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

HELLOU, Mehdi, VINANZI, Samuele and CANGELOSI, Angelo

Citation:

HELLOU, Mehdi, VINANZI, Samuele and CANGELOSI, Angelo (2023). Bayesian Theory of Mind for False Belief Understanding in Human-Robot Interaction. In: 2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN). IEEE, 1893-1900. [Book Section]

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

Mehdi Hellou*[1], Samuele Vinanzi[1,2] and Angelo Cangelosi[1]

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 simulated experiments show that the agent is able to predict human mental states and detect when false beliefs have arisen. We then explored the set-up in a real-world human interaction to assess the feasibility of such an experiment with a humanoid social robot.**

I. INTRODUCTION

As autonomous robots become more prevalent in our daily lives, they need to be capable of adapting to a variety of social situations. In particular, society has started taking an interest in social robots, which are intelligent agents specifically designed to operate in human environments, interact with people, and adapt their behavior to their partners' needs, preferences, and personalities. The emphasis on the robot's ability to adapt to different users is often known as "personalization" [1]. 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 [1], [2].

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. To do so, we tap into the domain of psychology to computationally model a cognitive skill known as Theory of Mind (ToM), which is defined as the ability to infer others' mental states, such as beliefs, desires, and intentions (often known as BDI), to predict behavior [3]. It is a cognitive process we unconsciously practice to understand other people's behavior and actions in the environment.

Our ability to understand and predict the actions of others is closely linked to our visual perception of their movements. For instance, if a restaurant customer suddenly stands up from their table and heads toward the kitchen, an observer may infer that the person is either seeking the restroom or some additional service, such as the bill. But if they also notice that there are still plates on the table and the customer seems a little lost, they are likely to infer that the person is looking for the restroom and may offer directions to help. Such cognitive processes illustrate the human capacity to comprehend and anticipate other people's actions, based on their mental states, including beliefs. This is why ToM has been extensively studied in psychology, particularly for understanding the cognitive development of infants and their understanding of others' minds [4], [5].

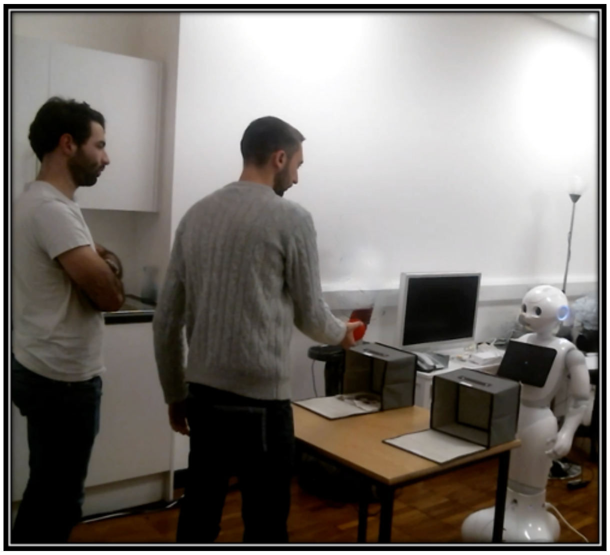

Fig. 1. Lab experiment setting with the boxes and the robot that will be deployed to infer a person's mental state. Here, one person is manipulating a toy the robot is tracking while another is standing close to him.

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 can understand when a person has a belief that contradicts reality [6], [7]. A famous experiment, still used at this time as a test to evaluate ToM with children with Autism Spectrum Disorder, is the "Sally-Anne" test from Baron-Cohen et al. [7] that evaluates whether the child can detect when people have false beliefs concerning the location of the toy.

This paper presents an artificial intelligence system capable of detecting false beliefs, a critical skill for social agents involved in collaborative tasks or caring for elderly people in retirement homes. For instance, a robot keeping track of a

* mehdi.hellou@postgrad.manchester.ac.uk
[1] Manchester Centre for Robotics and AI, University of Manchester
[2] Department of Computing, Sheffield Hallam University

patient's medication could benefit from this ability to prevent the patient from taking the wrong medication. Our approach is inspired by a human-human experiment involving a toy swapping game [8], where we evaluate the robot's ability to detect when a user has a false belief and determine the best collaborative course of action. To achieve this, we integrated the Bayesian theory of mind (BToM) framework developed in [9], [10] and modeled it in a simulated environment of the toy swapping game. We then conducted an HRI experiment replicating similar situations in [8] to analyze how a robot endowed with a ToM reasoning would respond (Figure 1).

The goal of our study is to demonstrate the efficacy of using a validated model for reasoning about people's mental states to improve robot behaviors in human-robot interaction (HRI) situations. By taking inspiration from a specific false beliefs experiment, we show how a robot endowed with theory of mind (ToM) abilities can reason about people's behaviors. We conducted two experiments to test the applicability of our cognitive model in HRI situations and to determine whether it could be implemented in a real interaction with a social robot. We analyzed the results of these experiments quantitatively through simulations and interactively through an HRI experiment to determine if they were consistent. Overall, our study aims to contribute to the development of more effective and adaptive social robots through the use of ToM reasoning.

