In D. Reitter & F. E. Ritter (Eds.), Proceedings of the 14th International Conference on Cognitive Modeling (ICCM 2016). University Park, PA: Penn State.

A Metacognitive Agent for Training Negotiation Skills

Christopher A. Stevens (c.a.stevens@rug.nl), Harmen de Weerd (h.a.de.weerd@rug.nl),
Fokie Cnossen (f.cnossen@rug.nl), Niels A. Taatgen (n.a.taatgen@rug.nl)
Department of Artificial Intelligence, Nijenborgh 9
9747AG Groningen, The Netherlands

Abstract

Training negotiators remains a difficult and expensive proposition. Negotiators require complex cognitive skills such as theory-of-mind to be successful, but these skills can be difficult to train and measure. Here we present an agent designed to model theory-of-mind for learners and serve as a practice partner for complex negotiations. This agent employs instance-based learning to make decisions about its own actions and to reflect on the behavior of the opponent. This reflection process is used to provide a source of explicit feedback on the opponent's strategy and behavior. In this paper we present evidence that the model is a plausible opponent for students learning negotiation. It is expected that practicing with this agent will improve theory-of-mind abilities in learners and, in turn, improve negotiation performance.

Keywords: Metacognition, Negotiation, Theory-of-mind, Autonomous Agents

Training Negotiation with Artificial Agents

Negotiation is a complex human activity that permeates many aspects of business, politics, and even daily life. There is a myriad of possible negotiation settings, and in most of these settings, there are many possible outcomes. Training negotiators is difficult because the optimal strategy often depends both on the negotiation setting and on the strategy used by one's negotiation partner (Fisher & Ury, 1981; Raiffa, 1982). For instance, an aggressive, unyielding strategy (a.k.a. the "Boulware" strategy (Cross, 1977)) may work very well when the partner is agreeable or flexible. However, the same strategy will lead to stalemate against an aggressive opponent and could harm the potential for future negotiations with fair-minded opponents (Tinsley, O'Connor, & Sullivan, 2002) Therefore, a good negotiator may benefit from using theory-of-mind to infer an opponent's preferences, diagnose an opponent's strategy, and select an appropriate counter strategy. One promising emerging training technique is practicing with artificial agents. Although few studies exist that explicitly evaluate artificial agents as training partners, there is evidence that training with an artificial agent is at least as good as training with a human (Lin, Gal, Kraus, & Mazliah, 2014). In this paper, we present a novel cognitive agent designed to imitate the strategies and theory-of-mind capabilities of human negotiators. Further, we present results from a small-scale pilot study that suggest the agent is capable of performing similarly to humans in a multi-issue bargaining scenario.

Simulation and behavioral studies have shown that theory-of-mind can improve negotiation outcomes. Specifically, agents with more complex theory-of-mind can achieve greater individual and collective outcomes than those with less complex or no theory-of-mind (de Weerd, Verbrugge, & Verheij, 2015). Moreover, negotiating with an agent that has theory-of-mind can encourage humans to use more complex theory-of-mind (de Weerd, Broers, & Verbrugge, 2015). For this reason, we developed an agent that explicitly reasons about the preferences and strategies of its opponent. We expect that practicing with an agent that has these capabilities will provide better learning outcomes than practicing with an agent without theory-of-mind.

The Metacognitive Agent

Our negotiation agent is based on ACT-R's declarative memory system (Anderson, 2007) and instance-based learning (IBL) theory (Gonzalez & Lebiere, 2005). Instance-based learning was chosen as the reasoning mechanism because it provides a flexible method to reason about novel situations based on examples from previous interactions. Moreover, we believe that theory-of-mind in this domain requires explicit, declarative reasoning rather than procedural knowledge. Many negotiation contexts are relatively novel and our participants were not experienced in professional negotiation, so it seems unlikely that people in our target population have sufficient practice to develop a comprehensive set of production rules that match each possible situation.

