I. SOCIAL INFLUENCE AND THE GENERATION OF JOINT MENTAL ATTITUDES IN MULTI-AGENT SYSTEMS

Pietro Panzarasa, Nicholas R. Jennings

Department of Electronics and Computer Science
University of Southampton, Southampton SO17 1BJ, U.K.
{pp, nrj}@ecs.soton.ac.uk

ABSTRACT
This work examines the social structural and cognitive foundations of joint mental attitudes in complexly differentiated multi-agent systems, and incorporates insights from a variety of disciplines, including mainstream Distributed Artificial Intelligence, sociology, administrative science, social psychology, and organisational perspectives. At the heart of this work lies the understanding of the on-going processes by which socially and cognitively differentiated agents come to be socially and cognitively integrated. Here we claim that such understanding rests on the consideration of the nature of the influence processes that affect socialisation intensity. To this end, we provide a logic-based computational model of social influence and we undertake a set of virtual experiments to investigate whether and to what extent this process, when it is played out in a system of negotiating agents, results in a modification of the agents' mental attitudes and impacts on negotiation performance.

KEYWORDS
Socially motivated mental attitudes; agreement; negotiation

1 INTRODUCTION
Building on both mainstream Distributed Artificial Intelligence research and more traditional organisational and structural analysis, the aim of this paper is to place the study of social influence processes on a more secure and formal footing. To this end, we formalise a model of social influence using a multi-modal, first-order, linear-time, quantified epistemic logic based on possible-worlds semantics (Hintikka, 1962). Building upon this framework, we explore the dynamic implications of social influence over time, and we provide new insights into the attainability of higher-level mentalistic phenomena in complexly differentiated multi-agent systems (Friedkin, 1998; Harrison and Carroll, 1991). In particular, we run a set of virtual experiments to investigate the role of social influence in the negotiation process underpinning the sale of a privately held company through the public offering of shares and, possibly, through the involvement of an active investor. By combining logical modelling with computer simulations, an attempt is made to determine on theoretic grounds the impact of differing social influence patterns on the agents' individual mental attitudes and on the attainability of higher-order mentalistic phenomena.

Our approach is to develop a logic-based architecture for formalising the cognitive agent's mental state in terms of mental attitudes such as beliefs, goals and intentions. Canonical examples of such architectures include IRMA (Bratman et al., 1988) and PRS (Georgeff and Lansky, 1987). As we augment the architecture with the agent's ability to engage in inter-agent social behaviour, we expect that the mechanisms and structures that enable the social nature of agenthood to impact upon the individual agent's mental state and behaviour will be identified and characterised. Against this background, we wish to pave the way for a cognitively and structurally motivated theory of the content, form, function, antecedents, and consequences of higher-order mental models. In this theory, the notion of social influence will serve as the key concept that allows us to gain generalisability over a wide range of joint mental attitudes (e.g. culture, agreement, shared beliefs, joint goals, joint intentions), admit cumulative theory building, and consolidate theories of "group mind"-like constructs with theories of individual cognition. The remainder of the paper is structured as follows. Section 2 provides a formalisation of the model. In Section 3 findings from a set of computer simulations of the model are reported. Finally, Section 4 summarises the major results.

2 THE MODEL

Social influence is here described as a *socio-cognitive* process. In order to emphasise its social and cognitive components, in Panzarasa et al. (2001a) we proposed and defined the term "social mental shaping". More precisely, by social influence we mean the process through which the mere social nature of agents affects, and thereby alters, their mental states from what they would otherwise have been, had the agents not engaged in any form of social behaviour. The typical outcome of social influence is the modification of an agent's mental state, via either the adoption of a new *socially motivated* mental attitude or the modification of an individually motivated one. An agent's adoption of mental attitudes can be seen as partly driven by the social environment in which it is located. In this view, the social environment can play an active causal role, and govern an agent's behaviour in the same way that its individual mental state usually does. In this paper, we elaborate on this idea and we use a formal modelling approach to extend the formalism we developed in Panzarasa et al. (2001a) to a setting in which the process of social influence plays a causal role in generating higher-order joint mentalistic phenomena.

