Bayesian Theory of Mind for False Belief Understanding in Human-Robot Interaction

Mehdi Hellou*, Samuele Vinanzi and Angelo Cangelosi
Cognitive Robotics Lab,
School of Computer Science,
The University of Manchester

Correspondence: mehdi.hellou@manchester.ac.uk

***Abstract*—In order to achieve a widespread adoption of social robots in the near future, we need to design intelligent systems that are able to autonomously understand our beliefs and preferences. This will pave the foundation for a new generation of robots able to navigate the complexities of human societies. To reach this goal, we look into Theory of Mind (ToM): the cognitive ability to understand other agents' mental states. In this paper, we rely on a probabilistic ToM model to detect when a human has false beliefs with the purpose of driving the decision-making process of a collaborative robot. In particular, we recreate an established psychology experiment involving the search for a toy that can be secretly displaced by a malicious individual. The results that we have obtained in simulation show that the agent is able to predict the mental states of the humans and detect when false beliefs have arisen. We plan to expand these findings to a real-world human-robot interaction setting.**

I. Introduction

As autonomous robots become more prevalent in our daily lives, it is important for them to be capable to adapt to a variety of social situations. A social robot is an intelligent agent specifically designed to operate in human environments, to interact with people and to adapt its behavior to their partner's needs, preferences and personality. The emphasis of the robot's ability to adapt to different users is often known as "personalization" [10]. The latter has been proven to enhance user engagement in long-term human-robot interaction (HRI) and to foster rapport and trust for tasks such as education, rehabilitation and elderly care [10, 11].

The aim of our study is to design an artificial cognitive architecture for autonomous robots that is able to personalize its behavior based on the user's mental states. In order to do so, we tap into the domain of psychology to computationally model a cognitive skill known as Theory of Mind (ToM): this is defined as the ability to infer other's mental states, such as beliefs, desires and intentions (often known as BDI), in order to predict behavior [14]. ToM is largely studied in the psychological literature, especially for the purpose of understanding the cognitive development of infants and how they perceive the world around them [5, 7].

Several experiments and procedures have been proposed, over the years, to assess ToM abilities in infants. One of the better-known tests is the "false belief understanding", which has been largely used to evaluate whether preschoolers can understand people's mental states, in particular their beliefs for the purpose of action anticipation. More specifically, some of these tests aim to evaluate whether a child could understand when a person has a belief that contradicts reality [17, 3].

In this paper, we present an artificial intelligence system that is able to detect false beliefs. This is a critical skill to possess for a social agent involved in collaborative tasks or caring for elderly people in a retirement homes. As an example, we could benefit from this ability in a robot that keeps track of some medications to prevent the patient from taking the wrong ones. We take inspiration from a human-human experiment involving a toy swapping game [4] to evaluate if the robot can detect when a user has a false belief understanding and determine the best collaborative course of action.

II. Previous work

Most of the literature about ToM comes from the psychological domain, but there has also been a growing interest from the fields of robotics and artificial intelligence. One of the most popular techniques to computationally design ToM-capable agents is the use of Bayesian Networks (BN), a graphical model for data analysis which can encode uncertainty in expert systems [9]. This kind of model can easily represent the knowledge and learning of infants by using a causal map: an abstract, coherent, learned representation of the causal relations among events [8, 6]. This approach was adopted by Vinanzi et al. [16], who have developed a robot learning architecture based on BNs able to estimate the trustworthiness of human partners based on the understanding of their mental states. Another relevant example comes from Baker et al. [1, 2], who implemented a dynamical BN known as "Bayesian model of ToM" (BToM), which uses Bayesian inverse planning to represent how people infer other's goals and preferences. This model is combined with Partially Observable Markov Decision Processes (POMDP) to represent the agent's planning and inference about the world. The model then uses Bayesian inference to invert the planning and reconstruct the agent's joint belief state and reward function, conditioned on the observations of the agent's behavior in some environmental context.