The agent's memory contains a set of examples (instances) that represent possible negotiation moves and possible contexts in which they may sensibly be used. Each instance is associated with a particular strategy (cooperative or aggressive). For example, when there is a deadlock, a player using a cooperative strategy might concede on a minor issue to break it. However an aggressive player might threaten to quit if his/her offer is not accepted. The agent uses this same knowledge to choose its own moves, to evaluate the player's strategy, and to make inferences about the player's preferences.

The Smoking Ban Negotiation Scenario

The specific setting we consider is a multi-issue bargaining scenario in which a representative of a city council and a representative of small business owners negotiate over the implementation of new anti-smoking regulations. The negotiation involves four issues, each with four or five different options. The task of the negotiators is to reach an agreement which assigns exactly one option to each issue. Despite this simple setup, this setting allows for 400 different possible negotiation outcomes in addition to the opt-out outcome. It also allows a rich set of possible negotiation moves (see Table 1 for definitions of possible moves).

In our setup, each negotiator has preferences that assign a value to each possible option. Higher values are associated with more preferable options. The value of a negotiated outcome is calculated as the sum of the values of the agreed-upon options. A negotiator therefore aims for a negotiated outcome that his preferences assign as high a value as possible. Importantly, preferences are private information. That is, each player knows their own preferences, but not the preferences of the other.

We pilot-tested this scenario to get a sense of the possible negotiation moves and strategies used by human negotiators. In this pilot test, four pairs of human participants negotiated an agreement for nine different problems with unique preference values. From this pilot testing, we made two observations. Firstly, the people in our sample did not always find optimal agreements and sometimes even accepted negative deals. Our measure of optimality is Pareto optimality. An agreement is Pareto optimal if it there are no possible alternative agreements that could raise one player's score without reducing the other player's score. The human dyads found Pareto optimal agreements in 61% of the problems. Moreover, 21% of the negotiated deals resulted in a negative score for at least one of the negotiation partners. These results confirm that our negotiation setting is sufficiently challenging for a training intervention.

Our second observation is that unlike automated negotiation agents, which submit offers as binding commitments, human players often make offers with lower levels of commitment. In other words, players can discuss preferences on various options without being bound to accept those options later in the negotiation. To capture this, we designed our artificial agent so that it can understand moves with differing levels of commitment. For example, exchanging preferences on an issue implies a low level of commitment, but accepting an offer implies a higher level of commitment.

Instance-based Decision-making

Instance-based learning was implemented here using a modified version of Java ACT-R (Salvucci, 2013). An instance is a set of slot-value pairs that represents a context, an action, and a utility value for that action (Gonzalez & Lebiere, 2005). Table 2 contains the specific slots used in the instances in our metacognitive agent.

To select a move, the model uses the current negotiation context to retrieve an instance from memory. The instances are retrieved using ACT-R's partial matching mechanism (Anderson, 2007). The more similar the instance is to the current context, the more active the instance will be and therefore the more likely the instance will be retrieved. Each instance also has a base activation level. This base activation is constant across all instances in the agent, save for a very small amount of noise (s=.01). The instances in the current model were written by the modelers to ensure a stable, challenging opponent whose cooperative and aggressive strategies were consistent with those found in the literature (Fisher & Ury, 1981; Raiffa, 1982).

As an example, suppose the agent retrieves an instance like the one in Table 3. In this context, the agent is playing aggressively. The agent has proposed an option that is worth a lot of points (4 is the maximum in this setting), and it believes that this option is bad for its partner (the option would cause the partner to lose points). The agent's partner has proposed an option that is worth 3 points to the agent, which is still a very good value for the agent, and it happens to be the agent's next best option. Moreover, the agent is already doing well in the negotiation, because on the other resolved issues, it has already gained 3 points. In this case, if the agent retrieves this instance, it will try to pressure its opponent to take a loss even though a more mutually beneficial option is probably available. This is indicated by the Insist move. By contrast, a cooperative instance would likely accept the opponent's bid in this situation.