II. Related works

A. *Psychological studies in ToM*

ToM has been extensively studied in the field of psychology, with a particular focus on children's understanding of people's mental states. Several researchers have specified that children undergo cognitive development that enables them to understand other people's mental states, such as beliefs, desires, emotions, and intentions [11]. Psychologists have investigated the years in which infants started developing their ToM and the ways in which it manifests [11], [12], [13], [14], [6]. Various tasks have been provided to evaluate this phenomenon in children, but one of the main measurements used is the ”false belief understanding.” In particular, some tests aim to evaluate whether a child can understand when a person has a belief that contradicts reality [6], [7]. Researchers have observed the reactions of children in false belief situations to assess how they understand other people's mental states such as beliefs [12] and their impact on other mental states, such as desire [13]. However, earlier studies primarily indicated that infants were aware that something strange had occurred during the experiments, rather than clearly demonstrating their understanding of false beliefs. More recent experiments have implemented tasks that require active behavior from children to measure their understanding of these particular situations, involving their active participation [8], [15], [16].

B. *ToM applied in robotics and AI*

Although most research on ToM has been conducted in the field of psychology, there is growing interest in using ToM principles to develop autonomous and intelligent systems in the fields of computer science and robotics. One popular technique for designing ToM-capable agents is the use of Bayesian Networks (BN), which are graphical models for data analysis and a popular tool for encoding uncertain expert knowledge in expert systems [17]. This type of model is well-suited for representing the knowledge and learning of infants, who are thought to represent the world by constructing a causal map: an abstract, coherent, learned representation of the causal relations among events [18], [19].

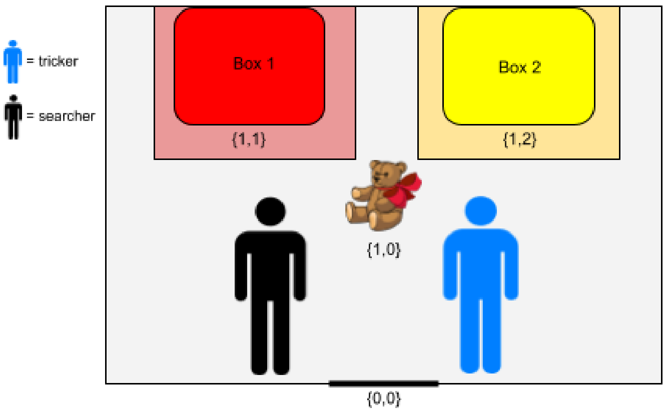

Fig. 2. Simulation setting. The labels denote the different possible states of the agents and the object in the experiment.

For example, Vinanzi et al. [20] have developed a robot learning architecture based on BNs that is able to estimate the trustworthiness of human partners based on an understanding of their mental states. Another notable example comes from Baker et al. [9], [10], who implemented a dynamical BN as the "Bayesian model of ToM" (BToM), which uses Bayesian inverse planning to represent how people infer others' goals or preferences. This model is combined with a partially observable Markov decision process (POMDP) to represent an 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 observations of the agent's behavior in some environmental context.

Other methods for incorporating ToM into autonomous systems exist in the literature. For instance, Patacchiola et al. [21] developed a cognitive architecture for trust and ToM in humanoid robots using an actor-critic framework [22]: a model called ERA [23] which utilizes self-organizing maps as function approximators, and a BN to represent the intrinsic values of the robot's environment

III. Proposed Method

A. *BToM for false beliefs understanding in HRI*

The BToM model, first implemented by Baker et al. [9] and used in other social scenarios [10], is a probabilistic model that leverages Bayesian inverse planning to represent how people infer others' goals or preferences. It is the first model in the literature to reason jointly about beliefs, desires, and perceptions by observing an agent's behaviors. Designed as a dynamic Bayesian network (DBN), it symbolizes how external and internal elements, such as the agent's location, observations, preferences, and beliefs, can influence the agent's behaviors over time to complete a specific task.

To represent how an agent behaves in an environment regarding its beliefs and preferences about the world via the principle of rational belief, the model uses partially observable Markov decision processes (POMDPs). In following the schematic model of ToM proposed by psychological researchers [24], [9], [10], the BToM follows three causal principles to explain an agent's core mentalizing, or how an agent reasons about other agents' behavior: (1) the agent uses its perception or what is in its line of sight to constitute its world model; (2) the agent builds its beliefs regarding the combination of its visualizations and prior knowledge; and (3) the agent plans a rational sequence of actions that, given its beliefs, is expected to achieve its desires.