A different approach involves the use of neural networks. Rabinowitz et al. [15] created ToMmet, which uses deep learn-

ing methods to learn about a family of POMDPs. Patacchiola et al. [13] implemented a cognitive architecture for trust and ToM in humanoid robots that relies on self-organizing maps [12].

III. Methodology

A. False belief understanding in a collaborative task

Our objective for this experiment is to implement a computational model of ToM that will allow a social robot to understand when someone has a false belief and to help them in the task at hand. To do so, we wish to replicate a psychological experiment by Buttelman et al. [4]. The latter is based on a setting which includes two human actors (the searcher and the tricker), two boxes, a toy and a child who is observing the scene. During the experiment, the searcher places the toy in one of the boxes and right after the tricker attempts to switch its position from the original container to the other one. This can happen either in the presence of the original actor, who is then aware of the swap (true belief condition), or in their absence (false belief condition). At this point, the searcher would approach the box in which they had originally placed the toy and unsuccessfully attempted to open it. The child was then asked to provide some form of help.

The results of the original experiments showed that, in the true belief condition, the children would help the searcher to open the box they committed to, even though the toy was not there anymore. The participants figured out that the adult had acknowledged the switch and that they wanted to open the box for some other reason. In the false belief condition, the children would open the box in which the toy had been switched to, demonstrating their understanding of the searcher's false belief about the location of the object.

Our aim is to replicate the same experiment, substituting an autonomous humanoid robot in place of the child. Our artificial agent's responsibilities will be, then, to infer the mental states of the searcher to detect false beliefs and to perform appropriate collaborative actions.

B. ToM cognitive architecture

In order to implement the same level of ToM in a robot, we need a model capable of reasoning on the belief and desires of an agent based on the observation of its actions. In the context of this experiment, the robot needs to know where the person believes the toy is located and what is their preference regarding the boxes: are they interested in retrieving the object or are they keen on opening the other box? We decided to use BToM [1, 2] which uses Bayesian inverse planning to represent how people infer other's goals or preferences. More specifically, the model exploits POMDPs to represent how an agent behaves in an environment regarding its beliefs and preferences about the world via the principle of rational belief, which formalizes how the agent's belief is affected by observations in terms of Bayesian belief updating. The model represents respectively the agent's desires as an utility function and the agent's own subjective beliefs about the environment as a probability distribution, which may be uncertain and may differ from reality. As a result, the model is designed as a dynamic BN (DBN) to symbolize how external and internal elements, such as agent's location, observations, preferences and beliefs, can influence the agent's behaviors over time to complete a specific task. The process to compute the observer's belief and reward inference of the agent, is similar to the "forward-backward" algorithm in hidden Markov models.

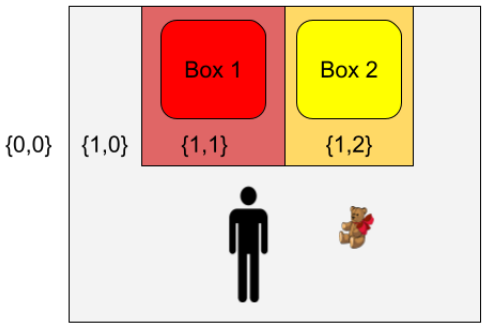

Fig. 1: Experimental setting. The labels represent the different possible states in which the human and he object can be.

This model was originally designed for an agent searching for its preferred food truck in a 2D environment. One of our contributions will be to adapt this model to perform its tasks in a more complex scenario that involves dynamical elements (such as the position of the humans and the toy).

C. Simulation

We performed a simulated experiment to analyze the performance of our model before deploying it on a real robotic platform. The simulation included the environment as depicted in the original paper [4] wherein a human, a toy and two boxes were present (Figure 1). The actor is able to enter and leave the room and move to one of the boxes, while the toy can either be placed outside the boxes or in one of them.