Theory-of-mind

Negotiators strike a delicate balance between cooperation and competition (Lax & Sebenius, 1986). Cooperation helps encourage agreement and trust between negotiators but it can be exploited by competitive negotiators. A good negotiator is mindful of this and takes steps to prevent exploitation. One way to achieve this is to use theory-of-mind to infer the opponent's strategy. Our agent achieves this by taking the perspective of its partner and using its own knowledge to evaluate the partner's strategy. The agent then attempts to match the toughness level of the opponent. If the partner is cooperative, the agent will also be cooperative. But if the partner is aggressive, the agent will become more aggressive. This meta-strategy has been observed in humans in negotiation and coordination games (Kelley & Stahelski, 1970; Smith, Pruitt, & Carnevale, 1982) and has been shown to be effective at encouraging cooperation in the Prisoner's Dilemma (Stevens, Taatgen, & Cnossen, 2016).

Each time the user makes a negotiation move, the agent assumes the perspective of the user and uses its own instances and decision processes to infer the user's strategy and beliefs. Of course, the agent does not have access to the same information when interpreting the opponent's actions as it does when it is choosing its own (e.g. exact preference values, user's chosen strategy). In these cases, the agent fills in its best guess or leaves the slot empty. Fortunately, ACT-R's declarative memory system is robust to missing information and memory retrievals can be made without specifying all of the slots. As more memory retrievals are made, the agent updates its best guess about the user's strategy. The agent evaluates both the user's reaction to the agent's move (if applicable) and the countermove made by the user. This results in up to two memory retrievals per negotiation turn.

The agent's memory holds three sets of preference values: the agent's own preferences, the agent's beliefs about the user's preferences, and the agent's beliefs about the user's beliefs about the agent's preference values. At the beginning of the negotiation, the agent has no beliefs about its opponent's preferences or beliefs. As the negotiation progresses, the agent gradually adds information to these sets based on the information found in instances retrieved during theory-of-mind. For example, suppose the agent retrieves the in-

Table 1: Overview of possible negotiation moves.

Move	Explanation
Invite	Elicit an offer from the trading partner on at most one issue. *Example:* "What would you like for the scope of the smoking ban?"
Inform	Inform the trading partner that the player likes or dislikes a single option of a given issue (can also indicate no preference). Liking an option that the partner has suggested or liked results in an agreement (see Agree) *Example:* "For me, a 10% increase in tobacco taxes would be difficult."
Suggest	Ask the partner to commit to a specific option on 1 or 2 issues. **Subtypes:** ***Concede -*** Suggest an option that you haven't Suggested before; ***Insist -*** Suggest an option that have Suggested before; ***Exchange -*** Suggest one option from two different dimensions. Both options are conditional. If one is rejected, so is the other. *Example:* "Would you agree to all outdoor smoking allowed in exchange for a 25% increase in tobacco taxes?"
Agree	Agree to the most recent Suggest or Inform move of the trading partner. This is a non-binding commitment. *Example:* "We can do it as you suggested."
Finalize	Commit to a given negotiation outcome. If the partner accepts, this commitment is binding. *Example:* "So to summarize, I think we should go for all outdoor smoking allowed, no change in tobacco taxes, anti-smoking television advertisements, and a ban on tobacco vending machines"
Accept	Accept the most recent Finalize move of the trading partner. This is a binding commitment. *Example:* "That's a deal."
Withdraw	Causes immediate negotiation failure. *Example:* "We cannot seem to reach agreement. Let's stop negotiating."
Final Offer	Commit to a given negotiation outcome and force the trading partner to either accept this outcome or withdraw from negotiation. This is a binding commitment. The negotiator cannot resume the negotiation if the partner rejects the offer. *Example:* "I think we should settle on all outdoor smoking allowed, no change in tobacco taxes, anti-smoking television advertisements, and a ban on tobacco vending machines. This is my final offer."

stance in Table 3 to interpret its partner's move. According to the instance, one situation in which a player might insist is when they have a strong preference for their current bid and a negative preference for the opponent's bid. Therefore, the agent guesses that the opponent's preference for the their bid is 4.0 and the opponent's preference for the agent's bid is -2.0. These two values are then used to retrieve instances in later turns by filling in the "my-bid-value-me" and "opp-bid-value-me" slots respectively.

