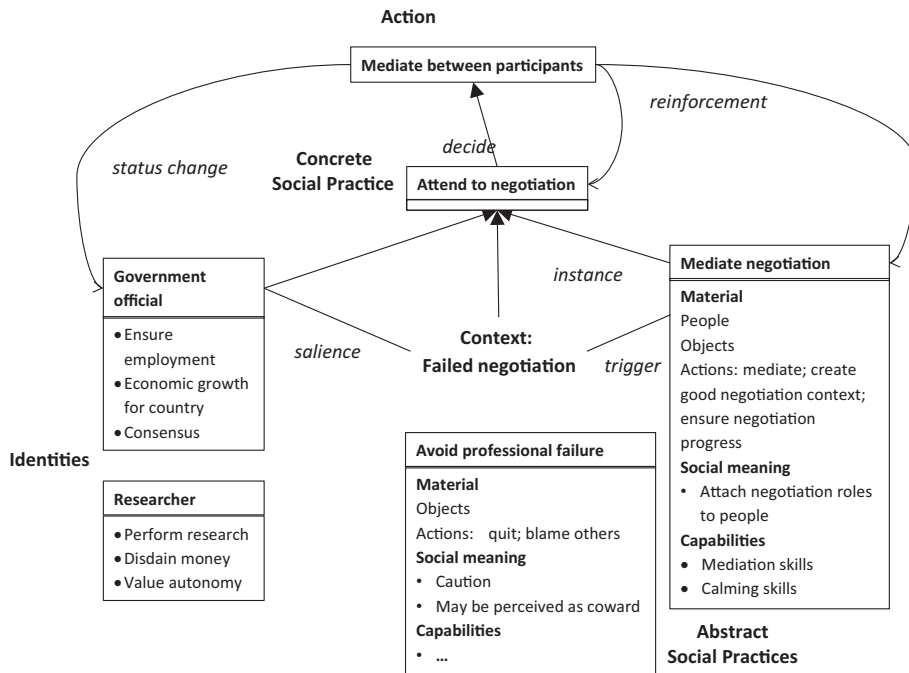
Fig. 4. Social deliberation for a government official.

to oppose the official. As a result the official now finds himself in an untenable situation. By taking into account his own goals (e.g. negotiation success) and motives, this changes the context and triggers another social practice for the official, i.e.,"avoid professional failure", which leads him to quit the negotiations and blame the companies for it.

Another example is the case that the official is himself a former researcher. In this case, he may take its affiliation motive into account, that makes him wanting to express its belonging to the group of researchers, and decide to favor company A in the negotiations.

These examples show that many different behaviours can be expressed by the integration of the deliberation architecture described in Section 3 and the social reasoning architecture depicted in Fig. 2. Even though more extensive evaluations are needed, this scenario illustrates the applicability of the main concepts.

5. Social practices as process

The previous section shows how existing social practices can be used to support individual deliberation about actions. However, that perspective does not highlight the social process of using social practices. When a social practice is used it always is done in a social context and, thus, has other individuals involved. The consequence is that when a social practice is successfully executed all the parties involved will get a positive feedback. This fact has as consequence that these parties will use the same social practice again (in the same context). Thus, the fact that the social practice is successfully used leads to a high chance that it will be re-used. This, by itself, makes the social practice already more successful in the coordination process. I.e. if you can expect in a certain context that all parties involved will use a similar social practice this will facilitate the effective coordination. It right away shows the main value

of standardization in business processes. It is not so much the exact format of the standard, but the fact that a standard is used which makes it effective.

The spread of a social practice depends on a number of factors. First of all, of course, how successful it has been to support a coordination. There is a balance here between generality of the practice and effectiveness of the support. If a social practice is applicable in many situations and contexts then inevitably the exact action patterns are also more general. If the generality of the actions leads to more mis-coordination or coordination effort (to determine the exact actions to be performed) this might lead the social practice to be perceived as less successful. However, if the social practice is very specific and only applicable in very specific situations as well it will not be used often (because it usually does not fit the context). This observation leads to a natural requirement to structure the social practices in a kind of hierarchies from general to very specific, such that individuals can easily move around the hierarchy in order to find the most specific level that is applicable in a context and that gives the most specific action support. If this practice is successfully executed not only that specific practice gets a positive feedback, but also the more general practices, thus, reinforcing that whole branch.