To evaluate the predictions generated by our model, we have randomly generated sequences of behaviors for the simulated human agent. This process is conducted in two steps: initialization and generation. The initial phase is the same for every iteration: the agent is outside the room, and the toy is outside the boxes. The human moves into the room, takes the toy and places it in one of the boxes. The second phase depends on a set of parameters:

The "rate of false belief" R_{fb} 2 [0, 1] determines the probability that the human agent will leave or not the room before the toys are switched, i.e. the rate of true or false belief instances.

The "alternate false belief" A_{fb} 2 $[True, False]$ instructs the generator to change successively or randomly the belief condition between iterations.

The "rate of preference" R_p 2 [0, 1] represents the probability that the human is interested or not in retrieving the toy.

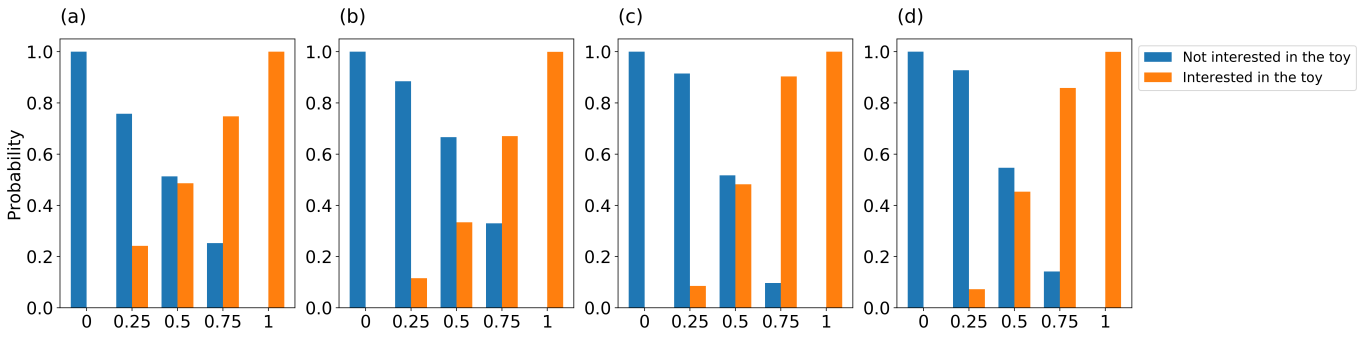

Fig. 2: Human agent's preference prediction, obtained by fixating $R_{fb} = 0.5$ and varying R_p. (a) $A_p = True$, closed. (b) $A_p = True$, open. (c) $A_p = False$, closed. (d) $A_p = False$, open.

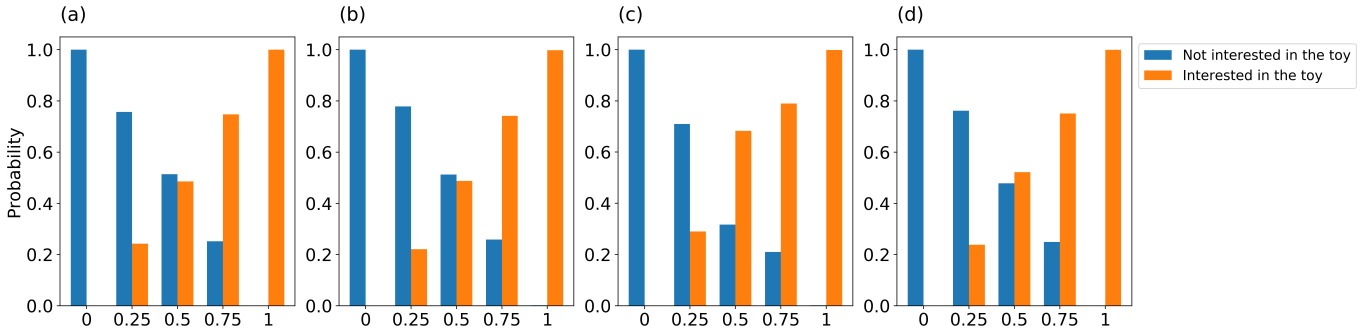

Fig. 3: Human agent's preference prediction, obtained by fixating $R_{fb} = 0.5; A_{fb} = True; R_p = 0.5; A_p = True$ and closing the model. Each graph is obtained by setting the rewards as displayed in Table I.