In a similar way, when the agent submits a move, the agent makes a guess about how its move influences its opponent's beliefs. For example, suppose the agent indicates that it likes a particular option ("Positive Inform"). The instance retrieved by the agent indicates that the this move is appropriate when the agent has a positive preference for the option (e.g. the value in the "next-bid-value-me" slot is 2.0). Now, the agent will believe that its partner believes that it has a preference of 2.0 for the option. This belief is then used to fill in slots during the agent's theory-of-mind reasoning process.

The process of inferring the user's strategy is similar to that of inferring preferences. As each new instance is retrieved, the agent notes whether the instance is cooperative, aggressive, or neutral. When the instances are cooperative or aggressive, the agent becomes more confident that the user is using that strategy. The confidence value for a given strategy is the activation level of that strategy in memory divided by the total activation of all strategies in memory. When this value exceeds a certain threshold, the model will switch to the appropriate counter-strategy. The threshold is a free parameter of the agent, and can be changed depending on the negotiation context. By default it is set to 0.55.

The agent adjusts its strategy according to the perceived aggression of the opponent. The agent has three modes: cautious, cooperative, and aggressive. The agent begins in cautious mode. This mode is designed to encourage cooperation from the opponent while still guarding against aggression. In this mode, the model prefers neutral moves, followed by cooperative, and then aggressive. If the agent believes the opponent is behaving cooperatively, it will enter cooperative mode, in which the agent favors cooperative moves, followed by neutral, and then aggressive. Finally, if the agent is confident that the opponent is unconditionally aggressive, then it will switch to aggressive mode, in which it favors aggressive moves, followed by neutral, and then cooperative.

Table 2: Structure of an instance in the metacognitive agent

Move type	Explanation
Strategy	The strategy associated with the instance
My-bid-value-me	The number of points the agent's bid is worth to the agent.
My-bid-value-opp	The number of points that the agent believes its bid is worth to the user.
Opp-bid-value-me	The number of points the user's bid is worth to the agent.
Opp-bid-greater	True if the user's bid is at least as much as the agent's current bid, False otherwise.
Next-bid-value-me	The number of points that the next best option is worth. The next best option is defined as the option closest in value to the current one (Not including those that are worth more than the current option.)
Overall-value	The total value of all options that have been agreed upon so far. This is a measure of how the negotiation is going. If it is negative, negotiation is likely to result in an unacceptable outcome.
My-move	The move that the agent should take in this context.

Table 3: An example of an aggressive instance.

Slot name	Value
Strategy	Aggressive
My-bid-value-me	4.0
My-bid-value-opp	-2.0
Opp-bid-value-me	3.0
Opp-bid-greater	false
Next-bid-value-me	3.0
Overall-value	3.0
My-move	Insist

Graphical Interface

Learners can interact with the agent through a graphical interface (see Figure 1). The interface contains five zones. The top zone shows the state of the negotiation. This includes a representation of the negotiation agent's current cooperativeness and a transcript of the negotiation. In this transcript, simulated negotiation dialog is shown in green and simulated agent dialog is shown in orange. This simulated dialog is taken from transcripts of human-human dyads participating in an earlier pilot study on the smoking ban scenario.

The second zone shows the possible ways in which the learner can respond to an offer made by the agent (if any). If the agent has made an offer, the learner may give it a positive, negative, or neutral evaluation. A positive evaluation indicates a tentative agreement, a negative evaluation indicates that an offer is undesirable, and a neutral evaluation states that the offer is under consideration. If the agent has not made an offer, this zone is disabled.

The third zone features the possible actions that can be performed by the user. Actions that are impossible to take at the moment are disabled. The third zone shows the four issues, each with its own options. The background color of each option indicates the evaluation of the option for the user. Darker red colors indicate options that are increasingly more negative, while darker blue options indicate increasingly more positive options. In addition, for each issue, colored triangles indicate the option most recently offered by the user (green) and the negotiation agent (orange). The final zone of the interface gives a preview of the move the user is about to make, and a button to submit that move.