A concern that is directly linked to the point above is the number of social practices that are applicable in a certain context. If only one practice is applicable it is easy for an individual to choose, but if this can only be achieved with very specific practices that are hardly ever applicable the maintenance and navigation of the social practice structure becomes very cumbersome. If the applicable contexts of social practices overlap a lot, the decision which social practice to choose might become difficult. This in itself would also lead to possible mismatches when different parties involved in the coordination in that context choose different social practices. Ideally we would have a unique social practice for each context. By linking them through generalization/specialization links along the different aspects (material/meaning/competence) it is then easy to navigate to the most specific uniquely applicable practice. Feedback on its performance should then again be given not only to the particular practice, but also (in lesser extent) to the practices linked to it.

Of course, in practice social practices can emerge in parallel. Whether one will dominate the others depends on the speed with which it spreads. This depends on how often parties involved in the practice participate in coordination with different parties. The better a party is connected the more influence it gets in spreading a social practice. Also when a social practice is successful in many contexts it can be used more often and, thus, will spread quicker.

Finally the above considerations depend a lot on the context in which the coordination takes place. If the environment is very static and regular, one or two coordination mechanisms suffice to provide effective coordination. Thus, few social practices will emerge and they will spread quick and converge quick. If the environment leads to many distinct contexts with their own specific requirements for coordination (like in human society) many more social practices may arise and also disappear again (when no longer useful). On an abstract level the process of convergence of social practices is similar to that of the convergence of terminology of ontologies in multi-agent systems. This also depends on (regular) interactions, the usefulness of the terms and the hierarchical shape of the ontologies. Some experimental results that can be used as examples for further simulations with social practices are described in [35].

6. Related work

There are, of course, many efforts to capture social behavior in software systems. However, most of them try to reproduce some small scale social behavior that is added to the deliberation and action processes of the agents. We try to take a more fundamental perspective in which all these social behaviors

should follow from underlying social agent models. In this perspective understanding social behaviour is the first step towards building systems capable of social deliberation. Social intelligence is defined as an aggregate of different capabilities, including awareness, social beliefs and attitudes, and the ability to change [8,18]. Thus we want to develop a framework in which all these aspects are integrated rather than extend existing systems that already try to incorporate one or two of these aspects in specific contexts.

Because we start with the aim of designing socially intelligent systems it seems logical to take computational cognitive models that are based on human cognitive models as basis. Computational cognitive models, such as ACT-R [2] and SOAR [10] produce intelligent behaviour by employing quantitative measures, which means that different social and individual aspects take the same form in the deliberation process. This makes it difficult to manage, control and vary different socio-cognitive aspects because these cannot easily be isolated in the decision rules. Moreover, once models get larger they lack transparency to link observed behaviour to the implementation, due to a lack of modularity.

As decision-making processes are influenced by individual and social sources [25], we look for architectures that model these influences separately and explicitly. Social influences are usually described in terms of social rules, such as 'obey your parents' or 'mimic the behaviour of your peers'. Individual influences are usually expressed in terms of personal goals or utilities and lead to 'rational' decision rules. The social sciences describe mechanisms or schemas used by humans to link these capabilities (e.g. salience, priming, motivation and regulation), determine how decisions are made and generate complex social behaviour [1,17]. Similar processes occur in human-agent interaction because social signals (like emotional expressions) produced by computational agents are processed by humans in a similar manner as signals produced by humans [37].

Existing architectures used to construct virtual agents and intelligent game characters, such as FAtiMA [12], GRETA [24] or CIGA [36] can achieve fairly realistic behaviours that are computationally efficient for specific domains of application. They do not employ the more general schemas described in social science either because these schemas are computationally inefficient, but more often because they are incomplete and non-operationalizable. Thus the gaps are filled in with pragmatic choices for the specific domain in which the systems are deployed. Given this domain-oriented focus, the implemented behaviours are not easily reusable in applications that require slightly different social aspects.

Recently Kaminka and Dignum et al. [15,22] discussed the many challenges of designing the social behaviour of agents. We agree with these challenges and one can actually see our framework as a first step to solve these challenges. Our main contribution is to provide a concrete architectural design based on social practices as a computational efficient way to deal with the complexity of socially intelligent behavior.

