Integrated Cognition
AAAI Technical Report FS-13-03

A Model of Social Dynamics for Social Intelligent Agents

Samuel Mascarenhas, Nuno Marques, Joana Campos and Ana Paiva
{smascarenhas,nuno.marques,joana.campos}@gaips.inesc-id.pt,ana.paiva@inesc-id.pt
INESC-ID and Instituto Superior Técnico,
Universidade de Lisboa, Portugal

Abstract

In this article we describe a general cognitive model of human social behavior that is meant to increase the social intelligence of autonomous intelligent agents in different contexts. Despite the remarkable improvements that have been made on human-agent interaction, agents still have a limited capacity to be aware of the social reality that is present in the human mind and significantly guides human behavior. The model discussed in this paper is a step toward increasing that capacity significantly. Two different case studies are described in which the proposed model is used to better explain and predict human behavior. The first case study is the well-known Ultimatum game. The second one is a variation of the "Game of Nines" played by children.

Introduction

One of the big dreams of Artificial Intelligence was, and still is, the creation of machines, or agents, that perform tasks and functions that require intelligence. Inspired by the seminal work of Newell and Simon (Newell and Simon 1963), for the past decades research in AI has grown significantly trying to capture computationally the human's ability for high level, formal symbolic reasoning. In that context, intelligence is mostly linked to rational, utility-based behaviour. However, it is undeniable that aspects like emotions or social behaviour are also required for achieving intelligence. So, nowadays, many researchers in Artificial Intelligence have been concerned with issues of believability and sociality in the development of autonomous agents. These aspects turn out to be particularly relevant when developing agents that are meant to interact with humans, understand humans and act in a natural manner with other agents (Reeves and Nass 1996).

Given that many different types of applications can benefit from having agents that are socially-intelligent, their research field is steadily growing. This includes applications for education (Aylett et al. 2005; Paiva and Machado 1998; Lester, Stone, and Stelling 1999; Lester et al. 1999), military training (Johnson and Rickel 1997; Johnson, Vilhjálmsson, and Marsella 2005; Solomon et al. 2008), health (Bickmore and Picard 2005; Johnson et al. 2003; Marsella, Johnson, and LaBore 2000), e-commerce (Cassell et al. 1999), entertainment (Becker-Asano and Wachsmuth 2010; Prada and Paiva 2009; Cassell et al. 2000; Tomlinson and Blumberg 2002; Rousseau and Hayes-Roth 1998), among others. In the last decades, notable improvements have been made in the creation of such agents, particularly in their emotional capabilities(Dias and Paiva 2005; Gratch and Marsella 2004; Marsella, Gratch, and Petta 2010; Becker-Asano and Wachsmuth 2010; Gebhard 2005; Rodrigues et al. 2009) and conversational skills (Traum and Rickel 2002; Cassell et al. 1994; Gratch et al. 2007). Moreover, the inclusion of personality factors (Dias and Paiva 2005; McRorie et al. 2009; Gebhard 2005; Egges, Kshirsagar, and Magnenat-Thalmann 2004; Rousseau and Hayes-Roth 1998) have facilitated the creation of agents capable of expressing different individual identities, enhancing their believability in a social context. However, despite the improvements made, the capacity for agents to simulate human social interaction is still strongly limited as many challenges have yet to addressed. One such challenge is the consideration of behavioural patterns that occur at a social level of analysis. Even though we like to view ourselves as strongly independent individuals, the reality is that humans only managed to survive by living in societal groups where resources had to be shared and labour had to be divided. To solve such problems, among many others, collaboration was essential. As a result, unwritten shared rules and beliefs began to emerge in the earliest tribes. The ones that were kept and passed on eventually became an integral part of the group's shared knowledge.

Humans have the unique capability of creating a social reality that is symbolic and exists only in the human mind. Even though this reality cannot be directly observed, its influence in behaviour can be as strong as the physical reality and often times stronger. This is evident when people go against their survival instincts to defend an ideal or when they intentionally starve in order to attain an idealized notion of beauty. Such type of behaviour is difficult to explain in agent models that only go so far as to include individual factors, such as personality or biological drives. For agents to interact with humans in a natural manner, they need to be able to consider and be influenced by, at least partially, the same type of social reality that drives human behaviour.

