On integrating Theory of Mind in context-aware negotiation agents

Dan Kröhling, Ernesto Martínez

INGAR (CONICET/UTN)
Avellaneda 3657
Santa Fe, Argentina

Abstract. Theory of Mind (ToM) is the ability of an agent to represent mental states of other agents including their intentions, desires, goals, models, beliefs, how the environment makes an impact on those beliefs, and the beliefs those agents may have about the beliefs others have about themselves. Integrating artificial ToM in automated negotiations can provide software agents a key competitive advantage. In this work, we propose integrating ToM into context-aware negotiation agents using Bayesian inference to update each agent's beliefs. Beliefs are about the necessity and risk of the opponent considering hypothesis about how it takes into account contextual variables. A systematic hierarchical approach to combine ToM with using evidence from the opponent actions in an unfolding negotiation episode is proposed. Alternative contextual scenarios are used to argue in favor of incorporating different levels of reasoning and modeling the strategic behavior of an opponent.

Keywords: Theory of Mind, Negotiation Agents, Context Awareness, Bayesian Models

1 Introduction

In the struggle for survival, humans were forced to compete and cooperate with others in the common environment they inhabit. Over the years, this need has led to the emergence of highly competitive environments where each individual must interact strategically with its pairs in order to accomplish its own goals, which in turn made environments even more complex. Strategic behavior in interacting socio-technical systems demand the development of increasingly sophisticated decision-making skills that required humans and software agents to reason about the strategies of other agents they interact with, to reason about how others reason, what do they think others know they know and think, and so on and so forth. This recursive type of mental reasoning about others has gained a lot of attention lately, and is usually referred to as Theory of Mind, or ToM.

Theory of mind denotes the ability to represent the mental states of others, including their beliefs, intentions and goals [2, 16], which can be used advantageously over shallower ways of reasoning [15]. This ability was found at first in humans, but it is also present in other species [7], which led to the possibility of implementing it in machines and artificial software agents [12, 13]. Recent research efforts have demonstrated ToM effectiveness at large [3, 4, 10, 17, 19], in a number of approaches according to Qi et al. [11], and these works seem to be the tip of the iceberg to further learning and reasoning in multi-agent systems.

Although most previous studies are focused in games as Colored-Trails [3], Common Pool Resource Games [18], Rock, Paper, Scissors and its variants [4] and in Game-Theoretic approaches, they address, at least partially, the negotiation problem [6]. However, to the best of our knowledge, there exists just a handful of works that combine the negotiation and its context [14, 20], and certainly none of the previous research efforts in the area of autonomous agents use ToM to predict other parties thoughts about others, to decide their next offers given the context and other agents beliefs and models.

There are three major contributions in this work. The first contribution is related to the use of theory of mind applied to negotiations that take the context into account, where agents will try to identify the intentions of their opponents through their offers and the use of a the Bayes rule to update an opponents beliefs. The second contribution is the use of private deadlines for each negotiation agent during a negotiation episode. In the literature about automated negotiation is it is assumed that the deadline is the same for both agents and is public knowledge. The third contribution is addressing most of the challenges and difficulties that ToM brings into automated negotiations when combined with identifying key contextual variables which are used to update prior beliefs from sparse data in a negotiation episode.

To begin with, an introduction to automated negotiations and the negotiation setting is presented in Section 2. In Section 3, the concept of theory of mind is described together with a group of ToM context-aware negotiation agents with increasing orders of sophistication. To continue with, in Section 4, a number of computational experiments made are discussed. Some concluding remarks about our research efforts in implementing ToM in automated negotiations are made in Section 5. Finally, in Section 6, future research avenues are presented.

2 Negotiation Setting

A number of approaches to address the problem of automated negotiations has been taken. The works of Fatima et al. [5] and Baarslag [1] made up a compendium of the state of the art in automated negotiations. In [5], differences between single and multiple issues, bilateral and multilateral negotiations are discussed, with discrete and continue values for the attributes of the issues being negotiated. Machine-machine and man-machine negotiations are addressed, as well as different negotiation protocols, domains, and negotiation agendas. In [1], different techniques for offering, accepting, and opponent modelling are presented, alongside a framework to develop negotiation software agents.

A group of negotiation parties that agree to negotiate over a certain issue is considered in this work. According to [8], negotiation parties are the agents that participate in different negotiation episodes, and issues are the resources to be allocated or the terms of agreement to be addressed during a particular negotiation episode. In this work, the same considerations will be used. The mechanism of bilateral negotiations between two agents taking actions seeking to agree about one single issue with discrete values using a discrete time line. The alternating-offers protocol [9] is used throughout.