The Language. The formalism used is a many-sorted, first-order, linear-time, multi-modal language which both draws upon and extends the work described in Panzarasa et al. (2001a; 2001b). Informally, the operators $\neg$ (not) and $\vee$ (or) have classical semantics, as does the universal quantifier $\forall$. The remaining classical connectives and existential quantifier are assumed to be introduced as abbreviations, in the obvious way. The semantics of the agent's mental attitudes is a natural extension of the traditional Hintikka possible-worlds semantics (Hintikka, 1962). In addition, we will introduce the key modal operator *Infl* to formalise the process of social influence in its basic forms. Finally, we write $Bel(a_i, \phi)(t_i)$ and $Att(a_i, \phi)(t_i)$ to indicate that agent a_i, at time t_i, maintains, respectively, a belief that ϕ holds, and a generic unspecified mental attitude (e.g. a belief, a goal, an intention) towards ϕ.

Role-Based Social Influence. We formalise the influence that occupying a role can have on the role-player's mental state by expressing the modal operator *Infl* in terms of three parameters, a_i, *Att* and r_i, where a_i is an agent, r_i is a role and *Att* has the meaning outlined above. We have:

$$\forall a_i \; \forall r_i \; \forall t_i \; Infl(Att(a_i, \phi), r_i)(t_i) \equiv \exists t_j \leq t_i \text{ s.t. } In(a_i, r_i) \wedge (In(a_i, r_i) \supset Att(a_i, \phi))(t_j, t_i) \quad (1)$$

Informally, the meaning of (1) is that at time t_i agent a_i is socially influenced by role r_i to have the attitude *Att* towards a state of the world ϕ iff during some interval in the past: (a) agent a_i occupied role r_i; and (b) agent a_i adopted the attitude $Att(a_i, \phi)$ as a consequence of a_i's taking on role r_i. Role-based social influence comes in two forms. Firstly, when there is a mandatory attitude attached to a role, the agent will *automatically* adopt such an attitude by occupying the role. Secondly, when there are optional role-based attitudes, the role-player may *decide* whether to adopt those attitudes or not. This distinction is fundamental to the problem of preventing automatic attitude-adoption whenever an agent occupies a role in a multi-agent system.

Inter-Agent Social Influence. There are a number of ways in which agents can influence one another's mental states. Some of the main modes of inter-agent social influence that are found in organisations are: authority (an agent may be influenced by another to adopt a mental attitude whenever the latter has the power to guide the behaviour of the former); helping disposition (an agent may be influenced by another to adopt a mental attitude simply because it intends to contribute to the welfare of the latter); trust (an agent may be influenced by another to adopt a mental attitude merely on the strength of its confidence in the latter); persuasion (an agent may be influenced to adopt another agent's mental attitude via a process of argumentation); threat (an agent may be threatened to adopt a mental attitude on the basis of future negative interference or denied help). Inter-agent social influence may take place either within or outside social relationships. In what follows, each of these two forms of inter-agent social influence will be dealt with in turn.

A) Inter-Agent Social Influence outside Social Relationships. The influence occurring between a pair of agents outside of any social relationship (e.g. when one of the agents is not aware of the existence of the other) can be formalised as follows:

$$\forall a_i \; \forall t_i \; Infl(Att(a_i, \phi))(t_i) \equiv \exists a_j \; \exists t_j \leq t_i \text{ s.t. } Att(a_j, \phi) \wedge Bel(a_i, Att(a_j, \phi)) \wedge (Bel(a_i, Att(a_j, \phi)) \supset Att(a_i, \phi))(t_j, t_i) \quad (2)$$

Informally, at time t_i, agent a_i is socially influenced to change its mental state by adopting a new mental attitude *Att* or by replacing one of its mental attitudes with *Att*. This form of social influence happens whenever an agent believes that another agent has a mental attitude, and *just for this reason* it adopts that mental attitude. This covers several forms of social influencing, from imitation to spontaneous goal-adoption, from benevolent (not due) adhesion to emulation.