With the goal of better capturing this social reality that exists in the human mind, in this article we describe the

Social Importance Dynamics (SID) model, which is a cognitive model of social behaviour that is grounded on the status-power theory proposed by Kemper (Kemper 2011). The model augments a typical BDI agent with a set of social dynamics that revolve around the concept of Social Importance (SI). This concept is meant to represent the behavioural dimension of status as proposed in Kemper's theory. In this theory, the author divides human interaction in two types: *technical* and *relational*. Technical activity corresponds to collaborative behaviour whose primary function is to achieve an instrumental goal such as building a house or winning a game of soccer. However, during technical activity and outside of it, humans are constantly engaging in relational behaviour. For instance, during a soccer match, team mates will encourage one another and will celebrate goals together. The aim of the SID model is to model the motivation behind such type of behaviour.

The outline of this paper can be described as follows. In the following section we describe the SID model, briefly discussing the theory in which the model is based and also presenting how the model has an impact on the agent's perception, deliberation and planning. Afterwards, we discuss the application of the model in a well-known scenario, the Ultimatum Game, which has been extensively used to study human decision-making. Finally, we explore the use of the SID model in a more complex negotiation context, by using it to conduct a preliminary analysis of children playing a variation of the "Game of Nines", a game designed to evoke conflict situations.

Social Importance Dynamics Model

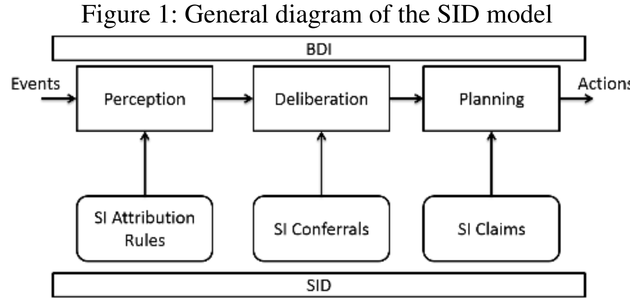
According to the status-power theory (Kemper 2011), the motivation behind all relational behaviour can ultimately be explained in terms of two dimensions, status and power. Status corresponds to our social standing in the eyes of another person, which is then reflected on that person's will to please and act according to our interests instead of their own. Conversely, power corresponds to our ability to coerce another person to comply with our wishes against their will.

The two dimensions have quite different dynamics associated to them. For instance, power can be obtained by holding a weapon against a person whereas status originates from positive social interaction. As our primary concern is to improve human-agent interaction, we started by only addressing the notion of status in the SID model. Thus, the model makes the assumption that neither agents nor users will attempt to coerce or manipulate others. As such, the inclusion of power is left as future work. It is also important to distinguish the concept of status as proposed in Kemper's theory from the common use of the word to refer to hierarchical roles in a social group such as "boss", "teacher", or "mother". To explain the distinction, consider the following examples. A boss can have a very high or very low status depending on how well he is respected and admired by his employees but, in both cases, his power is high given that he can fire his staff if they fail to meet his expectations. In the case of a typical mother-son relationship, although powerless, the son has a very high status, given that the mother is willing to do everything she can for her son. To avoid confusion between the two meanings, our proposed model refers to its implementation of status as social importance (SI).

The SI concept is modelled as a numeric variable that is associated to a relationship between two agents. More precisely, $SI(A, B)$, is a scalar that represents how much is agent A willing to act in the interest of agent B, which might be different than what agent B is willing to do for A, i.e. $SI(A, B)$ is not the same as $SI(B, A)$. But how is it determined? The amount of SI that a person attributes to another is determined by many different sources including friendship, family ties, group membership, behavioral conduct, admirable qualities, trust, reputation, among others. In this sense, SI constitutes an integrative measure of social factors in a similar manner that the concept of $utility$ represents an integrative measure of motivational factors. The main advantage of having this broad social concept is that it allows us to abstract from the particularities of the different relational factors and model a common underlying process. For instance, the relationship between father and son is clearly distinct from a lovers relationship in terms of appropriate behaviours. However, in both cases, the extent that one is willing to act in the interest of the other is similarly high, which is what the concept of SI represents.