During a bilateral negotiation of an issue, the two negotiation agents will make offers and concessions between an initial price and a reserve price, as shown in figure 1[1]. The initial price (IP) is the first offer made by a given agent in the negotiation episode. The reserve price (RP) of a negotiation agent is the ultimate offer that agent is willing to accept. Note that every offer outside this *zone of agreement* is detrimental to any concerned agent. Clearly, in a buyer-seller negotiation, a deal can only take place if the reserve price of the buyer is greater than the reserve price of the seller. In this situation, the zone of agreement of a negotiation episode is determined by the reserve prices of both agents involved.

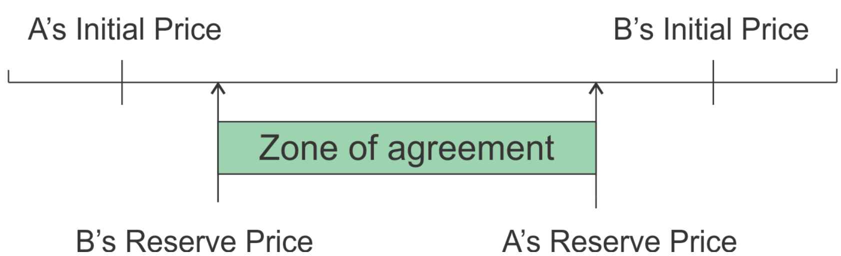

Fig. 1: Zone of agreement in a negotiation episode between two agents A and B.

In this work, the main assumption is that the initial price IP for each agent is private information, as well as its negotiation deadline, which will be named n. However, the fact that this deadline is private to each agent forces both agents to consider not only their private deadlines while negotiating, but also to reason about which would be their opponent's deadline.

One of the most novel aspects of our work is the negotiation context, and how do agents consider it. A colloquial definition of the negotiation context can be found in [14], in which the negotiation context is presented as the characteristics or circumstances under which the negotiation process occurs. More formally, in this work the context is represented by two abstract spaces: the agent's private information and the contextual variables which are assumed common knowledge.

An agent's private information is composed by all his internal or private variables, those that other agents cannot see but could attempt to model observing the actions the agent takes. The private information can refer to the needs or urgency of a factory to buy supplies due to low inventory and high demand, the rush of a family to buy a bigger car, or the urgency of a brand to avoid

[1] Adapted from [20].

bankruptcy. The *necessity* that an agent has given this private information will be summarized in a single variable ν, that varies between 0 and 1.

The negotiation context, on the other hand, is composed by all the external or public variables, those that every agent would access if deemed necessary. The public information may refer to changes in global currencies, variations in the weather forecast when selling wind energy, or particular events as the fall of a government leader. The *risk* that an agent internally perceives given this publicly available information will be summarized in a single variable ρ, that varies between 0 and 1.

3 ToM-Context-Aware Negotiation Agents

In this section, the proposed design for our ToM context-aware agent is presented, together with a short introduction to the concept of theory of mind. Following the definition given by Singer et al. [16], Theory of mind, or the equivalent term “mentalizing”, connotes an agent’s ability to cognitively represent the mental states of others, including intentions, knowledge, desires, goals, beliefs, and models. This ability combined with purposefully deceiving and manipulating opponents’ thoughts could provide agents with a social competitive advantage, as discussed in [4]. In this work, it is argued that by mentalizing and using Bayesian inference about opponents’ strategies (and their contextual-driven models) during a negotiation episode may provide a competitive advantage for maximizing an agent’s utilities. This approach differs from the works of Fatima et al. [5] and Baarslag [1], primarily because they concentrate in ToM-0 agents, where no opponent’s models are used, and ToM-1 agents, where agents suppose that others do not have the ability to mentalize about themselves.

Fig. 2 will serve us to explain how ToM works. Suppose there are two agents, A and B, negotiating in a certain environmental situation or context. As a ToM-0 agent, agent A models the negotiation context but does not explicitly take into account or model the other agent’s behavior, for instance B, during a negotiation episode. In other words, agent A only perceives the effect of agent B’s actions, but does not model this agent as an entity capable of reasoning about its opponent’s actions or the influence of contextual variables on the opponent’s behavior. As a ToM-1 agent, A distinguishes the presence of a situated negotiation agent as if it were a ToM-0 agent and model the negotiation context as immutable to actions taken by both negotiating agents. That is, agent A models agent B as if the latter were aware of the context and its contextual variables but could not consider how this context impacts on A itself. In the same way, a ToM-2 A would model B as if it were a ToM-1 agent, and so it continues. In general, a ToM-m A would model B as a ToM-m-1 agent. The ToM level of A will be represented by m^A.