B) Social Relationships and Inter-Agent Social Influence. We now want to formalise how an agent's mental state can be influenced by its being within a social relationship with another agent. This kind of social influence can be formally expressed as follows:

$$\forall a_i, a_j \ \forall t_i \ Infl(Att(a_i, \phi), rel(a_i, a_j))\ (t_i) \equiv$$
$$\exists t_j \leq t_i \text{ s.t. } rel(a_i, a_j) \wedge Att(a_j, \phi) \wedge Bel(a_i, Att(a_j, \phi)) \wedge (Bel(a_i, Att(a_j, \phi)) \supset Att(a_i, \phi))(t_j, t_i) \quad (3)$$

Informally, if an agent a_i, which is in a social relationship *rel* with another agent a_j, believes that a_j has a mental attitude *Att*, and *simply for this reason* it changes its mental state by adopting *Att*, then we can say that a_i is influenced by its being situated within a social relationship with a_j. In general, this form of social influence is based and depends on the agent's decision whether to adopt one of its acquaintance's mental attitudes or not. However, there are also circumstances in which an agent involved in a social relationship with another is *required* to adopt one or more of its acquaintance's mental attitudes (e.g. the boss is by right allowed to order other employees and impose upon them the goal of performing particular activities). In such cases, the agent might well *autonomously* decide whether or not to establish a relationship with another agent but, once established, the relationship may automatically impose a number of mental attitudes on the former's mental state. These are *relationship-based mandatory attitudes*.

Properties of the Model. Having defined social influence as a process whose typical outcome is the modification of an agent's mental state, we can now draw some implications with respect to the impact that social influence may have on the generation of joint mental attitudes. These can be conceived of as higher-order attitudes that emerge from and transcend the mental states of cognitively and socially interconnected individual agents (Panzarasa et al., 2001b). One way to make the mental states of the members of a multi-agent system converge in such a way that a joint mental state will ensue is therefore to affect the individual members' mental states.
Proposition 1. Social influence helps generate joint mental attitudes within systems of cognitively and socially differentiated agents.

Joint cognition and joint behaviour are dynamically intertwined, in that joint mental attitudes affect and are affected by the performance of joint behavioural processes (e.g. negotiation, collaborative decision-making). More specifically, joint behavioural processes find their roots in joint mental attitudes, as these provide the basic coordination principles that help socially and cognitively differentiated agents behave as socially and cognitively interconnected members of a multi-agent system (Carley, 1989). One way to impact upon the joint behaviour of a group of agents is therefore to impact upon their joint mental attitudes by exploiting the link that these have with the individual agents' mental states.
Proposition 2. Social influence helps improve the effectiveness and efficiency of joint behavioural processes.

3 VIRTUAL EXPERIMENT

We will now use the analytical structure introduced in the previous section to undertake a set of virtual experiments. The domain we chose is the sale of a privately held company through the public offering of shares to relatively small investors and through the involvement of an active investor interested in a large block of shares and willing to ultimately play an active role in the firm's management (Mello and Parsons, 1998). In order to sell shares, the seller can adopt a number of selling strategies, each reflecting a different pattern of inter-agent social influence that will be exerted upon the potential investors. We examine whether and to what extent the choice of a selling strategy, and therefore of the corresponding inter-agent social influence pattern, can impact upon the negotiating agents' mental states and, ultimately, upon the generation of an agreement and the performance of the going public process.