		Rewards			
Preference	**Action**	**a**	**b**	**c**	**d**
Interested in the toy	**Go to the box with the toy**	10	10	50	100
	Go to the box without the toy	-10	5	25	25
Not interested in the toy	**Go to the box with the toy**	-10	5	25	25
	Go to the box without the toy	10	10	50	10

TABLE I: Rewards for the policies generated by the model to infer the preferences of the agent.

The "alternate preference" A_p 2 $[True, False]$ instructs the generator to flip successively or randomly the preference of the human agent between iterations.

The generation process is described by following procedure:

1) Determine the belief condition (true or false) and the human's preference (interested in the toy or not).
2) The human moves back to the initial position (state $f1,0g$).
3) If the agent is in the false belief condition, it leaves the room (state $f0,0g$). If, instead, it is in the true belief condition, it stays in the room (state $f1,0g$).
4) The toy's position is randomly switched or preserved.
5) If the agent is outside, it re-enters the room.
6) The agent moves to box 1 (state $f1,1g$) or 2 (state $f1,2g$) according to its current belief and its preference: if it is interested in the toy, it will move to the box where it believes the toy is located; if it is not the case, then it will move to the box where it believes the toy is not located.

We define the size of the path S_{path} as the number of times in which the agent approaches the boxes during the generation phase, i.e. the number of times the above procedure is reiterated. This enables us to assess the model's performance over time and determine whether it can accurately track beliefs and preferences throughout several iterations. It is mostly important regarding the need of using social robots for long-term interactions and that can be aware of the mental state of human over the time.

IV. Results

To evaluate the performance of the model, we generate a number of different paths for the human agent by varying the associated parameters. For each of the trials described below, we generated 500 paths with $S_{path} = 8$.

Some of the following evaluations are calculated only when the human is located next to the boxes (states $f1,1g$ and $f1,2g$ in Figure 1. We call this the "closed" condition. If, on the contrary, the evaluation can happen at any point of the path, we say that we are evaluating the model in the "open" condition.

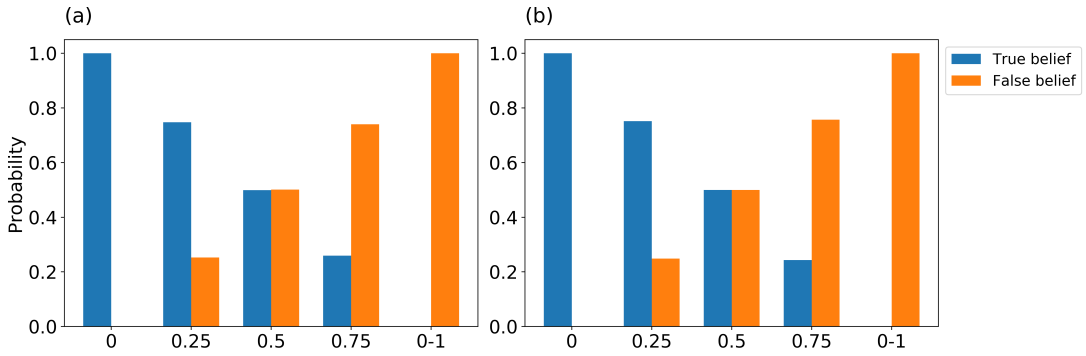

Fig. 4: Human agent's belief prediction, obtained by fixating R_p = 0.5 and varying R_{fb}. (a) A_{fb} = *True*. (b) A_{fb} = *False*.