The model represents the agent's desires as a utility function, which measures the subjective rewards received for taking actions in certain states, such as making decisions when close to the goal object. The agent's beliefs about the environment are represented as a probability distribution, which may be uncertain and different from reality. The contents of the distribution represent a possible world that the agent believes corresponds to reality. For instance, the agent's beliefs may be shared between different possible worlds if it is unsure about the contents of the environment, such as the exact location of an object.

Following Baker et al. [10], we modeled our false beliefs experiment as depicted in Figure 2. The experiment involves two humans, a toy, and two boxes, and the model focuses on inferring the blue agent's mental state, who plays the role of the searcher, and being tricked by the other agent (black one), who swaps the toy in his presence or not. We utilized POMDPs to represent the environment of the agent, where the state space $S = \{X, Y\}$ represents the set of agent's locations X and the possible location of the toy, Y. As previously explored in Baker's papers [9], [10], the world is dynamic and subject to change over time, i.e., y_t represents a possible world y at time t. The agent can take different actions a from a set of actions A to move into the environment, for example, to visit one of the boxes, and have observations o from the set of observations O, such as whether the agent can visualize the toy's location. Finally, a set of observation probabilities Ω represents the observations the agent can make regarding its location x and the world y, for instance, the agent can observe if the toy is located in the yellow box and if it is close to the box, i.e., state {1,2}.

Overall, the process of inferring the joint beliefs and desires is divided into two steps: 1) updating the agent's beliefs based on prior knowledge; and 2) performing joint inference on the posterior probability of the agent's unobservable mental states (beliefs, desires, and visualizations). Specifically, the belief that a world y_t is true at time t is denoted as $b_t(y_t)$, given its prior belief $b_{t-1}(y_{t-1})$, the likelihood $P\left(o_t \mid x_t, y_t\right)$ of observing the agent state x_t and the world y_t, and the probability $P\left(x_t, y_t \mid x_{t-1}, y_{t-1}, a_{t-1}\right)$ of observing the agent move from position x_{t-1} to x_t, and the world changing from y_{t-1} to y_t given the action a_{t-1}. The Bayesian belief update for $b_t(y_t)$ is given by:

$$
\begin{aligned}
b_t(y_t) \propto\, & P\left(o_t \mid x_t, y_t\right) P\left(x_t, y_t \mid x_{t-1}, y_{t-1}, a_{t-1}\right) \\
& b_{t-1}(y_{t-1})
\end{aligned}
$$

To infer joint beliefs and desires, the model utilizes a method similar to the belief filtering technique proposed in [25], adapted for the scenario of one agent reasoning about the behavior of another agent. In this approach, the observer agent's inference about the belief and desire of the other agent is akin to the backward-forward algorithm used in Hidden Markov Models. Specifically, given a state sequence up to a time T, the joint belief-desire at any time $t \leq T$, denoted as $P\left(b_t, d_t \mid x_{1:T}, y_{1:T}\right)$, can be computed.

B. Robot implementation

This section details an artificial cognitive architecture designed to enable social robots to reason about the preferences and false beliefs of other agents. We developed this system specifically for the Pepper robot from Aldebaran, a humanoid platform created for human interaction. Figure **??** provides an overview of the cognitive architecture.

The Vision Module serves as the interface between the robot and the external environment, and is responsible for collecting sensory information through the robot's RGB cameras. It includes the following components:

- **Face recognition**: This component is based on a Local Binary Patterns Histograms (LBPH) [26] model trained on a handmade face dataset. Its purpose is to classify the human currently interacting with the robot as the searcher or the tricker.
- **Toy recognition**: This component recognizes the toy and determines its location within the world. We use a red plastic ball and a variety of methods provided by the robot's API.
- **Box recognition**: This component keeps track of the two boxes, which are marked with special markers that assist the robot in identifying them.
- **People recognition**: This component uses the YOLOv4 neural network [27] to determine the number of people present in the robot's field of view. It is used to understand if the searcher is in the true or false belief condition (more details on the latter in Section IV-B).

The Cognitive Module gathers all the information provided by the Vision Module and uses it to construct a plan for the human, which is then analyzed to infer their beliefs and preferences. This is accomplished through the BToM model presented in Section III-A. The output of this module is forwarded to the Decision-Making Module, which is responsible for selecting appropriate actions.