As previously described in Section 2, the influence of the negotiation context on each agent will be modeled by two variables, ν and ρ. Then, our agent will use these variables to adapt its negotiation strategy using Bayesian ToM modeling of its opponent including beliefs, desires, intentions and preferences.

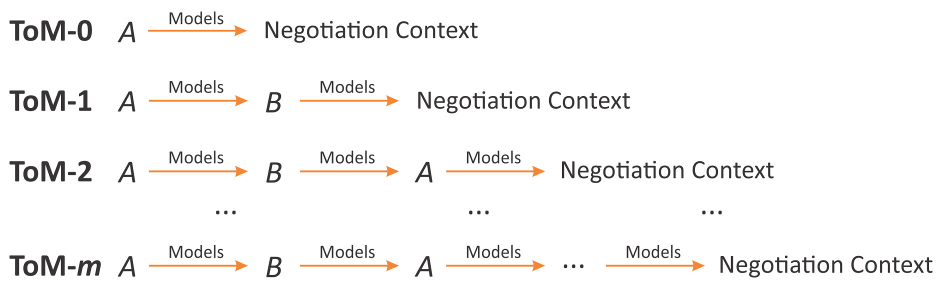

Fig. 2: Theory of mind explained.

It is worth highlighting here that one of the challenges for negotiation agents with levels of ToM higher or equal than 1 is to be aware that they cannot wait until their private deadlines to concede more. This is because private deadlines of their opponents may be shorter than theirs, and sticking to their own deadline can result in an unfavorable or unexpected negotiation outcome.

3.1 A ToM-0 Negotiation Agent

Adapted from Fatima et al. [5], a heuristic concession strategy for a ToM-0 negotiation agent A is defined as follows:

$$O_t^A = IP + (RP - IP) * \left(\frac{t}{n}\right)^{1/\beta} \tag{1}$$

where O_t is the offer agent A will make at time t, IP is the initial price, which is assumed to be the best deal an agent considers it can obtain from the negotiation, RP is the reserve price, which is the worst deal the agent may accept at the end of an episode. Finally, t is the time elapsed from the beginning of the negotiation episode, n is the agent's deadline, and β defines the concession rate, that is, how much a given agent would concede during the negotiation.

As previously said in Section 2, IP and n are assumed private information for each agent, and t will advance as the negotiation episode proceeds. RP and β will depend on the context, as equations 2 and 3 show:

$$RP = 1.0 - \nu \tag{2}$$

$$\beta = 1.5 * (1.0 - \rho) + 0.5 \tag{3}$$

This simplified negotiation strategy will suffice to prove the potential of resorting to an Artificial ToM and Bayes rule for modeling opponents behavior in automated negotiations and taking key contextual variables into account.

The acceptance strategy to be used by an agent A is taken from Baarslag [1]. A will accept an offer from B if the following acceptance condition is verified:

$$u(O_{t-1}^B) \geq u(O_t^A) \tag{4}$$

where $u(O_t^{agent})$ is the utility A will obtain from the offer O made by *agent* at time t. The efficacy of this rather simple acceptance strategy was demonstrated in [1], and consists of the acceptance of an offer from an opponent whenever such offer supposes a higher utility for agent A in comparison with the utility it would obtain from its own next offer.

The utility of every agent for an offer O_t will be given by:

$$u(O_t) = \begin{cases} O_t - RP & \text{if } A \text{ is seller} \\ RP - O_t & \text{if } A \text{ is buyer} \end{cases} \tag{5}$$

At the end of a negotiation episode, the offer used to compute the utilities for both agents will be the last offer made by the agent which ended the negotiation.

3.2 A ToM-1 Negotiation Agent

A ToM-1 agent A, as mentioned before, will model any other agent B as if it were a ToM-0 version of itself. Accordingly, this ToM-1 agent A will make some hypothesis about its opponents' behavior. The first hypothesis is that the opponent agent B resorts to the same public information as agent A does. In that sense, it models the risk ρ^B of B as 1 - ρ^A, that is, its own perceived risk if it were agent B. The second hypothesis is that agent B would negotiate as itself would do in a given context. With these two considerations in mind, agent A will model agent B's responses to its offers during the negotiation episode. As ρ^B is modeled directly from the knowledge agent A has about how the context influences its necessity and risk, agent A will try to model ν^B and n^B based on agent B's offers.