A. Evaluation of the human's preferences

To analyze if the model was able to correctly identify the simulated human's preferences, we have set R_{fb} = 0.5 and varied R_p between values of 0, 0.25, 0.5, 0.75, and 1. For example, a rate of preferences of 0.5 means that 50% of the time the agent will move to the box wherein it believes the toy is located, while the remaining 50% of the time it will go to the other box.

The evaluation of the agent's preferences is shown in Figure 2. The four charts are produced by varying A_p and by opening and closing the model. For example, Figure 2a reflects the condition where the model is closed and A_p = $True$, Figure 2b shows the performance measured when the model is closed and A_p = $False$ and so on. As described by the graphs, the model can infer the agent's preferences corresponding to the ratio's values we gave, even when we set to a rate equal to 0.5, wherein the agent's behavior will change half of the time. However, the model has a lower performance when inferring the agent's preferences by including all the states. Setting the ratio to 0.5, the model inferred that the agent is interested in the toy less than 40% of the time and close to 40% when we do not alternate the preferences. We can explain that the model is unsure about the agent's preferences when this one is not close to one of the boxes; otherwise, it can predict the preference regarding which boxes it is going to visit. Figure 3 represents the model performances with different rewards employed to compute agent's policies. There is not much impact on model performances when using different rewards, as we can notice on the bar charts, except for Figure 3c, wherein the model predicted that the agent is majorly interested in the toy when setting the ratio to 0.5. This result may indicate that the variation between the rewards is too important, compared to the reward in Figure 3b, or not that significant such as the reward in Figure 3d. We can only observe this inconsistency with R_p = 0.5, whereas for other values, we observe the same pattern as with the other Figures.

B. Evaluation of the human's beliefs

For the belief's inference, we have set R_p = 0.5 and varied R_{fb} between values of 0, 0.25, 0.5, 0.75, and 1. The results depicted in Figure 4 show us that the model can accurately infer the belief of the human agent regarding the value of R_{fb}. For example, when R_{fb} equals 0.5, which correspond to a change of belief condition half of the time, the model can infer that approximately 50% of the time, the agent has false beliefs understanding, and the remaining time, he has true beliefs.

The current preliminary results are promising and they validate the performance of our model in simulation. Indeed, our model can accurately infer the agent's preferences and its beliefs even when we increase the variation of the agent's behaviour.

V. Conclusion

In this paper, we have presented our preliminary results on a ToM-capable cognitive architecture for HRI. We have discussed on the importance of false belief understanding for collaborative agents and we have traced an arc that shows how this mental skill is studied in both the psychological and the computational domains. Finally, we have introduced BToM [1, 2], a DBN that was originally used to determine the desires of an agent in a 2D world. We have applied this model to a more complex environment involving an established psychology experiment used to test false belief understanding in children [4]. Finally, we have conducted simulated tests to evaluate the fitness of this model for our purposes and the data we have collected are in line with our expectations. The next step in this line of research will be to embed this model in a social robot involved in a real-wold replication of the false belief understanding experiment and to evaluate its ability to correctly help its human partners in joint action scenarios.

Acknowledgments

The PERSEO project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 955778.

References

[1] Chris L. Baker, Rebecca R. Saxe, and Joshua B. Tenenbaum. Bayesian theory of mind: Modeling joint belief-desire attribution. In *In Proceedings of the Thirtieth Third Annual Conference of the Cognitive Science Society*, pages 2469–2474, 2011.

[2] Chris L. Baker, Julian Jara-Ettinger, Rebecca Saxe, and Joshua B. Tenenbaum. Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, 1, 2017.

[3] Simon Baron-Cohen, Alan M. Leslie, and Uta Frith. Does the autistic child have a "theory of mind" ? *Cognition*, 21(1):37–46, 1985. ISSN 0010-0277. doi: https://doi.org/10.1016/0010-0277(85)90022-8. URL https://www.sciencedirect.com/science/article/pii/0010027785900228.