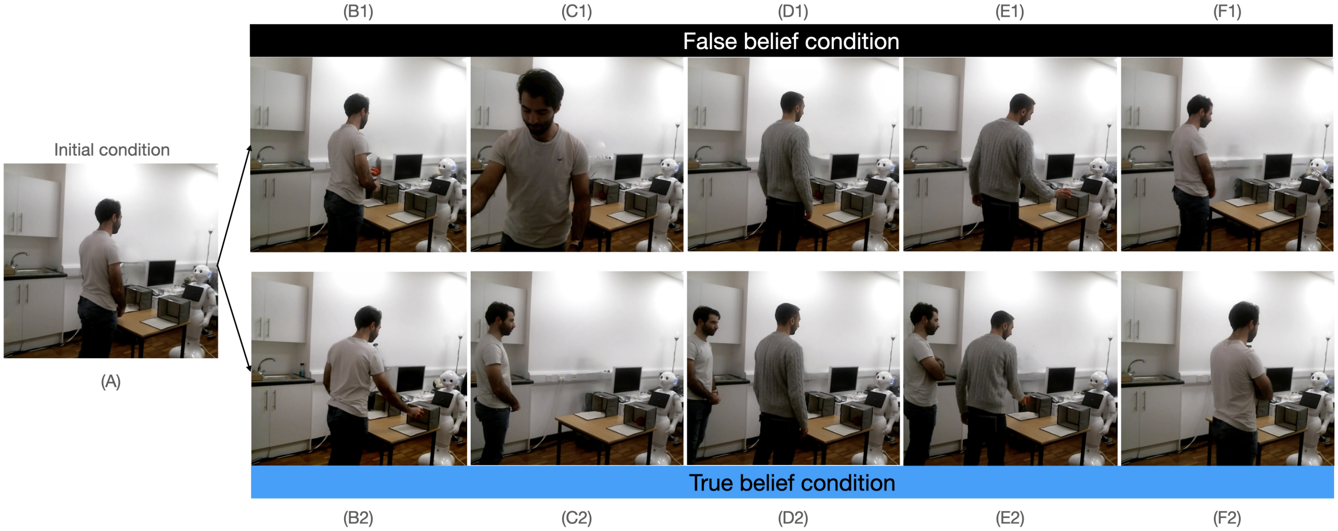

Fig. 3. HRI experiment involving two human actors, a searcher (wearing a white shirt), a tricker (wearing a grey shirt), and the social robot Pepper with a total view of the scene. The mosaic of pictures depicts the two experiments we set up, where pictures A and B correspond to the initial condition, and pictures C to F correspond to the rest of the experiments under different conditions. The top pictures are related to the false belief condition, and the bottom ones correspond to the true belief condition (refer to Section IV-B).

IV. Experiments

A. Simulation

As a preliminary evaluation of our architecture, we conducted a simulation experiment to assess the performance of the Cognitive Module under conditions similar to those in which the robot would be subsequently tested. Following the methodology of [8], the simulation included an environment with two humans (a searcher $\mathcal{S}$ and a tricker $\mathcal{T}$), a toy and two boxes b_1 and b_2. This is depicted in Figure 2). In the simulation, $\mathcal{S}$ is capable of entering and leaving the room and moving to one of the boxes, while $\mathcal{T}$ has the ability to place the toy outside the boxes or in one of them.

To evaluate the predictions generated by our model, we randomly generate sequences of behaviors for $\mathcal{S}$ and $\mathcal{T}$ during trials. This process is conducted in two steps: "initialization" and "generation". The initialization phase is the same for every trial: $\mathcal{S}$ is outside the room, $\mathcal{T}$ is inside the room, and the toy is outside the boxes. Then, $\mathcal{S}$ 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" $R_{fb} \in [0,1]$ determines the probability that $\mathcal{S}$ will leave or not the room before $\mathcal{T}$ switches the toy's location, i.e. the rate of true or false belief instances.
- The "alternate false belief" $A_{fb} \in [True, False]$ instructs the generator to alternate the belief condition between iterations in a sequential or random manner.
- The "rate of preference" $R_p \in [0,1]$ represents the probability that the human is interested or not in retrieving the toy.
- The "alternate preference" $A_p \in [True, False]$ instructs the generator to flip the preference of the human agent between iterations either successively (i.e., True) or randomly (i.e., False).

The generation process is described by following procedure:

1) Set the belief condition (true or false) and the human's preference (interested in the toy or not) regarding the values of the parameters described above.
2) The human moves back to the initial position (state {1,0}).
3) If the agent is in the false belief condition, it leaves the room (state {0,0}). If, instead, it is in the true belief condition, it stays in the room (state {1,0}).
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 b_1 (state {1,1}) or b_2 (state {1,2}) 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.

The set of parameters above contextualizes each trial, depending on how we define the size of the path S_{path}, which represents the number of times $\mathcal{S}$ will reach boxes $b1$ and $b2$ to retrieve the toy or not. This allows us to assess the model's performance over time and determine whether it can accurately track beliefs and preferences throughout multiple iterations. This is particularly important for using social robots in long-term interactions where they need to be aware of the human's mental state over time.