According to the status-power theory (Kemper 2011), our actions towards others can be perceived as either claims or conferrals, from a relational standpoint. A claim occurs whenever someone makes a request to another. The bigger the request, the higher the claim. After a claim is performed, a conferral act is expected in response, with different actions conferring different amounts of SI. For illustration purposes, consider a professor that is about to give an invited talk on a conference but is not able to find the room. The professor can approach a random person in the lobby and ask for directions and this would be perceived as a relatively low claim. Then, based on the SI that the other person attributes to the professor, a decision is made on how much will the person confer. The following three actions, exemplify three possible conferrals that vary on the amount of SI conferred (from lowest to highest): (1) come up with an excuse to avoid having to explain the directions, (2) take the time to explain the directions or (3) personally accompany the professor to the room. A person who highly admires the professor will be inclined to choose the third option. Oppositely, a person who has very little respect for the professor is more likely to choose the first option instead.

Figure 1: General diagram of the SID model



Based on the aforementioned notions, the SID model is based on the following three components: (1) $SIAttributionRules$, (2) $SIConferrals$, and (3) $SIClaims$. As illustrated in Figure 1, each of these components has an impact on a different process of a general BDI architecture. The $SIAttributionRules$ are used to determine how much SI should be attributed to another agent. These are formally defined as a <T,A,V> tuple where: T corresponds to the target of the rule, A is a list of conditions that specify when the rule is activated, and V specifies the amount of SI the target of the rule gains/loses while the rule is active. The impact of these rules on the agent's perception corresponds to having the agent actively checking which rules can be activated for every other agents, every time they update their beliefs. Additionally, agents also infer how much SI do they have in the perspective of others, assuming that these others adopt an identical set of rules.

The $SIConferrals$ have an impact on the agent's deliberation by automatically generating social goals to confer status to others, which according to (Kemper 2011), corresponds to an intrinsic human motivation that explains why people care so much about greeting rituals and other activities that achieve no instrumental function outside of the social world. Formally, an $SIConferral$ is defined by the tuple <T,C,A,V> where: T is the target agent to whom the conferral is directed, C is a set of activation conditions that dictate the context in which the conferral is relevant, A is the name of the conferral action that needs to be performed by the agent and finally V defines the amount conferred by the conferral.

If there is a target agent that has the same or more SI than the amount conferred by an $SIConferral$ and the context conditions of the conferral are all verified, a goal to execute the conferral act is automatically generated. The utility of this goal is linearly proportional to the amount of the SI the act confers. Consequently, if an agent has two or more conferral goals activated that have the same target, the conferral selected will be the one that confers the most, without exceeding the SI of the target.

The last component of the SID model is the $SIClaims$. These are represented as a <A,V> pair where: A is the the name of the action that symbolizes a claim and V is the amount of SI that is claimed by the action. They affect the planning process in the following way. While the agent is planning a course of actions to achieve its intention, it will check if any action of the current plan corresponds to an $SIClaim$. When this happens, the agent will check its assumption of how much SI it has on the perspective of the target agent. If the agent infers that its SI is not enough, then the action is removed and the agent searches for an alternative plan that also achieves the same goal. This implies that agents will never try to claim more SI than what they believe they have. This is a simplification we plan to address in future work, as sometimes this form of behaviour is performed by humans in an intentional manner.

To clarify how these components can be used to model social behaviour, the previous situation of the lost professor will be used as an example. This particular situation can be modelled with the following $SIAttributionRules$:

- <T = x, isPerson(x) = True, V = +30>
- <T = x, isAdmired(x) = True, V = +10>

Only one claim is needed, namely:

- <A = ask-directions, V = 30>.

Finally, in response to this claim, consider that the possible conferrals are:

- <T = x, C = performed(x,ask-directions), A = give-excuse,V = 20>
- <T = x, C = performed(x,ask-directions), A = give-directions,V = 30>
- <T = x, C = performed(x,ask-directions), A = accompany-to-destination, V = 40>

When the professor and the random person perceive each other, they will attribute some amount of SI to one another. Based on the first rule, the professor will attribute to the person an SI of 30. However, the random person will attribute to the professor an SI of 40 or 30, depending if that person admires the professor or not (second attribution rule). The professor, aware that he is lost, forms an intention to obtain directions. When planning to achieve this intention, he plans to perform the action of asking directions to the random person. Because asking directions is a claim of 30, the professor will consider that this is a socially appropriate action to perform to the person and does so. In response to the claim made by the professor, the person activates all the three possible conferrals and selects the highest one that does not exceed the SI attributed to the professor. This means that if the person admires the professor, then the third conferral is chosen and the person will accompany the professor. If not, then the person will select the second conferral and will just explain the directions.