The action selected in the third zone determines what can be selected in the fourth zone. To help the user, actions are grouped by their level of commitment. In addition, two separate buttons are used for proposing and exchanging offers. Proposing offers are offers that assign a single option to exactly one issue, while exchanging offers are offers that assign a single option to exactly two issues. Note that an exchanging offer is interpreted as a temporary offer. If an exchanging offer is not accepted, the triangles indicating the most recently offered option revert to their previous positions. The interface is of course more restrictive than a real-life negotiation. For example, the interface does not allow users to make offers on more than two separate issues. In addition, the interface automatically handles proper Agree, Accept, Finalize, and Final Offer moves. This means that a user can only make an Agree move when the negotiation agent has made an offer and Accept when the agent has made a Finalize or Final Offer move. A user can only make a Finalize or Final Offer move if a green triangle indicates an option for each issue. This means that users cannot attempt to Agree to offers that have not been made, or make partial Finalize moves.

Agent Feedback

During the negotiation, the agent accumulates data about the players' actions to present as feedback. The agent provides feedback on two different aspects of the learner's performance: negotiation style and outcomes. Negotiation style concerns the learner's strategy (cooperative or aggressive) and outcomes refers to the utility of the agreement reached for both players.

Throughout the negotiation, the metacognitive agent evaluates its trading partner on his/her negotiation style. After each action, feedback on the perceived cooperativeness of the action, the agent's beliefs about the preferences of the player, and the accumulated perception of the cooperativeness of the agent's trading partner is available immediately to display as feedback. This includes changing the facial expression of the agent to happy (cooperative), angry (aggressive), or neutral