Agent A resorts to Bayesian inference [2] to update its predictions about ν^B and n^B through the Bayes' Rule which changes prior beliefs agent A has as shown in Eq. 6.

$$P(H|E) = \frac{P(E|H) * P(H)}{P(E)} \tag{6}$$

In this sense, the new evidence E that makes possible for agent A to update its beliefs about its opponent's desires and contextual model are agent B's offers, and the revised beliefs H that A has about its opponent are agent B's necessity ν^B and deadline n^B. Note that the agent A has a prior belief about B's parameters ν^B and n^B. Initially, this is an uniform distribution, that is, every possible combination of ν^B and n^B is considered as equally likely. As the negotiation episode proceeds, this prior distribution is updated based on the offers received, discarding those combinations or hypotheses that have negligible probability of occurring given received data from the previous offers made by agent B in the negotiation episode. While A tries to model B, it will make no concession, as if its ν^A where equal to 0.0 (unless its deadline is close enough). When A thinks it has enough informative beliefs about the situation agent B is in, it will concede as little as it can, trying to force an agreement at the reserve price of B, RP^B. If agent A is not too confident about the beliefs it has regarding the risk and necessity of its opponent, it behaves simply as a ToM-0 agent, given its context.

3.3 A ToM-m Negotiation Agent

The recursive notion behind ToM may go up towards any level of abstraction m. However, modeling thoughts and the way of behaving of other agents becomes more and more elaborated, and the assumptions made about their thoughts, and what the others think about our own thoughts, require more precise models as ToM levels of reasoning are more abstract and self-referenced.

Algorithm 1: Pseudocode of a ToM-m context-aware negotiation agent A

Data: necessity ν^A, risk ρ^A, deadline n^A, ToM-level m^A, initial price IP^A
Result: negotiation outcome

begin negotiation episode;
for $i \leftarrow 1$ **to** m^A **do**
    initialize ToM-i Bayesian model of B with uniform priors of ν and n;
send first offer IP^A;
for $t \leftarrow 1$ ***to*** n^A **do**
    receive offer O^B_{t-1};
    for *ToM Bayesian models of* B **do**
        update priors of ν^B and n^B given evidence O^B_{t-1};
    if *there is only one ToM-m model that explains* ν^B *and* n^B **then**
        if $n^B \leq n^A$ **then**
            $\nu^A \leftarrow 1 - \nu^B$ and $n^A \leftarrow n^B$;
        else
            compute ν^A so that $O^A_{n^A}$ equals $O^B_{n^A}$;
        if $\nu^A \geq \nu^A_{t=0}$ **then**
            play ToM-0 strategy: $\nu^A \leftarrow \nu^A_{t=0}$ and $n^A \leftarrow n^A_{t=0}$;
    else if *there is more than one ToM-m model that explains* ν^B *and* n^B **then**
        play a deceitful strategy varying ν^A and n^A;
    else
        play ToM-0 strategy: $\nu^A \leftarrow \nu^A_{t=0}$ and $n^A \leftarrow n^A_{t=0}$;
    compute new offer O^A_t given ν^A, ρ^A, n^A, t, IP^A;
    if $u(O^B_t) \geq u(O^A_{t+1})$ **then**
        accept last offer made by B;
        negotiation episode ends;
    else
        send offer O^A_t;
    for *ToM Bayesian models of* A **do**
        update priors of ν^A and n^A given evidence O^A_t;
reject last offer;
negotiation episode ends;

In Algorithm 1, a pseudocode of the behaviour for the proposed ToM-m context-aware negotiation agent A during an episode is presented, given its internal need or urgency ν^A, perceived risks ρ^A, private negotiation deadline n^A and initial price IP^A.

To make the rationale behind Algorithm 1 more clear, some remarks are given. In lines **7** and **8**, models of agent B are updated. Suppose the case of a ToM-4 agent A. To achieve such a sophisticated level of mentalizing, agent A needs to revise the model ToM-1 B has about ToM-0 A and the model ToM-3 B has about a ToM-2 A. In the same way, in lines **26** and **27**, models of agent A are updated. A will update the model ToM-2 A has about ToM-1 B and the model ToM-4 A has about a ToM-3 B. In these lines, the complexity of recursive modelling can be seen, and the reader can have a glimpse of the problems an incorrect model in the lower levels of ToM may give rise to in the accuracy of the whole mental model of the opponent.

Perhaps the key issue in this algorithm can be found in line **17** where, if agent A has not yet been able to identify the strategy B is using, it will play a deceitful strategy as an incentive for B to show its true type, that is, the true values ν^B and n^B.