Ultimatum Game

Originally, the SID model was used to facilitate the creation of agents with different cultural biases in their behaviour. This involved the implementation of the model in FAtiMA (Dias, Mascarenhas, and Paiva 2011), an existing BDI architecture for embodied agents (Dias, Mascarenhas, and Paiva 2011). The resulting architecture was then applied to create an intercultural training tool where users can socially interact with agents that have different cultural profiles (Mascarenhas et al. 2013).

Although the SID model was originally meant to augment the social intelligence of a typical BDI agent, we are also interested in exploring its use in more formalized scenarios such as the Ultimatum Game. The motivation for doing so comes from the fact that this scenario has been studied extensively and it is a good example of how human behaviour deviates from optimal game-theoretic predictions. As such, we are interested in determining if the social dynamics proposed in the SID model are able to more closely reflect how humans behave when playing this game.

In the Ultimatum Game, first developed by (Güth, Schmittberger, and Schwarze 1982), two players must divide a given sum of money between them. One of the players, the

Allocator, first chooses how that amount should be be split. The other player, the Recipient, can then either accept or reject the offer. If he or she accepts, then each player receives the agreed amount. Otherwise, both get nothing. In its standard form, the game is one shot.

Despite its very simple structure, the Ultimatum Game can be found in many every day situations (eg. bargaining), making it a specially attractive tool in studies of social decision making (Handgraaf, Van Dijk, and De Cremer 2003). From a game theoretic perspective, the Allocator should propose the minimum possible amount, which would always be accepted by the Recipient (getting something is better than getting nothing). However, empirical data shows that the fair offer (a 50/50 split) is the most common outcome (Hoffman, McCabe, and Smith 1996). Thus, it seems as people take other factors into consideration besides their personal expected return when playing the Ultimatum Game.

The aforementioned factors appear to be of socio-cultural nature. For instance, (Henrich et al. 2001) found that the mean Allocator offers and rejection rates may depend on the nationality of the participants, and (Bolton and Zwick 1995) showed that a lower social distance between the participants result in higher offers and higher acceptance rates. These findings are reinforced by neuroscience research linking the activation of specific areas of the brain with rewards of social nature, such as equality and reciprocity (Sanfey et al. 2003; Sanfey 2007).

In the context of our work, this effects may be represented by employing the SID model in an Ultimatum Game Scenario. An Allocator agent (A) determines his offer, p, using the amount of SI he attributes to the Recipient, such that:

$$p = SI(A, R) * k_1$$

In the case of a Recipient agent (R), the decision whether to accept or reject depends on the agent's expectation over the offer he will receive. This expectation is a scalar $e \in [0, 1]$, which is linearly proportional to the Receiver's estimation of the SI attributed to him by the Allocator, represented by $SI_R(A, R)$. Thus:

$$e = SI_R(A, R) * k_2$$

This expectation can then be used to determine the Recipient's response:

$$f(p) = \begin{cases} accept & \text{if } p - e > \varepsilon \\ reject & \text{otherwise} \end{cases}$$

where ε is the error margin associated with the expectation.

We are preparing a user study where human opponents play against agents, so we can gather data on how people perceive the emerging behaviour. This will imply first manipulating the SI attributed to the agent, for example via in-group/out-group membership information. Then we can check if the agent predictions match the way the human plays, and thus if it reacts realistically.

In addition, the social dynamics of the game will also be explored by varying the amount of SI and, consequently, the amounts, as a result of making unfair offers or rejecting fair offers. We will analyse at what level does the "social fairness" factor, represented by the SI, interacts with the self-interest drive, and balance their effects in order to achieve more human-like behaviours.

We should note, however, that other factors, such as Kemper's notion of power and a self-interest drive are still absent and their inclusion is seen as future work. However, this preliminary approach will likely help us in refining the model.

Game of Nines

Although the repeated ultimatum game is a good game to explore the impact of social relations in decision making, we also wanted to investigate how social importance is linked to the way we negotiate. To do that, and thus investigating the types of interactions that emerge and how the social importance is linked to negotiation, we have designed an experiment with the "Game of Nines". The "Game of Nines" is a mixed-motive bargaining game and it was firstly used by Kelley et al. (Kelley, Beckman, and Fischer 1967). The two-person game requires that the players start a negotiation process to divide a joint reward between themselves.