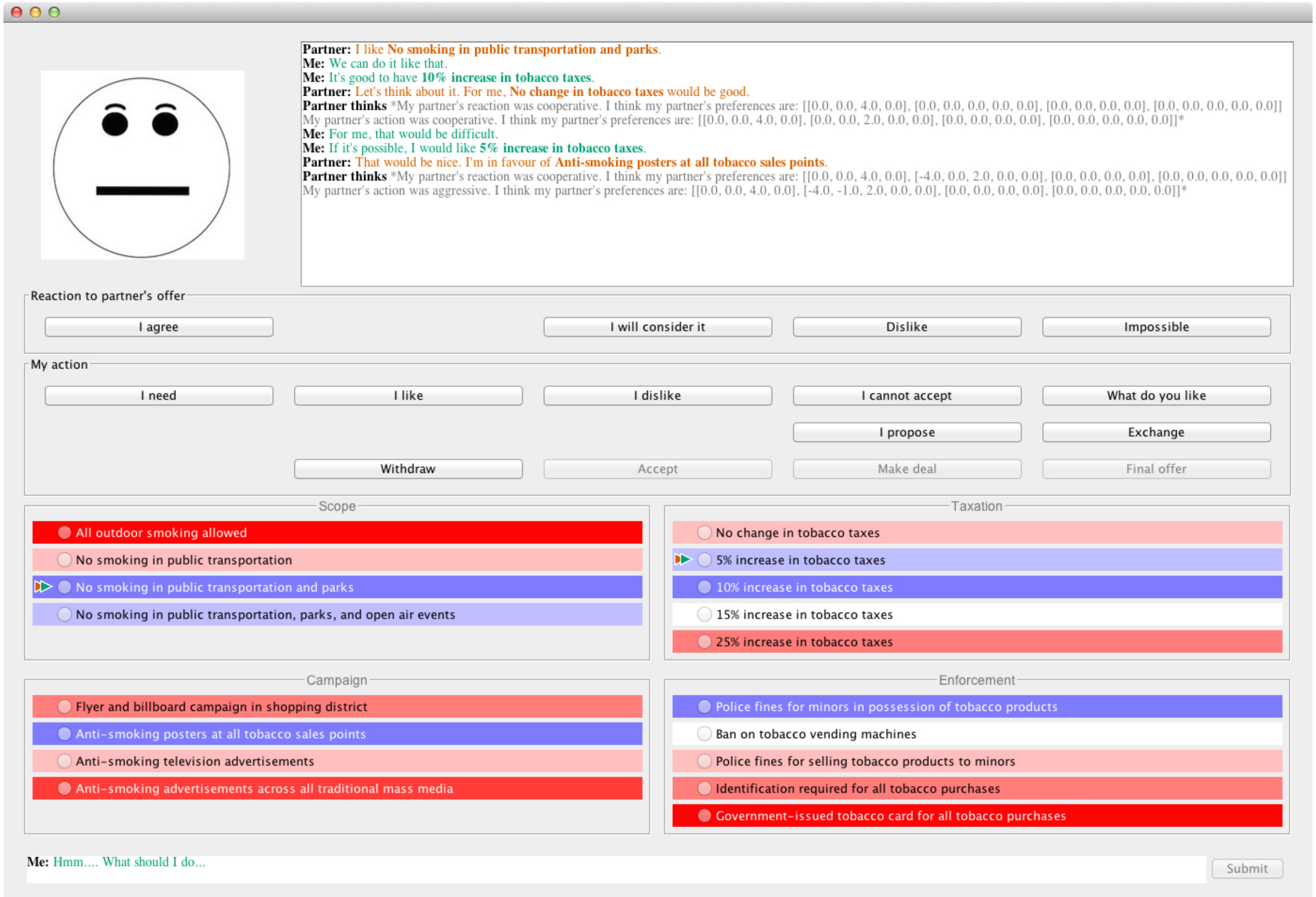

Figure 1: The graphical interface of the agent.

via a line-drawing of a smiley face. Moreover, a line of text is added to the transcript showing whether the agent believes the move was cooperative or aggressive. This line also includes a representation of the agent's beliefs about the user's preferences (see Figure 1).

Once a negotiation has finished and the outcome is known, additional feedback is available about the economic outcome. First, the GUI informs the player how many points he/she gained or lost as a result of the deal. These points are based directly on the colors of the preference panel in the GUI. A good negotiator does not accept a deal that is worse than no deal (i.e. has a negative score), and the agent will inform the learner when this occurs. In addition to the individual outcome, the agent evaluates the Pareto optimality of the agreement and displays it in the GUI.

Pilot experiment

To test the agent's ability to negotiate rationally with human users, we conducted a pilot study. In this study, six human users negotiated with the agent on the same nine problems we used in the pilot experiment with human dyads. The agent always represented the business side. Overall, the results of the pilot study are encouraging, and suggest that the agent is a competent negotiation partner. The agent did not exploit, nor was it exploited by the human users. The agent earned an average of 32.2 (out of 88) possible points while the human users scored an average of 33.7 (out of 86) points. Thus, agreements were similarly beneficial for both partners. The human-human dyads also achieved a relatively even point distribution, albeit with a slight advantage for the business side. Moreover, we compared the agent-human dyads to the human-human dyads on rate of agreement, ability to find Pareto optimal outcomes, and acceptance of negative outcomes (see Table 4). We find that the humans reach a similar number of agreements when negotiating with the agent as when negotiating with each other. Furthermore, the human-agent dyads find a similar (though slightly lower) percentage of Pareto optimal deals. Finally, the human-agent dyads reached fewer negative deals than the human-human dyads. This does not necessarily mean that the human-agent dyads were superior at avoiding negative deals. It is possible that the Withdraw option was more salient in the human-agent dyads due to the GUI. Also, in the human-agent case, the human users received explicit feedback about their scores after every round. Therefore it was clearer how the colors mapped onto overall scores.

Table 4: Comparison of human-human dyads to human-agent dyads

	Human-Human (n = 4)	Human-Agent (n = 6)
Mean business score	33 (20)	32 (3)
Mean council score	29 (7)	34 (8)
% Agreement	78	74
% Pareto optimal	61	52
% Negative deals	21	2

Note. In the human-agent dyads, the business side was always played by the agent. All percentages represent percentage of the total number of trials. SD's are presented in parentheses.

Future Directions

The present pilot study of course does not test the educational outcomes resulting from training with the agent. Evaluation of learning gains is ongoing. In future studies, we aim to test the extent to which training with the agent improves outcomes not only in the smoking ban scenario, but also in other negotiation contexts.