We argue that a novel architecture is needed because generic existing cognitive models in AI are mostly suitable for well-defined problem domains, are platform- or domain-specific, or computationally too complex [10,23,32]. Deliberative agent models, such as BDI [38], have formal logic-grounded semantics, but require extensive computational resources to deal with social contexts, or use game-theoretic rules that are too simple to capture the rich interactions that take place in real-world scenarios [6]. BDI does use different modules for beliefs, desires and intentions. However, these are geared towards individual influences on decision making. These models thus lack an explicit representation for social influences. One can represent all these social influences in the beliefs or goals of an agent, but that leads to the same objection as against the cognitive models; the rules become convoluted and different aspects cannot easily be managed separately.

Finally, social practice theories are, until now, mostly concentrating on how social practices relate, and compete. Recently, a model was proposed to investigate the emergence of social practices [19]

that comprises components and processes which are hypothesised to explain the emergence of social practices, i.e., behaviour that is habitual and similar for a large part of a society. Although we agree that social practices can be seen as independent social constructs that can be studied as first class entities and as such may give new insights in all kinds of social phenomena, we also think that they can be used as input and filter on the individual's deliberation process. From that perspective there is no related work yet in this area.

7. Conclusions

We have shown that intelligent systems will need to be more socially aware of their context in order to take the appropriate action. As a result the deliberation of these systems should keep track of both their social and physical context when they deliberate about actions. If one would add the social context as an extra module (or aspect) to traditional intelligent systems this would overload the deliberation cycle and render it inefficient to cope with real time situations.

As a solution we have indicated several elements that can support the social behavior and that lead to a new type of architecture. Social identities, norms and habits are used as mechanisms to prioritize potential behaviours. They also serve to keep individual behaviour consistent over time. Norms and identities also have a strong social component in that they are shared by individuals in a society. Therefore, they can also be used to predict behaviour of others in known situations. The use of social practices facilitates the combination of social and physical aspects of a situation. Although these social practices might over constrain the combinations they lead to a good and natural reduction of the complexity.

This paper presents a first step towards truly socially intelligent systems. We used some mechanisms from social science and psychology to structure the deliberation process in a way that keeps it efficient in the advent of the increased complexity caused by adding the social context to a situation. We have shown how the use of social practices in the architecture can lead to the fast and slow thinking patterns that are also described in the psychology literature. Using these patterns allows for quick reactions whenever possible while there is a natural entrance for more extensive deliberation as well.

There are many issues that are interesting for future research. We have shown how deliberation can take place when a social practice is recognized and triggered. What will happen if more than one social practice fits the situation? Or if no social practice fits (completely)? More work on these aspects will need a more rigorous (formal) definition of social practices and the interpretation processes involved. Some first steps on how this can be tackled using belief networks is described in [5], but more work is needed. Another very interesting issue is the emergence and evolution of social practices. We already indicated that there is a feedback from the actions performed in a social practice, but how this will exactly adapt the social practices is an interesting question.