Each player holds cards from one to eight in their hands. At each trial, the negotiators must agree on the cards they play such that their sum is nine or less. Furthermore, for each bargaining problem each negotiator is privately assigned to a *minimum necessary share* (MNS). The negotiator must play a card above the MNS if he wants the agreement to be profitable to him (e.g. if a player has a MNS equals to 3 and plays a 6 he will get 3 as a reward). If no agreement is reached both players get zero. A person does not have direct knowledge of the other's MNS value so the negotiation occurs under incomplete information. Adding to that, the negotiation must occur in a limited amount of time.

These conditions create a mixed-motive relationship in the sense that it is of their mutual interest to agree upon a division that gives them a share larger than their MNS in the minimum use of time. On the other hand, it is of each person's individual interest to guarantee to have the largest share in the division at each trial.

This bargaining game creates an interesting setting, where the negotiators face dilemmas concerning their goals and forms of communication. We used this game to investigate forms of social conflict and the dynamics behind the social interactions (thus linked to the SID model we have been investigating). In particular, we wanted to understand how children perceive situations of disadvantage and conflicting interests and which strategies they employ in different relational contexts.

We have designed an experiment, involving 11 pairs of children aged 10 to 12 years-old. Opt-out consent forms were provided to all parents or guardians of those children. All games were video and audio recorded.

Method

The basic paradigm of the "Game of Nines" was used in our experiment, a constant-sum game with incomplete information. Yet, as the experiment subjects were children we eliminated the time constraint[1].

The experiment consisted in 5 rounds. At the beginning of each round the subjects took from an envelop a number that represents their minimum necessary share (MNS). The players were told to never show that number to the opponent during the trial and never agree on a value below that number. At the end of the round the participants have to show their MNS value to the other.

On each bargaining trial the players have to jointly agree on a possible contract. Each contract corresponds to a card that is going to be played by player A and a card played by player B, so that their sum does not exceed 9. For example, if player A plays the card 7 a possible contract is player B to play card 2. The interests of the parties are always directly opposed. Besides the cards ranging from one to eight that each player holds, they also have a card that allows them to give up if they feel they are not able to achieve a viable agreement.

The MNS values throughout the rounds are in table 1. The combinations of values at each trial, although different, they were chosen to describe the same situation, i.e. the distribution of viable contracts is the same throughout the trials. An exception occurs at the 4th round, which given the MNS values there are not any mutually profitable agreement.

	Player A	**Player B**
Round 1	2	2
Round 2	1	3
Round 3	3	1
Round 4	5	5
Round 5	2	2

Table 1: MNS values for each player per round

To motivate the participants to do well, we told the players that the winner of the game would win a prize. The winner is the player who accumulates more points throughout the rounds. This is a constant-sum game, so the points at round 3 are the sum of points gained at the previous rounds. The number of points a players gets at each round is determined by the card played minus his minimum at that round. After the explanation of the rules, the participants were "walked through" two rounds of the game to learn its mechanics. Following that they were left alone to play the game.

Analysis of Negotiation Behaviours

This experimental setting incorporates motivational, cognitive, perceptual and decision making processes. As a possible way to understand its dynamics we used the Social Importance Dynamics Model to examine how these processes unfold.

For illustration purposes, consider the following interaction between two girls (see Figure 2) during the first round of the game (the MNS values are in Table 1). Between brackets is the analysis of the interaction according to the SID model. The acronyms used in the excerpt are described in Table 2.

[1]We verified in pilot sessions that this factor was making them not to pay attention to what was happening in the game.