Currently, the agent possesses instances that were hand-coded by the authors and the base activation level does not change. However, allowing the agent to learn the utilities of the instances could result in a more dynamic, and potentially more intelligent, opponent. This is possible, but challenging, for a task like negotiation. Instance-based learning requires a measure of utility, and the utility of a particular negotiation move is not always immediately clear. Therefore, implementing learning in such an agent requires careful consideration of the learning and social context of the negotiation to avoid chaotic agent behavior.

This agent is not designed to be restricted to a point-and-click interface. Rather, it is meant to be a component of a larger system known as Metalogue, a large, multimodal negotiation trainer capable of simulating a real negotiation dialogue (Helvert, Rosmalen, Borner, Petukhova, & Alexandersson, 2015). In the coming months, the agent will be incorporated into this system, and will function as a decision engine. As it does in the GUI setting, the agent will play the role of a negotiation partner and trainer discussing the options of a smoking ban. However, in this case, the learner will be able to interact with the model through speech rather than through clicking buttons. Moreover, the agent will be portrayed by a virtual avatar with speech and gestures of its own.

Summary

Here we have presented a novel cognitive agent that reasons about the goals and strategies of human partners to successfully engage in a negotiation task. This agent leverages established cognitive theories, namely ACT-R and instance-based learning, to generate plausible, flexible behavior in this complex setting. Our preliminary results suggest that our cognitive agent could play a role in training effective negotiators.

Acknowledgments

This work was funded by European Union Grant 611073: Multiperspective Multimodal Dialogue (METALOGUE).

References

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* Oxford University Press.

Cross, J. G. (1977). Negotiation as a learning process. *Journal of Conflict Resolution*, *21*(4), 581–606.

de Weerd, H., Broers, E., & Verbrugge, R. (2015). Savvy software agents can encourage the use of second-order theory of mind by negotiators. In *Proceedings of the 37th annual conference of the cognitive science society.* (pp. 542–547). Pasedena, California.

de Weerd, H., Verbrugge, R., & Verheij, B. (2015). Negotiating with other minds: the role of recursive theory of mind in negotiation with incomplete information. *Autonomous Agents and Multi-Agent Systems*.

Fisher, R., & Ury, W. L. (1981). *Getting to Yes: Negotiating Agreement Without Giving In*. London: Penguin Group.

Gonzalez, C., & Lebiere, C. (2005). Instance-based cognitive models of decision-making. In D. Zizzo & A. Courakis (Eds.), *Transfer of knowledge in economic decision making.* New York: Palgrave McMillan.

Helvert, J. v., Rosmalen, P. V., Borner, D., Petukhova, V., & Alexandersson, J. (2015). Observing, coaching and reflecting: A multi-modal natural language-based dialogue system in a learning context. In *Proceedings of the 11th international conference on intelligent environments* (p. 220).

Kelley, H. H., & Stahelski, A. J. (1970). Social interaction basis of cooperators' and competitors' beliefs about others. *Journal of Personality and Social Psychology*, *16*(1), 66–91.

Lax, D., & Sebenius, J. (1986). *The Manager as Negotiator*. Boston: Harvard Business School Press.

Lin, R., Gal, Y., Kraus, S., & Mazliah, Y. (2014). Training with automated agents improves people's behavior in negotiation and coordination tasks. *Decision Support Systems*, *60*(1), 1–9.

Raiffa, H. (1982). *The art and science of negotiation*. Cambridge: Belknap Press.

Salvucci, D. D. (2013). Integration and reuse in cognitive skill acquisition. *Cognitive Science*, *37*(5), 829–860.

Smith, D. L., Pruitt, D. G., & Carnevale, P. J. (1982). Matching and mismatching: The effect of own limit, other's toughness, and time pressure on concession rate in negotiation. *Journal of Personality and Social Psychology*, *42*(5), 876–883.

Stevens, C. A., Taatgen, N., & Cnossen, F. (2016). Instance-based models of metacognition in the prisoner's dilemma. *Topics in Cognitive Science*, *8*(1), 322–334.

Tinsley, C. H., O'Connor, K. M., & Sullivan, B. a. (2002). Tough guys finish last: The perils of a distributive reputation. *Organizational Behavior and Human Decision Processes*, *88*(2), 621–642.