For instance, suppose we have a trial of size $S_{path} = 8$ with the following parameter values: $R_{fb} = 0.75,\ A_{fb} = True,\ R_p = 0.5,\ A_p = False$. This means:

- The tricker $\mathcal{T}$ will exchange the location of the toy six

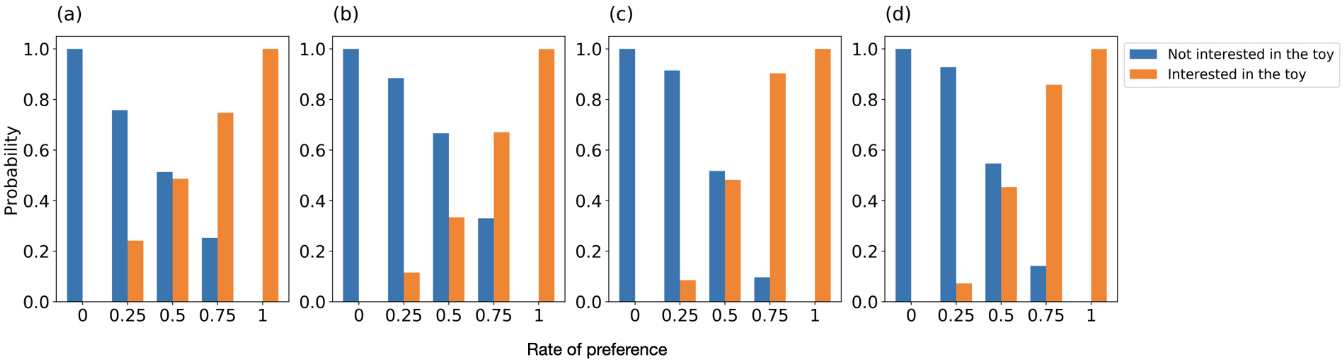

Fig. 4. Human agent's preference prediction, obtained by fixating $A_{fb} = False$, $R_{fb} = 0.5$ and varying R_p. (a) $A_p = True$, closed condition. (b) $A_p = True$, opened condition. (c) $A_p = False$, closed condition. (d) $A_p = False$, opened condition.

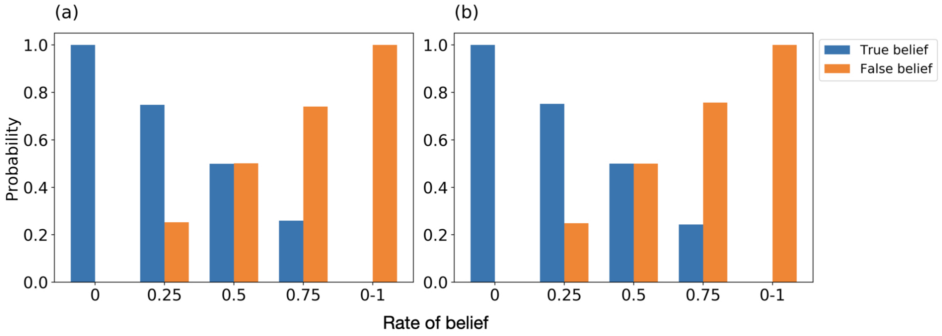

Fig. 5. Human agent's belief prediction, obtained by fixating $R_p = 0.5$ and varying R_{fb}. (a) $A_{fb} = True$. (b) $A_{fb} = False$.

times while $\mathcal{S}$ is outside of the room and two times with $\mathcal{S}$ present.

- We determine the belief condition successively with respect to R_{fb}.
- $\mathcal{S}$ will approach the box where they believe the toy is located four times and the other box for the remaining times.
- We randomly assign preferences to $\mathcal{S}$ with respect to R_p.

B. Human-robot interaction

An important aspect of this study is to implement a human-robot interaction experiment with the described model and analyze its performance. The goal is to prove whether the robot can observe and analyze people's behaviors to infer their mental states and how it should affect its interaction with people. To achieve this, we followed the setup of the false beliefs experiment in [8]. In this experiment, a robot observes two humans manipulating and putting a red ball into a box. The robot is placed in front of a table with two boxes to store the ball. As in the initial experiment, one person plays the role of the searcher who will be tricked by a second person putting the toy in a different box from the initial one.

Figure 3 shows pictures of the interactions, which indicate the scenarios and the robot's responses under different conditions. During the first step, the robot learns to recognize who the searcher and the tricker are (picture A). Then, according to the conditions, we have two scenarios where we evaluate the robot's response at the end:

1) False Belief condition: : The searcher puts the toy in the first box (picture B1) and leaves the room (picture C1). In their absence, the tricker approaches the robot (picture D1) and switches the toy's position from the original container (picture E1) to the other one. Finally, the searcher returns to the table and stands next to the first box, the one where they believe still contains the toy (F1).

2) True Belief condition: : In contrast with the first condition, after placing the toy in the aasecond box (B2), the searcher does not leave the room (C2). They are then aware of the tricker approaching the robot (D2), and swapping the toy's location (E2). When the malicious user leaves, the searcher steps back in front of the robot, standing next to the box where the toy was originally located (F2).

At the conclusion of both tasks, the robot must make a decision about what advice to provide based on its understanding of the searcher's mental state. Our expectation is that in the "False Belief" condition, the robot will point to the box containing the toy to help the searcher retrieve it, while in the "True Belief" condition, the robot will verbally acknowledge that the searcher is not interested in it.