Experimental Design. In order to manipulate social influence, we introduce the following selling strategies: (i) selling a controlling block before the Initial Public Offering (IPO) (strategy A); (ii) an IPO followed by the sale of a controlling block (strategy B); (iii) an IPO followed by an informationally enriched sale of a controlling block (strategy B*); (iv) a single-period parallel sale of dispersed shares and a controlling block (strategy C); and (v) a public offering of all shares at a uniform price (strategy D) (Mello and Parsons, 1998). The interpretation of these selling strategies in terms of the pattern of inter-agent social influence they incorporate is as follows. First, each of the above strategies reflects a persuasion-based influence pattern through which the seller tries to impact upon either the active investor's or the small investors' mental states. Second, persuasion is carried out via a process of argumentation which, in turn, can be characterised in terms of the amount of information on which it is based. In the light of this, the social influence pattern reflected by each selling strategy can be operationalised and manipulated through the amount and quality of the information its corresponding argumentation process conveys. More specifically, with strategy A, the seller offers the active investor a price with no additional information, and then approaches the small investors by offering them a price and letting them know whether or not a controlling block has been sold. With strategy B, the seller offers the small investors a price with no additional information, and then approaches the active investor by offering him a price and letting him know whether or not an IPO has occurred. Strategy B* differs from strategy B in that the seller tries to more convincingly exert social influence upon the active investor by providing him with more detailed information on which the price being offered is made contingent (e.g. the price already established with other investors). As a result of this, the active investor is expected to be more motivated to accept the offer received and act accordingly than would be the case if the seller had forwarded only a price without any further information that might justify it. With strategy C, the seller approaches both the active investor and the small investors simultaneously by letting them know about the parallel negotiation processes. Finally, with strategy D, the seller approaches the small investors only with a price and no further information.

Additionally, we introduce a second independent variable – the accuracy of the agents' cognitive representations. Cognitive accuracy will be operationalised through the quality of the agents' beliefs about each other's value for the shares of the firm. In turn, the quality of these beliefs will be measured by the value the agents assign to the standard deviation of the idiosyncratic component of each other's pricing function: the higher the standard deviation, the less accurate the agents' beliefs. Our dependent variable, negotiation performance, is rated by two measures of negotiation effectiveness and efficiency: (i) the revenue raised in the aggregate sell of shares; and (ii) the number of messages that the agents must send to each other before they can reach a final agreement. Our experiments are therefore intended to examine whether and to what extent the choice of a social influence pattern (reflected by the selling strategy) and the accuracy of the agents' cognitive representations can have individual and combined effects on the effectiveness and efficiency of the negotiation process. To ensure that the results reflect the underpinning structure of the model and not merely stochastic processes, we adopt a Monte Carlo approach to average out differences arising from distinct independently specified instantiations of the model parameters (e.g. the firm's book value, the negotiation protocol).

Results. Figure 1 shows the impact that each combination of social influence pattern and degree of cognitive accuracy has upon the revenue raised in the aggregate sell of shares (negotiation effectiveness). In general, for each social influence pattern, the higher the degree of cognitive accuracy, the better the negotiation performance. One way of interpreting these results is that the agents are willing to make more concessions to each other, when their cognitive representations are less accurate. In particular, the lower the degree of cognitive accuracy, the lower the minimum price that the seller is willing to accept from either the small investors or the active investor and, hence, the lower the price that may be agreed upon during negotiation. Thus, the impact that social influence has on negotiation performance becomes weaker as the agents' cognitive representations become less accurate.

There is an interesting exception to this result that allows us to speculate as to whether and to what extent the effectiveness of using more information to exert social influence may depend on how accurately the information is perceived. At high degrees of cognitive accuracy, strategy B does not work well. This result contrasts with the performance of strategy B*, that, at high degrees of cognitive accuracy, gives the highest revenue. How can this phenomenon be explained? The answer is found in the relationship that the information embedded in the social influence pattern has with different degrees of cognitive accuracy. The higher the degree of cognitive accuracy, the more self-confident the agents are in making their proposals and counter-proposals and therefore the less prone they are to be conciliatory and make concessions to one another. This, in turn, might have two consequences: 1) it might well be the case that the seller's price range will never overlap with either the