CLM-H	High Claim
CLM-L	Low Claim
CONF-L	Conferral lower than expected by the claim
CONF-EXP	Conferral expected in response to the claim
CONF-SP	Spontaneous conferral with no explicit claim

Table 2: Acronyms used in the analysis with the SID model.

```
Player B: So...
Player A: OK. You play first. (CONF-SP)
Player B: I play a ....<hesitation>.
     But first we have to see our
     minimums don't we?
A: First we have to take the
  minimum.
< They take the envelop for that round
and save their card>
B: I play the card "6". (CLM-H)
  Do you agree or not?
A: I agree. (CONF-EXP)
B: Clan I play the card?  (CLM-H)
  <She places the card on the table>
A: Yes. (CONF-EXP)
A: I play... <hesitation>
B: <Coff>. huh?...
A: <Look at the cards>
B: Are you going to play?
  You play the card "3". Yes? (CLM-H)
A: Wait. Wait. Wait. (CONF-L)
B: I'm confused....
A: Wait. I play the card 6. (CLM-H)
B: What? (CONF-L)
A: I play 6. (CLM-H)
B: Huh? (CONF-L)
A: I play 6. (CLM-H)
B: But that doesn't work. 6+6 is 12.
(CONF-L)
A: That's right...
B: The maximum is 9.
  Now I'm a little bit confused.
A: OK. I play 4.  (CLM-L)
B: That doesn't work either. (CONF-L)
A: 7,8,9.
B: It can't be. Gives 10.
A: I play 3. (CONF-EXP)
A: <laugh>  (CONF-SP)
B: <laugh> (CONF-SP)
```

This example suggests that Player A attributes more SI to Player B than the opposite. The first sign of this occurs at the start of the interaction where Player A proposes that Player B should be the first to play. Then, Player B makes a really high claim (she proposes to play 6) and Player A accepts it straight away. But at some point in the interaction, while

Player A is deciding the card she will play, she realizes that Player B is being a bit greedy. But instead of directly asking Player B to play a lower card, she says that she will also play 6, possibly expecting that player B goes back in her initial proposal. Player B quickly expresses her surprise and confusion regarding Player A's attempt. She then continuously rejects all the offers that Player A makes, until, finally, Player A accepts to play 3. Interestingly, they both laugh at the end. This exchange of conferrals possibly serves the purpose of re-establishing any loss of SI that occurred during the round.

It should be emphasized that this is a tentative explanation for the observed phenomena. Other factors may be influencing the player's behaviour, such as a limited understanding of game rules. A more in-depth analysis is thus warranted in order to isolate the effects of Social Importance in this negotiation game.

Figure 2: Two girls playing the *Game of Nines*

We are now analyzing and annotating the interactions of all the eleven dyads using the SID model, not only to inform us about the social dynamics that unfolds during the games, but also, as a way to inform the model itself and thus, further refine it.

Conclusion

In this paper we argued about the importance of increasing the social intelligence of autonomous agents by making them more aware of the relational aspects of human behaviour. More than physically instrumental, our actions carry out symbolic meanings and their performance signifies our relation with others. In this paper, we described a cognitive model of social behaviour that takes into account these relational meanings and endows agents with a set of social dynamics concerning them. More precisely, the proposed SID model affects the perception of agents by making them aware of their relational standing with other agents, which is treated as a numeric variable, named SI. The model affects deliberation by automatically creating conferral goals to convey how much SI do agents attribute to others. The model also affects planning by making agents avoiding actions that would claim more SI than what other agents attribute to them.

For the sake of exploring how the proposed SID model can be used to simulate human decision-making, we discussed its implementation in the Ultimatum Game. This is a well-researched game that is played by two players in which one of them makes a monetary offer to the other and the other can either accept or reject the offer made. Previous studies have shown that people are significantly influenced by socio-cultural factors when playing this game and we are attempting to represent such factors with the SID model. We plan to conduct different manipulations on the SI that either the player or the agent attributes to one another and observe how it affects their strategy. We are also interested in seeing how the SID model can be useful to model more complex games, that involve active negotiation between the players. For this purpose, we have recorded 11 pairs of children playing a variation of the "Game of Nines", a game in which two players must bargain a division between nine points with one another across multiple rounds. We are currently testing the predictive power of the SID model by using it to analyse and describe the observed interactions between the children playing the game as it was exemplified in this article.

Acknowledgements

This work was partially supported by national funds through FCT - Fundação para a Ciência e a Tecnologia, under project PEst-OE/EEI/LA0021/2013. It was also partially supported by two FCT scholarships (SFRH BD /62174/2009), (SFRH/BD/75342/2010) and by the European Community (EC), being partially funded by the ECUTE project (ICT-5-4.2 257666), the EMOTE project (FP7 ICT-317923) and the SIREN project (FP7 ICT 258453). The authors are solely responsible for the content of this publication. It does not represent the opinion of the EC or the FCT, which are not responsible for any use that might be made of data appearing therein.