V. Results

A. Simulation

To evaluate the performance of our model, we generated 500 paths with $S_{path} = 8$, varying the associated parameters.

We also conducted the evaluation process in two different ways: the "closed" and "open" conditions. In the "closed" condition, we only assessed the model's predictions when $\mathcal{S}$ was located next to the boxes (states 1,1 and 1,2 in Figure 2). In the "open" condition, we measured the model's inference at any possible state of $\mathcal{S}$.

The evaluation consisted of two parts: the evaluation of the human's preferences and the evaluation of the human's beliefs.

1) Evaluation of the human's preferences: To assess the model's ability to identify the simulated human's preferences, we set $R_{fb} = 0.5$ and varied R_p between the values of [0, 0.25, 0.5, 0.75, 1]. For instance, a preference rate of 0.5 indicates that the agent would move to the box where it believes the toy is located 50% of the time, while it would choose the other box 50% of the time.

Figure 4 displays the evaluation of the agent's preferences. The four charts correspond to different combinations of A_p and the evaluation type (closed or opened). For instance, Figure 4a represents the closed evaluation with $A_p = True$, and Figure 4b shows the opened evaluation with $A_p = False$. As the graphs demonstrate, the model can accurately infer the agent's preferences based on the provided preference ratios, even when the preference rate is set to 0.5, where the agent's behavior alternates more frequently. However, the model's performance is relatively lower when inferring the agent's preferences by considering all the states. Specifically, when the preference rate is set to 0.5, the model predicted that the agent was interested in the toy less than 40% of the time and close to 40% of the time when the preferences were not alternating. This might be attributed to the model's uncertainty about the agent's preferences when the agent is not close to either of the boxes.

2) Evaluation of the human's beliefs: For the belief inference evaluation, we set $R_p = 0.5$ and varied R_{fb} between the values of [0, 0.25, 0.5, 0.75, 1]. The results, depicted in Figure 5, show that the model can accurately infer the human agent's beliefs regarding R_{fb}. For instance, when R_{fb} is 0.5, corresponding to a belief condition change 50% of the time, the model can correctly infer that $\mathcal{S}$ has false beliefs approximately 50% of the time, and true beliefs for the remaining time.

The evaluations of both beliefs and preferences validate the performance of the BToM in this false-beliefs understanding experiment. The results shown in Figure 5 demonstrate that the model can accurately track the searcher's beliefs. Joint inference of beliefs and preferences is depicted in Figure 4, where we can observe how beliefs affect the agent's desires. As observed in the evaluation of beliefs, we notice a similar symmetry related to the assigned values of R_p. However, the inconsistent predictions in the opened condition indicate that the model cannot accurately determine the searcher's preference and assumes that the agent is not interested in the toy. This can be explained by the fact that the model's understanding of the agent's intention becomes clearer when it is approaching one of the boxes. In any case, the model is mainly used to explain the reason behind the agent's choice of a particular box.

B. Human-robot interaction

Following the promising results from the simulation presented in Section V-A, we aimed to confirm the model's performance through an HRI experiment wherein a robot tracks the mental states of a human and takes appropriate actions based on the context. Our main objective in this experiment was to evaluate the robot's behavior in response to the task and assess how well the robot follows the searcher's beliefs and preferences.

In terms of both conditions discussed in Section IV-B, we observe similarities and differences with the results depicted in Figures 6 and 7 that reflect how the robot's cognitive model infers those mental states over time. In both conditions, the model starts with no knowledge of the environment, which is reflected by the equiprobability of the human's beliefs and preferences during the introductory step. However, over time, the model begins inferring those mental states by observing human behavior.

1) Evaluation of the human's beliefs: In the False Belief (FB) condition, the robot accurately infers that the searcher has false beliefs by determining that they believe the toy is in box 1, even after the tricker swaps its location (Figure 3 E1). Conversely, in the True Belief (TB) condition, the robot infers that the searcher has a true belief regarding the toy's location (Figure 3 E2).

2) Evaluation of the human's preferences: The robot demonstrated its ability to infer the searcher's preferences in both conditions. In both conditions, the cognitive model predicts that the searcher is interested in the toy when they try to put it in box1 (Figure 3 B1) or box2 (Figure 3 B2). When they move away from it (Figure 3 C-E), the model assumes that they are no longer interested in the toy. Finally, when they move back close to it, the preference levels rise again in the FB condition (Figure 3 F1), whereas in the TB condition, the searcher's preference remains low. By understanding this and analyzing the searcher's position with respect to the two boxes, the robot can successfully determine which action to take: in the FB condition, it will direct the searcher to the correct location of the toy, while in the TB condition, it will acknowledge that they are no longer interested in it.

These outcomes are consistent with our initial hypothesis about the impact of beliefs and preferences on human behavior and how a cognitive robots can use this information to drive their action selection capabilities. Furthermore, they validate the use of BToM as a model for empowering social robots to understand and predict human actions, granting them a greater degree of social autonomy.