Agent A may decide to return to a basic strategy, namely its 0 level of ToM, for two different reasons. The first one is in line **15**, whenever agent A deduces that, given ν^B and n^B and its own ν^A and n^A, there is no way an agreement can be reached. The second reason is in line **19**, where A understands it has failed to identify agent B's strategic profile, and plays a conservative strategy ToM-0 so as not to concede beyond what is strictly necessary. Finally, in lines **21** and **22** the acceptance condition mentioned in equation 4 is presented.

4 Computational Experiments and Results

4.1 Experimental Design

A number of negotiation episodes has been made; each of them considers a different context, where agents will apply different levels of ToM. The contexts are formally defined by different values of necessity and risk for both agents as shown in Table 1.

Context	ν^A	ρ^A	n^A	ν^B	ρ^B	n^B
1	0.1	0.3	20	0.7	0.7	20
2	0.5	0.5	10	0.5	0.5	15
3	0.5	0.5	15	0.5	0.5	10
4	0.6	0.2	25	0.6	0.8	10
5	0.8	0.8	10	0.7	0.2	20

Table 1: The different contexts in which agents A and B are to negotiate.

As said before, there is a single issue under negotiation, and the initial price IP for each agent is assumed as given. Two agents are considered for the negotiation episodes, A and B. Agent A wants to negotiate a price for this issue with the owner, agent B. While the initial price for A, IP_A, is 0, IP_B is equal to 1000, for every negotiation episode. On this basis, the influence of different contexts on the mental models used for taking actions could be properly compared.

4.2 Results

Context 1

Figure 3 depicts the negotiation episodes played between A and B in the context 1, considering that both agents may use different levels of ToM.

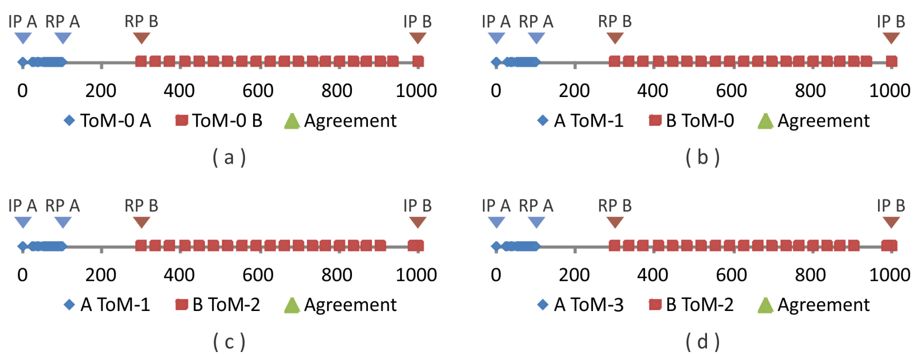

Fig. 3: Negotiation episodes conducted in context 1 with ToM agents.

In this first context, there is not too much to do for these agents, as the reserve price of the seller B is greater than the reserve price of the buyer A. This means that the necessities of A to buy the issue are not so high, and the true necessity that B has to sell is not high enough. For this context, thus, there is no agreement zone: every possible agreement is detrimental to at least one of the negotiation agents. Hence, using ToM does not provide any advantage in such a context, as can be seen from Fig. 3(a) through Fig. 3(d).

Context 2

Figure 4 summarizes actions taken by both agents in the negotiation episodes played in context 2, considering they use different levels of ToM.

In this second context, agent A and agent B consider that the issue being negotiated has a RP of 500, given their necessities. This is a special case, where the only possible agreement is that both agents concede until their RP.

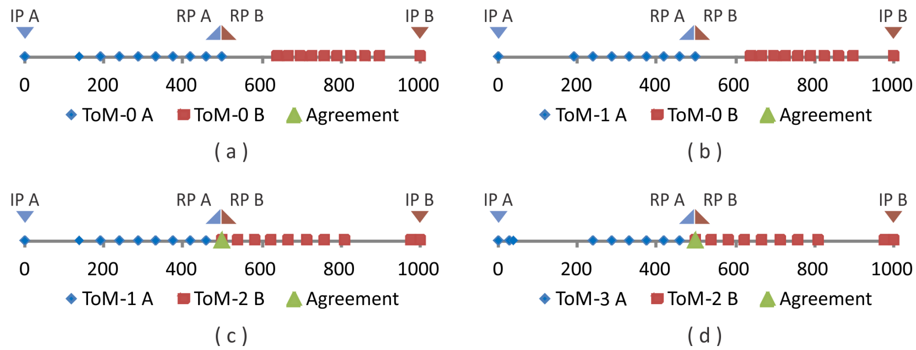

Fig. 4: Negotiation episodes conducted in context 2 with ToM agents.