References

Aylett, R.; Louchart, S.; Dias, J.; Paiva, A.; and Vala, M. 2005. Fearnot! - an experiment in emergent narrative. In Panayiotopoulos, T.; Gratch, J.; Aylett, R.; Ballin, D.; Olivier, P.; and Rist, T., eds., *Intelligent Virtual Agents*, volume 3661 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg. 305–316.

Becker-Asano, C., and Wachsmuth, I. 2010. Affective computing with primary and secondary emotions in a virtual human. *Autonomous Agents and Multi-Agent Systems* 20:32–49.

Bickmore, T. W., and Picard, R. W. 2005. Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction* 12(2):293–327.

Bolton, G. E., and Zwick, R. 1995. Anonymity versus punishment in ultimatum bargaining. *Games and Economic Behavior* 10(1):95–121.

Cassell, J.; Pelachaud, C.; Badler, N.; Steedman, M.; Achorn, B.; Becket, T.; Douville, B.; Prevost, S.; and Stone,

M. 1994. Animated conversation: rule-based generation of facial expression, gesture & spoken intonation for multiple conversational agents. In *Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, SIGGRAPH '94, 413–420. New York, NY, USA: ACM.

Cassell, J.; Bickmore, T.; Billinghurst, M.; Campbell, L.; Chang, K.; Vilhjálmsson, H.; and Yan, H. 1999. Embodiment in conversational interfaces: Rea. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, CHI '99, 520–527. New York, NY, USA: ACM.

Cassell, J.; Ananny, M.; Basu, A.; Bickmore, T.; Chong, P.; Mellis, D.; Ryokai, K.; Smith, J.; Vilhjálmsson, H.; and Yan, H. 2000. Shared reality: physical collaboration with a virtual peer. In *CHI '00 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '00, 259–260. New York, NY, USA: ACM.

Dias, J., and Paiva, A. 2005. Feeling and reasoning: a computational model for emotional agents. In *Proceedings of 12th Portuguese Conference on Artificial Intelligence, EPIA 2005*, 127–140. Springer.

Dias, J.; Mascarenhas, S.; and Paiva, A. 2011. Fatima modular: Towards and agent architecture with a generic appraisal framework. In *Proceedings of the Workshop on Standards in Emotion Modelling*. Leiden: Lorentz Center.

Egges, A.; Kshirsagar, S.; and Magnenat-Thalmann, N. 2004. Generic personality and emotion simulation for conversational agents. *Computer Animation and Virtual Worlds* 15(1):1–13.

Gebhard, P. 2005. Alma: a layered model of affect. In *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*, AAMAS '05, 29–36. New York, NY, USA: ACM.

Gratch, J., and Marsella, S. 2004. A domain-independent framework for modeling emotion. *Cognitive Systems Research* 5(4):269–306.

Gratch, J.; Wang, N.; Gerten, J.; Fast, E.; and Duffy, R. 2007. Creating rapport with virtual agents. *Intelligent Virtual Agents* 125–138.

Güth, W.; Schmittberger, R.; and Schwarze, B. 1982. An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization* 3(4):367–388.

Handgraaf, M. J.; Van Dijk, E.; and De Cremer, D. 2003. Social utility in ultimatum bargaining. *Social Justice Research* 16(3):263–283.

Henrich, J.; Boyd, R.; Bowles, S.; Camerer, C.; Fehr, E.; Gintis, H.; and McElreath, R. 2001. In search of homo economicus: behavioral experiments in 15 small-scale societies. *The American Economic Review* 91(2):73–78.

Hoffman, E.; McCabe, K.; and Smith, V. 1996. On expectations and the monetary stakes in ultimatum games. *International Journal of Game Theory* 25:289–301.

Johnson, W. L., and Rickel, J. 1997. Steve: An animated pedagogical agent for procedural training in virtual environments. *ACM SIGART Bulletin* 8(1-4):16–21.

Johnson, W. L.; Shaw, E.; Marshall, A.; and LaBore, C. 2003. Evolution of user interaction: the case of agent adele. In *Proceedings of the 8th international conference on Intelligent user interfaces*, IUI '03, 93–100. New York, NY, USA: ACM.