VI. CONCLUSIONS

In this paper, we presented the results of our HRI experiment on a ToM-capable cognitive architecture and discussed

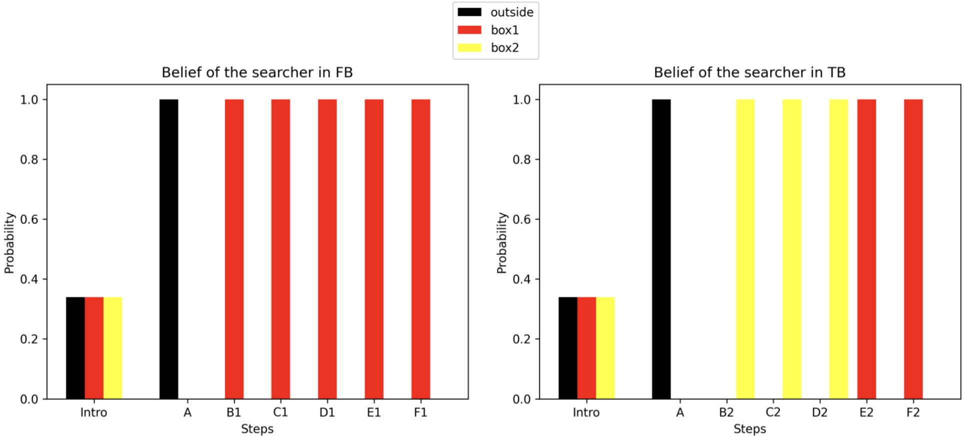

Fig. 6. The robot's reasoning on the searcher's beliefs during the two interactions under the "False Beliefs" (FB) and "True Beliefs" (TB) conditions. Each bar graph represents the searcher's belief about the toy's location.

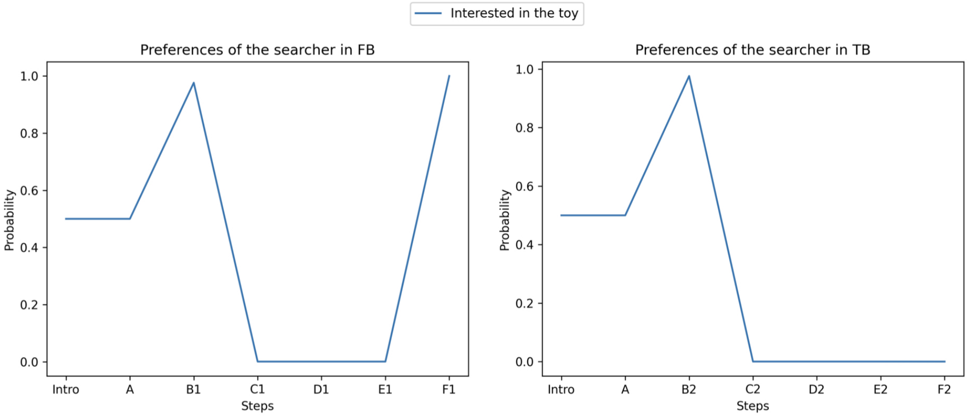

Fig. 7. The robot's reasoning on the searcher's preferences during the two interactions under the "False Beliefs" (FB) and "True Beliefs" (TB) conditions. The curve depicts the interest level of the searcher regarding the toy.

the importance of false belief understanding for collaborative agents. We traced the development of this mental skill in both the psychological and computational domains. Our methodology involved the use of BToM [9], [10], a DBN originally used to determine an agent's desires in a 2D world, and applied it to a more complex environment involving a psychology experiment used to test false belief understanding in children. Our simulated tests provided data in line with our expectations, and our real-world HRI experiment replicated the false belief understanding experiment to observe and analyze the behavior of a social robot embedded with a ToM cognitive model. The results demonstrate the effectiveness of using a BToM model for enabling a social robot to understand and reason about human beliefs and preferences in a false-belief understanding experiment. Our study suggests that by tracking the mental states of humans, robots can better assist and collaborate with them. While our scenario provides a clear demonstration of our model's performance in HRI, we recognize the need to implement more complicated and real-world-oriented situations to evaluate the adaptability of our system. In the future, we plan to conduct more complex scenarios with multiple actors and objects. Additionally, we aim to conduct user studies to test whether the predictions made by the BToM model align with how humans reason about false beliefs. Such studies would provide valuable insights into how robots can better understand human behavior and improve their ability to interact and collaborate with humans in real-world settings.

ACKNOWLEDGMENT

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.

Cangelosi and Vinanzi's work on this paper was supported by the Air Force Office of Scientific Research, USAF under Award No. FA9550-19-1-7002 and by the UKRI TAS Node on Trust.

REFERENCES

[1] M. Hellou, N. Gasteiger, J. Lim, M. Jang, and H. Ahn, "Personalization and localization in human-robot interaction: A review of technical methods," *Robotics*, vol. 10, p. 120, 11 2021.