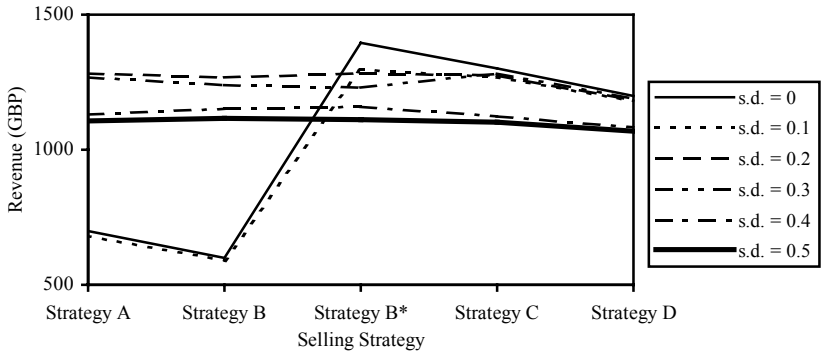

Figure 1.The Impact of Social Influence on Negotiation Effectiveness at Different Levels of Cognitive Accuracy

active investor's or the small investors' price range; or 2) the overlapping area of the agents' price ranges becomes narrower as the spread between the minimum the seller is willing to accept and the maximum the investors are willing to offer becomes lower. In the former case, negotiation fails and no agreement is made. In the latter, it may become more difficult to settle for a price that is acceptable to both parties. In such situations, agreement may be reached in a more effective way if the agents influence one another by exchanging more information so as to support their requests during negotiation. Figure 1 indicates that, at high degrees of cognitive accuracy, the more informative strategy B* is more effective than its variant B as well as strategy C. For example, when the agents' cognitive representations are perfectly accurate (i.e. the standard deviation is zero) strategy B* gives the highest revenue, about 1394.98 (GBP), whereas strategy B gives only 600 (GBP) as no agreement is made between the seller and the active investor. This is because, with strategy B*, the seller influences the active investor by communicating an additional piece of information so as to justify the offer made. In turn, a greater amount of information helps the two agents to reach an agreement. Therefore, these results indicate that the decision as to whether or not to increase the amount and/or quality of information for social influence purposes depends on the agents' cognitive accuracy. In particular, the impact that using more information has on negotiation effectiveness becomes stronger as cognitive accuracy increases.

Another result that is evident from Figure 1 is that, at medium and low degrees of cognitive accuracy, a public offering of all shares at a uniform price is associated with the worst performance. In such situations, a sequential sale involving an active investor (strategy A) is more effective as it allows the seller to obtain more information on which subsequent social influence processes exerted upon the small investors can be made contingent. Furthermore, at medium and low degrees of cognitive accuracy, strategy C mostly dominates strategy D and is dominated by strategy A. The reason for this result is that, with strategy C, the small investors are more effectively influenced to accept the offer than with strategy D where social influence is exerted only by offering a price with no additional information. On the other hand, strategy A dominates strategy C because with the former the seller informs the small investors about the controlling block sold at earlier stage, whereas with the latter the seller only lets the small investors know that a parallel negotiation is simultaneously being carried out with the active investor. The impact of social influence on negotiation between the seller and the active investor follows a similar path. Most of the time, at medium and low levels of cognitive accuracy, strategy B is more effective than strategy C. Again, this result is explained by the additional information conveyed by strategy B and used to persuade the active investor to accept the price offered.

Figure 2 shows the impact that each combination of social influence pattern and degree of cognitive accuracy has upon the number of messages sent (negotiation efficiency). The behaviour of strategies A and B supports the hypothesis that the more accurate the agents' cognitive representations are, the less efficient a social influence pattern is in generating an agreement. This is because the agents are more reluctant to concede and therefore it takes more messages for them to influence one another and reach an agreement.