It will be shown that the use of ToM in this context helps negotiation agents reach an agreement as they understand that the other party also has a negotiation deadline. In Fig. 4(a), none of the agents use ToM, in which case their negotiation strategy is just to concede according to their internal needs ν and risks ρ. In Fig. 4(b), agent A infers that B has a certain deadline, but it cannot do anything more than conceding to its RP, because its own deadline is closer to the actual time t. In other words, it is B's problem that the negotiation had not reached an agreement. In 4(c), B can model A as a ToM-1 agent and model its deadline n_A, and they can finally reach an agreement by accepting their RP.

Context 3

Figure 5 shows the negotiation episodes conducted in context 3. This context is quite similar to context 2, with the exception that agent A has a longer deadline than agent B. As a result, as shown in Fig. 5(b), as soon as agent A properly infers this situation, it can adjust its own deadline to reach an agreement.

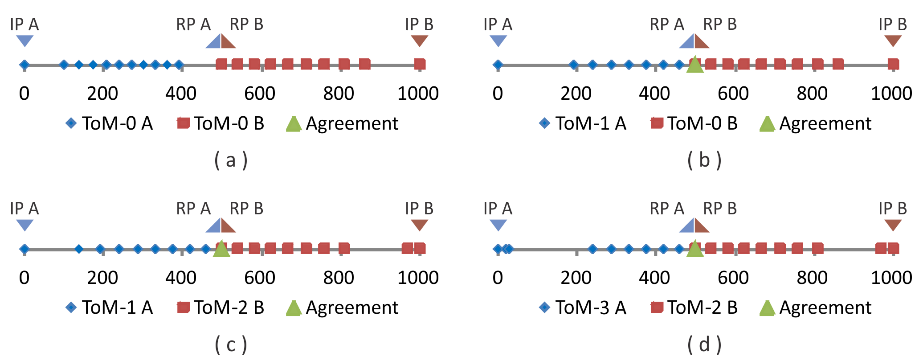

Fig. 5: Negotiation episodes conducted in context 3 with ToM agents.

Context 4

Figure 6 shows the negotiation episodes conducted between A and B in context 4. In this context, ToM plays an interesting role. In Fig. 6(a), when both agents negotiate according to their necessities and risks, they do not reach an agreement, despite there is an agreement zone where both agents can win if an deal is reached. In Fig. 6(b), ToM-1 A adapts its strategy to reach an agreement while using Bayesian inference to model the need of B and its deadline.

Then, in Fig. 6(c), ToM-2 B interprets that A is trying to model its actions. For that reason, B deceives A by playing as if its internal need was lower than it really is. Then, when A thinks that its best choice is to concede until RP_A, B adapts its strategy to concede accordingly. Finally, in Fig. 6(d), as A interprets that B is a ToM-2 agent, it waits until B models its deceitful strategy and reveals its RP. Just then, A concedes until RP_B.

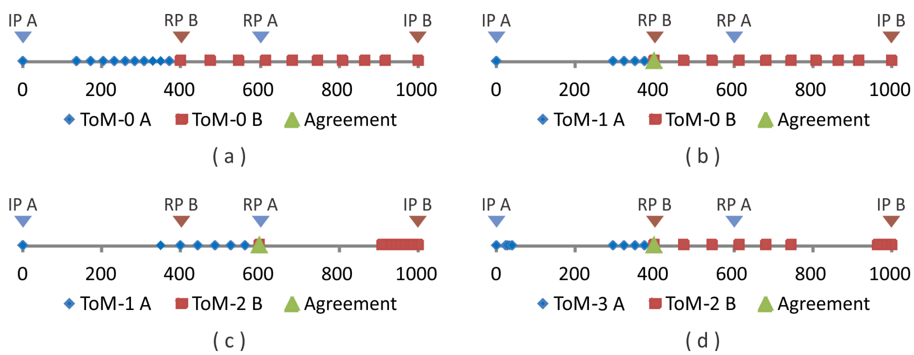

Fig. 6: Negotiation episodes conducted in context 4 with ToM agents.

Context 5

Figure 7 depicts the negotiation episodes conducted between A and B in context 5. In this last context, the benefits of using ToM are undeniable, as it plays a crucial role during the negotiation episode. This is a novel situation in which, even when agents play the strategy suggested by their necessities and risks, agreement takes place, as shown in Fig. 7(a).