[2] B. Irfan, A. Ramachandran, S. Spaulding, D. F. Glas, I. Leite, and K. L. Koay, "Personalization in long-term human-robot interaction," in *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 685–686, 2019.

[3] D. Premack and G. Woodruff, "Does the chimpanzee have a theory of mind?," *Behavioral and Brain Sciences*, vol. 1, no. 4, p. 515–526, 1978.

[4] J. H. Flavell, "Cognitive development: children's knowledge about the mind.," *Annual review of psychology*, vol. 50, pp. 21–45, 1999.

[5] A. Gopnik and H. M. Wellman, *The theory theory*, p. 257–293. Cambridge University Press, 1994.

[6] H. M. Wellman, D. Cross, and J. Watson, "Meta-analysis of theory-of-mind development: The truth about false belief," *Child Development*, vol. 72, no. 3, pp. 655–684, 2001.

[7] S. Baron-Cohen, A. M. Leslie, and U. Frith, "Does the autistic child have a "theory of mind" ?," *Cognition*, vol. 21, no. 1, pp. 37–46, 1985.

[8] D. Buttelmann, M. Carpenter, and M. Tomasello, "Eighteen-month-old infants show false belief understanding in an active helping paradigm," *Cognition*, vol. 112, pp. 337–342, Aug. 2009.

[9] C. L. Baker, R. R. Saxe, and J. 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*, pp. 2469–2474, 2011.

[10] 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*, vol. 1, 2017.

[11] A. N. Meltzoff, "Understanding the Intentions of Others: Re-Enactment of Intended Acts by 18-Month-Old Children," *Developmental Psychology*, vol. 31, no. 5, pp. 838–850, 1995.

[12] A. Gopnik and V. Slaughter, "Young children's understanding of changes in their mental states," 1991.

[13] B. M. Repacholi and A. Gopnik, "Early reasoning about desires: Evidence from 14-and 18-month-olds," 1997.

[14] H. M. Wellman, "The child's theory of mind.," *The child's theory of mind.*, pp. xiii, 358–xiii, 358, 1992. Place: Cambridge, MA, US Publisher: The MIT Press.

[15] D. Buttelmann, H. Over, M. Carpenter, and M. Tomasello, "Eighteen-month-olds understand false beliefs in an unexpected-contents task," *Journal of Experimental Child Psychology*, vol. 119, pp. 120–126, 3 2014.

[16] B. Priewasser, E. Rafetseder, C. Gargitter, and J. Perner, "Helping as an early indicator of a theory of mind: Mentalism or teleology?," *Cognitive Development*, vol. 46, pp. 69–78, 4 2018.

[17] D. Heckerman, D. Geiger, and D. M. Chickering, "Learning bayesian networks: The combination of knowledge and statistical data," in *Uncertainty Proceedings 1994* (R. L. de Mantaras and D. Poole, eds.), pp. 293–301, San Francisco (CA): Morgan Kaufmann, 1994.

[18] A. Gopnik, C. Glymour, D. M. Sobel, L. E. Schulz, T. Kushnir, and D. Danks, "A Theory of Causal Learning in Children: Causal Maps and Bayes Nets.," 2004.

[19] N. D. Goodman, C. L. Baker, E. B. Bonawitz, V. K. Mansinghka, A. Gopnik, H. Wellman, L. Schulz, and J. 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.

[20] S. Vinanzi, A. Cangelosi, and C. Goerick, "The collaborative mind:

intention reading and trust in human-robot interaction," *iScience*, vol. 24, no. 2, p. 102130, 2021.

[21] M. Patacchiola and A. Cangelosi, "A developmental cognitive architecture for trust and theory of mind in humanoid robots," *IEEE Transactions on Cybernetics*, vol. 52, no. 3, pp. 1947–1959, 2022.

[22] V. Konda and J. Tsitsiklis, "Actor-critic algorithms," *Society for Industrial and Applied Mathematics*, vol. 42, 04 2001.

[23] A. F. Morse, J. de Greeff, T. Belpeame, and A. Cangelosi, "Epigenetic robotics architecture (era)," *IEEE Transactions on Autonomous Mental Development*, vol. 2, no. 4, pp. 325–339, 2010.

[24] C. L. Baker, R. Saxe, and J. B. Tenenbaum, "Action understanding as inverse planning," *Cognition*, vol. 113, pp. 329–349, 12 2009.

[25] L. S. Zettlemoyer, B. Milch, and L. P. Kaelbling, "Multi-agent filtering with infinitely nested beliefs," in *Proceedings of the 21st International Conference on Neural Information Processing Systems*, NIPS'08, (Red Hook, NY, USA), p. 1905–1912, Curran Associates Inc., 2008.

[26] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognition*, vol. 29, no. 1, pp. 51–59, 1996.

[27] A. Bochkovskiy, C. Wang, and H. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," *CoRR*, vol. abs/2004.10934, 2020.