Johnson, W. L.; Vilhjálmsson, H. H.; and Marsella, S. 2005. Serious games for language learning: How much game, how much ai? In Looi, C.-K.; McCalla, G. I.; Bredeweg, B.; and Breuker, J., eds., *AIED*, volume 125 of *Frontiers in Artificial Intelligence and Applications*, 306–313. IOS Press.

Kelley, H. H.; Beckman, L. L.; and Fischer, C. S. 1967. Negotiating the division of a reward under incomplete information. *Journal of Experimental Social Psychology* 3(4):361 – 398.

Kemper, T. 2011. *Status, power and ritual interaction: a relational reading of Durkheim, Goffman, and Collins*. England: Ashgate Publishing Limited.

Lester, J. C.; Voerman, J. L.; Towns, S. G.; and Callaway, C. B. 1999. Deictic believability: Coordinating gesture, locomotion, and speech in lifelike pedagogical agents. *Applied Artificial Intelligence* 13:383–414.

Lester, J. C.; Stone, B. A.; and Stelling, G. D. 1999. Lifelike pedagogical agents for mixed-initiative problem solving in constructivist learning environments. *User Modeling and User-Adapted Interaction* 9(1-2):1–44.

Marsella, S.; Gratch, J.; and Petta, P. 2010. *Computational Models of Emotion.* Oxford University Press.

Marsella, S. C.; Johnson, W. L.; and LaBore, C. 2000. Interactive pedagogical drama. In *Proceedings of the fourth international conference on Autonomous agents*, AGENTS '00, 301–308. New York, NY, USA: ACM.

Mascarenhas, S.; Prada, R.; Paiva, A.; and Hofstede, G. J. 2013. Social importance dynamics: A model for culturally-adaptive agents. In Aylett, R.; Krenn, B.; Pelachaud, C.; and Shimodaira, H., eds., *Intelligent Virtual Agents*, volume 8108 of *Lecture Notes in Computer Science*, 325–338. Springer.

McRorie, M.; Sneddon, I.; Sevin, E.; Bevacqua, E.; and Pelachaud, C. 2009. A model of personality and emotional traits. In Ruttkay, Z.; Kipp, M.; Nijholt, A.; and Vilhjalmsson, H., eds., *Intelligent Virtual Agents*, volume 5773 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg. 27–33.

Newell, A., and Simon, H. 1963. GPS: a program that simulates human thought. In Feigenbaum, E. A., and Feldman, J., eds., *Computers and Thought*. McGraw-Hill. 279–293.

Paiva, A., and Machado, I. 1998. Vincent, an autonomous pedagogical agent for on-the-job training. In Goettl, B.; Halff, H.; Redfield, C.; and Shute, V., eds., *Intelligent Tutoring Systems*, volume 1452 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg. 584–593.

Prada, R., and Paiva, A. 2009. Teaming up humans with autonomous synthetic characters. *Artificial Intelligence* 173(1):80–103.

Reeves, B., and Nass, C. 1996. *The Media Equation - How*

People Treat Computers, Television, and New Media Like Real People and Places. Cambridge University Press.

Rodrigues, S. H.; Mascarenhas, S. F.; Dias, J.; and Paiva, A. 2009. i can feel it too!: Emergent empathic reactions between synthetic characters. In *Proc. of the 3rd International Conference on Affective Computing and Intelligent Interaction (ACII) 2009*, 1–7. IEEE.

Rousseau, D., and Hayes-Roth, B. 1998. A social-psychological model for synthetic actors. In *Proceedings of the second international conference on Autonomous agents*, 165–172. ACM.

Sanfey, A. G.; Rilling, J. K.; Aronson, J. A.; Nystrom, L. E.; and Cohen, J. D. 2003. The neural basis of economic decision-making in the ultimatum game. *Science* 300(5626):1755–1758.

Sanfey, A. G. 2007. Social decision-making: insights from game theory and neuroscience. *Science* 318(5850):598–602.

Solomon, S.; van Lent, M.; Core, M.; Carpenter, P.; and Rosenberg, M. 2008. A language for modeling cultural norms, biases and stereotypes for human behavior models. In *BRIMS*.

Tomlinson, B., and Blumberg, B. 2002. Social synthetic characters. *Computer Graphics* 26(2).

Traum, D., and Rickel, J. 2002. Embodied agents for multiparty dialogue in immersive virtual worlds. In *Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 2*, AAMAS '02, 766–773. New York, NY, USA: ACM.