In Fig. 7(b), ToM-1 A interprets that B behaves as a ToM-0. In this case, there is no much for A to do, more than to model its opponent and concede, taking into account that its deadline is possibly shorter than B's. For this reason, it cannot wait until B reaches RP_B: agent A estimates the last offer B is going to make in time n^A and concedes until that offer is made. B can surpass that constraint in Fig. 7(c) because, according to its model, A has less time than itself to negotiate, and thus should concede faster until RP_A. Finally, in Fig. 7(d), A

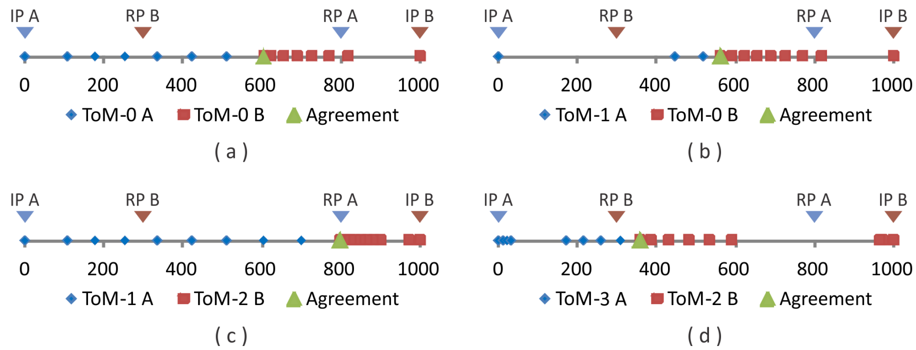

Fig. 7: Negotiation episodes conducted in context 5 with ToM agents.

knows that B makes a ToM-1 model. For that reason, A deceives B and forces it to concede a little more. It is worth noting that B obtains a better outcome in Fig. 7(b) than in Fig. 7(d), despite in the latter it uses a higher level of ToM.

5 Concluding Remarks

Theory of mind constitutes a set of social skills that allows agents to reason about the world they inhabit, to reason about how others reason, what do they think others know they know and think, and so forth. These social skills provide agents desirable abilities when interacting with others, abilities that can increase their utilities when competing and cooperating.

In this work, the role of ToM in automated negotiations between context-aware agents was addressed. After a number of negotiation episodes conducted in different contexts, the importance of understanding this recursive modelling technique (taking into account the models others made out of the context and their pairs) was highlighted.

Some problems remain open, though, in this new research area. The first problem is the assumptions an agent makes when modelling others. For instance, if B does not act the way a ToM-1 A imagines or assumes B should act in a given context, then the model of this ToM-1 agent fails, and further levels of ToM will inevitably fail as well. In the same manner, the assumption of A that B has a similar way of modelling the world that it does can be misleading. For instance, if the risk B perceives is different to the risk A thinks B perceives, then the model A uses to take actions fail. This first problem is concerned with the difficulties of correctly modelling the world and the opponents.

The second problem is about the incorrect interpretation of the level of ToM used by the opponent. If the agent A does not make the right assumptions about the reasoning deepness that the agent B makes of itself, then A's actions may be detrimental. Figure 8 highlights this problem in context 5. In (a), A uses a level of ToM exceedingly high, performing even worse than it would do by using no ToM at all, as no agreement is reached. In (b), A uses the same level of ToM as B, but B takes far from optimal decisions. It is worth noting that A performs better here than in 7 (d), when it uses the "correct" level of ToM, $m = 3$. In 7 (d), B reckons that its ToM model fails to deliver, and resorts to its ToM-0 strategy. In 8 (b), B correctly identifies the deadline of A, n_A, but incorrectly identifies its RP_A, and it is not aware of such an error. In this case, A was benefited by the confusion of B, but the same confusion may happen to A itself in other contexts. This second problem illustrates the difficulties of correctly identifying the right level of ToM used by the opponent.

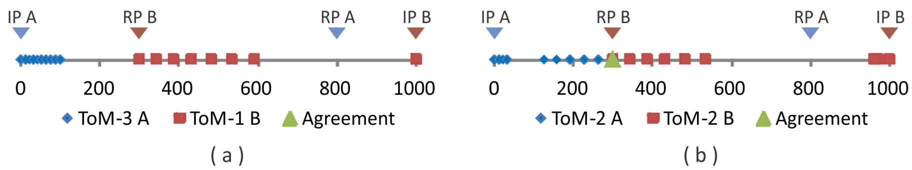

Fig. 8: Negotiation episodes conducted in context 5, where a ToM agent fails in identifying the ToM's level of its opponent. In (a), ToM-3 A fails to model its opponent as a ToM-1. In (b), both agents A and B fail to model their opponent's ToM's level.