Throughout all levels of cognitive accuracy the lowest number of messages is associated with strategy D. This result suggests that the number of messages may be positively correlated with the number of agents among which agreement must be made (public offering does not involve the active investor). Strategy A turns out to be less efficient than strategy C: here, the effect of persuasion upon the small investors seems to be more than compensated for by the effect of persuasion upon the active investor. The most interesting result in Figure 2 concerns the role of social influence in negotiation between the seller and the active investor. In general, strategy

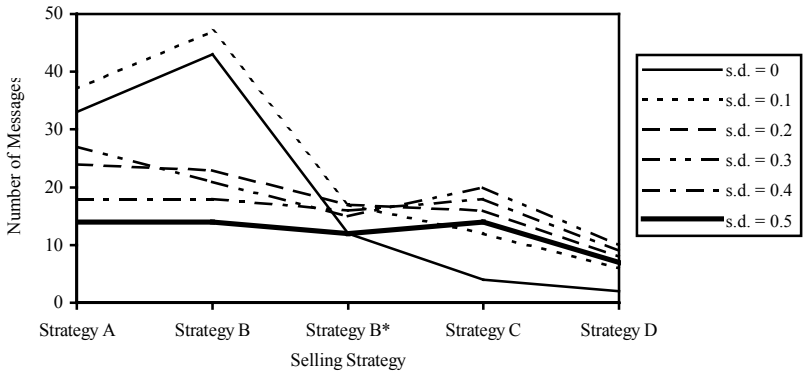

Figure 2.The Impact of Social Influence on Negotiation Efficiency at Different Levels of Cognitive Accuracy

B* requires fewer messages than strategy B. This suggests that the additional piece of information sent together with the main message helps the seller to influence the active investor and reach an agreement more quickly. Additionally, at medium and low levels of cognitive accuracy, strategy B* generates fewer messages than the parallel strategy. This supports our expectation that the more informative a persuasion-based social influence pattern is, the more efficient it is in impacting upon the agents' mental states.

4 CONCLUSIONS

A view of the cognitive agent as a kind of *associative* entity, engaged in an iterated series of social actions and interactions aimed at *completing* its mental state was presented. A key element of this view is the notion of social influence, here described as the process through which the social environment (e.g. roles, other agents within or outside social relationships) complements and augments the agent's bare individual mental attitudes. By combining multi-modal logical modelling with computer simulations, an attempt was made to elucidate on theoretic grounds the systematic implications of the influence process in terms of the generation of joint mental attitudes as a consequence of socialisation pressures.

Herein, using our computational model, a series of virtual experiments were undertaken to explore the individual and combined effect of social influence and cognitive accuracy on the performance of a negotiation process. The domain we chose was the sale of a privately held company through the public offering of shares and through the involvement of an active investor. The results supported our initial expectations that social influence can enhance the performance of joint behavioural processes (Proposition 2) via the generation of joint mental attitudes (Proposition 1). Additionally, the results shed light on what are normally hidden cognitive processes. In particular, cognitive accuracy has an impact on the value of the available information, and this in turn influences the effectiveness of various social influence patterns that are different in terms of the amount of information conveyed.

REFERENCES

Bratman, M. E., Israel, D., and Pollack, M. E. (1988). Plans and resource-bounded practical reasoning. *Computational Intelligence*, 4/4, 349-355.

Carley, K. (1989). The value of cognitive foundations for dynamic social theory. *Journal of Mathematical Sociology*, 14/2-3, 171-208.

Friedkin, N. E. (1998). *A Structural Theory of Social Influence*, Cambridge University Press, Cambridge.

Georgeff M. P., and Lansky, A. L. (1987). Reactive reasoning and planning. *Proceedings of the Sixth National Conference on Artificial Intelligence* (pp. 677-681).

Harrison, J. R., and Carroll, G. R. (1991). Keeping the faith: A model of cultural transmission in formal organizations. *Administrative Science Quarterly*, 36, 552-582.

Hintikka, J. (1962). *Knowledge and Belief*, Cornell University Press.

Mello, A. S., and Parsons, J.E.. (1998). Going public and the ownership structure of the firm. *Journal of Financial Economics*, 49, 79-109.

Panzarasa, P., Jennings, N.R., and Norman, T. (2001a). Social mental shaping: Modelling the impact of sociality on the mental states of autonomous agents. *Computational Intelligence*, forthcoming.

Panzarasa, P., Jennings, N.R., and Norman. T. (2001b). Formalising collaborative decision-making and practical reasoning in multi-agent systems. *Journal of Logic and Computation*, forthcoming.