Acknowledgements

References

[1]   R. Adolphs, Social cognition and the human brain, *Trends in Cognitive Sciences* **3**(12) (1999), 469–479.
[2]   J.R. Anderson, M. Matessa and C. Lebiere, Act-r: A theory of higher level cognition and its relation to visual attention, *Hum-Comput Interact* **12**(4) (Dec 1997), 439–462.
[3]   G. Andrighetto, G. Governatori, P. Noriega and L.W.N. van der Torre, *Normative Multi-Agent Systems*, Dagstuhl Follow Ups, Germany, (2013).
[4]   M. Argyle, *Social Interactions*, Transaction Publishers, USA, 2009.
[5]   A. Augello, M. Gentile and F. Dignum, Social practices for social driven conversations in serious games, in: *Games and Learning Alliance Conference*, submitted 2015.
[6]   R. Beheshti, Normative agents for real-world scenarios, in: *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-Agent Systems*, AAMAS '14, International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, (2014), 1749–1750.
[7]   P. Berger and T. Luckmann, The social construction of reality: A treatise in the sociology of knowledge, Penguin Books, New York, 1996.
[8]   R.E. Boyatzis, Learning life skills of emotional and social intelligence competencies, *The Oxford Handbook of Lifelong Learning* (2011), 91.
[9]   C.L. Breazeal, *Designing Sociable Robots*, MIT Press, Cambridge, USA, 2004.
[10] A.J. Butt, N.A. Butt, A. Mazhar, Z. Khattak and J.A. Sheikh, The soar of cognitive architectures, in: *CTIT 2013*, IEEE, (2013), 135–142.
[11] J. Dias, S. Mascarenhas and A. Paiva, Fatima modular: Towards an agent architecture with a generic appraisal framework, in: *Proceedings of the International Workshop on Standards for Emotion Modeling* (2011).
[12] J. Dias, S. Mascarenhas and A. Paiva, Fatima modular: Towards an agent architecture with a generic appraisal framework, in: *WS Standards for Emotion Modeling* (2011).
[13] F. Dignum, V. Dignum and C. Jonker, Towards agents for policy making, in: *MABS IX*, Springer, (2009), 141–153.
[14] F. Dignum, R. Prada and G. Hofstede, From autistic to social agents, in: *AAMAS 2014* (May 2014).
[15] F. Dignum, R. Prada and G.J. Hofstede, From autistic to social agents, in: *International conference on Autonomous Agents and Multi-Agent Systems, AAMAS '14*, IFAAMAS, (2014), 1161–1164.
[16] V. Dignum, Mind as a service: Building socially intelligent agents, in: *Coordination, Organisation, Interction and Norms in Multi-Agent System: Proceedings of COIN 2015* (2015).
[17] S.T. Fiske and S.E. Taylor, *Social Cognition: From Brains to Culture*, Sage, (2013).
[18] D. Goleman, Social intelligence: The new science of social relationships, New York, NY: Bantam, 2006.
[19] G. Holtz, Generating social practices, *JASSS* **17**(1) (2014), 17.
[20] M.A. Janssen and W. Jager, Fashions, habits and changing preferences: Simulation of psychological factors affecting market dynamics, *Journal of Economic Psychology* **22** (2001), 745–772.
[21] D. Kahneman, Thinking, fast and slow, *Farrar, Straus and Giroux*, (2011).
[22] G.A. Kaminka, Curing robot autism: A challenge, in: *International conference on Autonomous Agents and Multi-Agent Systems, AAMAS '13*, IFAAMAS, (2013).
[23] J. Laird, *The Soar Cognitive Architecture*, MIT Press, 2012.
[24] M. Mancini and C. Pelachaud, Dynamic behavior qualifiers for conversational agents, in: *Intelligent Virtual Agents*, Springer, (2007), 112–124.
[25] J.G. March, *Primer on Decision Making: How Decisions Happen*, Simon and Schuster, 1994.
[26] D.C. McClelland, *Human Motivation*, Cambridge Univ. Press, 1987.
[27] D.T. Neal, W. Wood, J.S. Labrecque and P. Lally, How do habits guide behavior? perceived and actual triggers of habits in daily life, *Journal of Experimental Social Psychology* **48**(2) (2012), 492–498.
[28] A. Reckwitz, Toward a theory of social practices, *European Journal of Social Theory* **5**(2) (2002), 243–263.
[29] A. Schutz, The phenomenology of the social world, Northwestern University Press, 1967.
[30] J.R. Searle, *The Construction of Social Reality*, Penguin, 1996.
[31] E. Shove, Beyond the abc: Climate change policy and theories of social change, *Environment and Planning A* **42**(6) (2010), 1273.
[32] B.G. Silverman, M. Johns, J. Cornwell and K. O'Brien, Human behavior models for agents in simulators and games: Part i: Enabling science with pmfserv, *Presence: Teleoperators and Virtual Environments* **15**(2) (2006), 139–162.
[33] H. Tajfel, Social identity and intergroup behavior, *Social Science Information* **13**(2) (1974), 65–93.
[34] J. Turner, M. Hogg, P. Oakes, S. Reicher and M. Wetherell, *Rediscovering the Social Group: A Self-Categorization Theory*, New York: Basil Blackwell, (1987).
[35] J. van Diggelen, R.J Beun, F. Dignum, R.M van Eijk and J.-J.Ch. Meyer, Ontology negotiation in heterogeneous multi-agent systems: The anemone system, *Applied Ontology, An Interdisciplinary Journal of Ontological Analysis and Conceptual Modeling* **2**(3–4) (2007), 267–303.