Despite some drawbacks, the correct use of ToM can help a negotiation agent to increase its utilities, as it is proven in this work. Further research is needed in order to overcome the mentioned problems, but the simulation results highlight that contextual models and ToM can be profitably integrated with Bayesian inference in automated negotiations.

6 Future work

Besides overcoming the problems mentioned above, there are a number of research lines to follow from now on. To begin with, the use of ToM could be applied in series of negotiation episodes rather than in a single negotiation episode. Keeping a history of past negotiations could help agents to better capture the level of ToM of other parties, as well as their true necessities and risks.

Another possible line of research is the use of ToM to explain why a certain deal or proposal occurs. This opens new opportunities to make automated negotiations richer and gives agents a cognitive ability with the possibility to explain their actions and own behaviour.

Finally, this work could be extended to contemplate man-machine negotiations in order to understand why people decide to bargain or not in different contexts and under various situations.

References

1. T. Baarslag. *Exploring the strategy space of negotiating agents: A framework for bidding, learning and accepting in automated negotiation.* Springer theses. Springer, Switzerland, 2016.

2. C. L. Baker, J. Jara-Ettinger, R. Saxe, and J. B. Tenenbaum. Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, 1(4):598, 2017.
3. S. de Jong, D. Hennes, K. Tuyls, and Y. Gal. Meta-strategies in the colored trails game. *Belgian/Netherlands Artificial Intelligence Conference*, 1-3(c):2–6, 2011.
4. H. de Weerd, R. Verbrugge, and B. Verheij. How much does it help to know what she knows you know? an agent-based simulation study. *Artificial Intelligence*, 199-200:67–92, 2013.
5. S. Fatima, S. Kraus, and M. Wooldridge. *Principles of Automated Negotiation*. Cambridge University Press, 2014.
6. Y. Gal, B. Grosz, S. Kraus, A. Pfeffer, and S. Shieber. Agent decision-making in open mixed networks. *Artificial Intelligence*, 174(18):1460–1480, 2010.
7. J. Joiner, M. Piva, C. Turrin, and S. W. C. Chang. Social learning through prediction error in the brain. *npj Science of Learning*, 2(1):8, 2017.
8. F. Lopes, M. Wooldridge, and A. Q. Novais. Negotiation among autonomous computational agents: Principles, analysis and challenges. *Artificial Intelligence Review*, 29(1):1–44, 2008.
9. M. J. Osborne and A. Rubinstein. *Bargaining and markets*. Economic theory, econometrics, and mathematical economics. Academic Press, San Diego, 1990.
10. J. Pöppel and S. Kopp. Satisficing models of bayesian theory of mind for explaining behavior of differently uncertain agents. *Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), Stockholm, Sweden.*, 1(July), 2018.
11. S. Qi and S.-C. Zhu. Intent-aware multi-agent reinforcement learning. *CoRR*, abs/1803.02018, 2018.
12. N. C. Rabinowitz, F. Perbet, H. F. Song, C. Zhang, S. M. A. Eslami, and M. Botvinick. Machine theory of mind. *ArXiv e-prints*, 2018.
13. T. Revell. Deepmind ai is learning to understand the 'thoughts' of others. *New Scientist*, 1(February):1–5, 2018.
14. J. Rodriguez-Fernandez, T. Pinto, F. Silva, I. Praça, Z. Vale, and J. M. Corchado. Context aware q-learning-based model for decision support in the negotiation of energy contracts. *International Journal of Electrical Power and Energy Systems*, 104(October 2017):489–501, 2019.
15. I. Sher, M. Koenig, and A. Rustichini. Children's strategic theory of mind. *Proceedings of the National Academy of Sciences*, 111(37):13307–13312, 2014.
16. T. Singer and A. Tusche. Understanding others: Brain mechanisms of theory of mind and empathy. *Neuroeconomics: Decision Making and the Brain*, pages 513–534, 2014.
17. C. A. Stevens, N. A. Taatgen, and F. Cnossen. Instance-based models of metacognition in the prisoner's dilemma. *Topics in Cognitive Science*, 8(1):322–334, 2016.
18. F. B. von der Osten, M. Kirley, and T. Miller. The minds of many: Opponent modelling in a stochastic game. *IJCAI International Joint Conference on Artificial Intelligence*, pages 3845–3851, 2017.
19. W. Yoshida, R. J. Dolan, and K. J. Friston. Game theory of mind. *PLoS Computational Biology*, 4(12), 2008.
20. D. Zeng and K. Sycara. Bayesian learning in negotiation. *International Journal of Human Computer Studies*, 48(1):125–141, 1998.