[4] David Buttelmann, Malinda Carpenter, and Michael Tomasello. Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition*, 112(2):337–342, August 2009. ISSN 0010-0277. doi: 10.1016/j.cognition.2009.05.006.

[5] John H. Flavell. Cognitive development: children's knowledge about the mind. *Annual review of psychology*, 50:21–45, 1999.

[6] Noah D. Goodman, Chris L. Baker, Elizabeth Baraff Bonawitz, Vikash K. Mansinghka, Alison Gopnik, Henry Wellman, Laura Schulz, and Joshua B. Tenenbaum. Intuitive theories of mind: a rational approach to false belief. In *Proceedings of the Twenty-Eigth Annual Conference of the Cognitive Science Society. Mahwah, NJ: Erlbaum*, 2006.

[7] Alison Gopnik and Henry M. Wellman. *The theory theory*, page 257–293. Cambridge University Press, 1994. doi: 10.1017/CBO9780511752902.011.

[8] Alison Gopnik, Clark Glymour, David M. Sobel, Laura E. Schulz, Tamar Kushnir, and David Danks. A Theory of Causal Learning in Children: Causal Maps and Bayes Nets. 1 2004. doi: 10.1184/R1/6490802.v1. URL https://kilthub.cmu.edu/articles/journal_contribution/A_Theory_of_Causal_Learning_in_Children_Causal_Maps_and_Bayes_Nets_/6490802.

[9] David Heckerman, Dan Geiger, and David M. Chickering. Learning bayesian networks: The combination of knowledge and statistical data. In Ramon Lopez de Mantaras and David Poole, editors, *Uncertainty Proceedings 1994*, pages 293–301. Morgan Kaufmann, San Francisco (CA), 1994. ISBN 978-1-55860-332-5. doi: https://doi.org/10.1016/B978-1-55860-332-5.50042-0. URL https://www.sciencedirect.com/science/article/pii/B9781558603325500420.

[10] Mehdi Hellou, Norina Gasteiger, Jong Lim, Minsu Jang, and Ho Ahn. Personalization and localization in human-robot interaction: A review of technical methods. *Robotics*, 10:120, 11 2021. doi: 10.3390/robotics10040120.

[11] Bahar Irfan, Aditi Ramachandran, Samuel Spaulding, Dylan F. Glas, Iolanda Leite, and Kheng Lee Koay. Personalization in long-term human-robot interaction. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 685–686, 2019. doi: 10.1109/HRI.2019.8673076.

[12] Anthony F. Morse, Joachim de Greeff, Tony Belpeame, and Angelo Cangelosi. Epigenetic robotics architecture (era). *IEEE Transactions on Autonomous Mental Development*, 2(4):325–339, 2010. doi: 10.1109/TAMD.2010.2087020.

[13] Massimiliano Patacchiola and Angelo Cangelosi. A developmental cognitive architecture for trust and theory of mind in humanoid robots. *IEEE Transactions on Cybernetics*, 52(3):1947–1959, 2022. doi: 10.1109/TCYB.2020.3002892.

[14] David Premack and Guy Woodruff. Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1(4):515–526, 1978. doi: 10.1017/S0140525X00076512.

[15] Neil C. Rabinowitz, Frank Perbet, H. Francis Song, Chiyuan Zhang, S. M. Ali Eslami, and Matthew Botvinick. Machine theory of mind. *CoRR*, abs/1802.07740, 2018. URL http://arxiv.org/abs/1802.07740.

[16] Samuele Vinanzi, Angelo Cangelosi, and Christian Goerick. The collaborative mind: intention reading and trust in human-robot interaction. *iScience*, 24(2):102130, 2021. ISSN 2589-0042. doi: https://doi.org/10.1016/j.isci.2021.102130. URL https://www.sciencedirect.com/science/article/pii/S2589004221000985.

[17] Henry M. Wellman, David Cross, and Julanne Watson. Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, 72(3):655–684, 2001. ISSN 00093920, 14678624. URL http://www.jstor.org/stable/1132444.