[36] J. van Oijen, L. Vanhée and F. Dignum, Ciga: A middleware for intelligent agents in virtual environments, in: *Agents for Educational Games and Simulations*, Springer, (2012), 22–37.
[37] J. Wagner, F. Lingenfelser, T. Baur, I. Damian, F. Kistler and E. André, The social signal interpretation framework: Multimodal signal processing and recognition in real-time, in: *21st ACM Conf Multimedia*, ACM, (2013), 831–834.
[38] G. Weiss, *Multiagent Systems*, MIT Press, 2013.

Authors' Bios

Frank Dignum is an Associate Professor at Computer Science department of Utrecht University and honorary senior research fellow at the University of Melbourne. He has been working on social aspects of software agents with applications in serious gaming and social simulations. He is well known for his work on norms and agent communication and lately for the combination of agents and games. His latest research focuses on creating new agent architectures to build agents that operate in real-time environments and have to cooperate with humans and other agents. He has organized many workshops and conferences on the topics and given tutorials and invited presentations at most major conferences and summer schools on them.

Virginia Dignum is an Associate Professor at the Faculty of Technology, Policy and Management, Delft University of Technology. She got her PhD in 2004 from the Utrecht University. Previously, she worked for more than 12 years in consultancy and system development in the areas of expert systems and knowledge management. Her research focuses on agent based models of organizations, in particular in the dynamic aspects of organizations, and the applicability of agent organizations to support knowledge creation, sharing and representation in distributed environments, and the interaction between people and intelligent systems in particular the behaviour of hybrid teams. Her work ranges from the engineering of practical applications and simulations to the development of formal theories that integrate agency and organization, and includes a strong design methodology component. In 2006, she was awarded the prestigious Veni grant from NWO (Dutch Organization for Scientific Research) for her work on agent-based organizational frameworks, which includes the OperA framework for analysis, design and simulation of organizational systems. She has organized many international conferences and workshops, and was co-organizer of AAMAS 2005. She has been involved as (S)PC and reviewer in all main conferences and journals in her field of research, and is one of the leading international researchers in Agent Organizations. She is involved in national and EU projects and has more than 180 peer-reviewed publications, including 15 books, yielding a h-index of 23. She is vice-chair of the Benelux AI Association and will be co-chair of the European AI conference in 2016.

Rui Prada is a Senior Researcher at INESC-ID and an Assistant Professor at the Computer Science Department of Instituto Superior Técnico - University of Lisbon (IST), where he teaches courses on User Centred Design, Socially Intelligent Agents and Game Design and Development. He has a degree in Computer Science with a specialization in Artificial Intelligence and a PhD in the same field, both from IST. At INESC-ID he develops his research interests in the fields of Social Intelligent Agents, Virtual Collaborative Environments, User Centred Design and Computer Games and has participated in 14 National and EU research projects, coordinating two of them. He was an invited researcher at the Global Lab, National Institute of Informatics (NII) - Tokyo for 5 months in 2009 and received a NIAS Fellowship for one semester in 2013. He is author and co-author of more than 90 peer-reviewed scientific papers presented at international events and journals. He was a member of the organization committee of

eight national and international conferences, including: ACII'2007, AAMAS'2008, IBERAMIA'2008, ACE'2011 and AAMAS'2015. He is co-founder and currently the president of the Portuguese Society of Sciences of Videogames (SPCVideojogos).

Catholijn M. Jonker is full professor of Interactive Intelligence at the Faculty of Electrical Engineering, Mathematics and Computer Science of the Delft University of Technology. She chaired De Jonge Akademie (Young Academy) of the KNAW (The Royal Netherlands Society of Arts and Sciences) in 2005 and 2006, and she was a member of the same organization from 2005 to 2010. She is a member of the Koninklijke Hollandsche Maarschappij der Wetenschappen and of the Academia Europaea. She was the president of the National Network Female Professors (LNVH) in The Netherlands from 2013 till 2016. Catholijn is ECCAI Fellow since 2015. Her publications address cognitive processes and concepts such as negotiation, teamwork and the dynamics of individual agents and organizations. In all her research lines Catholijn has adopted a value-sensitive approach. In particular, she works towards intelligent agents that can interact with their users in value-conflicting situations when also meta-values no longer solve the situation.