Where is My Favourite Toy? Inferring the Mental States of Users in False Belief Understanding

HELLOU, Mehdi, VINANZI, Samuele and CANGELOSI, Angelo

Published version

HELLOU, Mehdi, VINANZI, Samuele and CANGELOSI, Angelo (2024). Where is My Favourite Toy? Inferring the Mental States of Users in False Belief Understanding. In: 2024 IEEE International Conference on Development and Learning (ICDL). IEEE. [Book Section]

Where is my favourite toy? Inferring the mental states of users in false belief understanding

1[st] Mehdi Hellou
Manchester Centre for Robotics and AI
University of Manchester
Manchester, United Kingdom,
mehdi.hellou@manchester.ac.uk

2[nd] Samuele Vinanzi
Department of Computing
Sheffield Hallam University
Sheffield, United Kingdom
0000-0003-0241-9983

3[rd] Angelo Cangelosi
Manchester Centre for Robotics and AI
University of Manchester
Manchester, United Kingdom
0000-0002-4709-2243

***Abstract*—The increasing prevalence of social robots in today's world has made it crucial to have autonomous systems that can interact with humans and adjust to their behaviour. Doing so requires a deep understanding of the human mind, including complex mental states such as beliefs and preferences. To tackle this issue, we have developed a model that can identify false beliefs using the principles of Theory of Mind (ToM), a unique human cognitive mechanism that attributes mental states to others. False belief understanding has always been the primary benchmark used to evaluate ToM in psychology, and this still remain true when testing it in machine systems such as robots. Our model is a modified version of the Bayesian Theory of Mind (BToM), a probabilistic model that reasons on agents' mental states regarding their interactions within the environment. To test the model's performance, we set up a complex assistive scenario with a robot and two human agents playing with toys. In this scenario, the model serves as the cognitive component of the robot, responsible for organising a room with toys while considering the preferences and beliefs of the agents regarding the toys' locations. We have provided results to demonstrate the model's performance in different conditions. Additionally, we have used the Unity Engine as a platform to simulate the cleaning scenario and show the robot's role in such situations.**

***Index Terms*—Theory of Mind, False Belief understanding, Social Robotics, Bayesian Network, Reinforcement Learning**

I. Introduction

Nowadays, we are facing a rising amount of autonomous robots in our everyday lives, whether in personal environments, like household cleaning using devices such as Roomba [1], or in public spaces, including museums [2], [3]. Those robots are supposed to engage with people to facilitate communication with each other autonomously. They require exceptional behaviours to engage multiple people according to their preferences, personalities and needs. This adaptability is commonly referred to as "personalisation", the ability to tailor the robot's behaviour to various users according to their traits. Numerous studies have demonstrated that personalisation is a key factor in fostering trust, engagement and long-term human-robot interaction (HRI) [2], [4], [5].

Identify applicable funding agency here. If none, delete this.

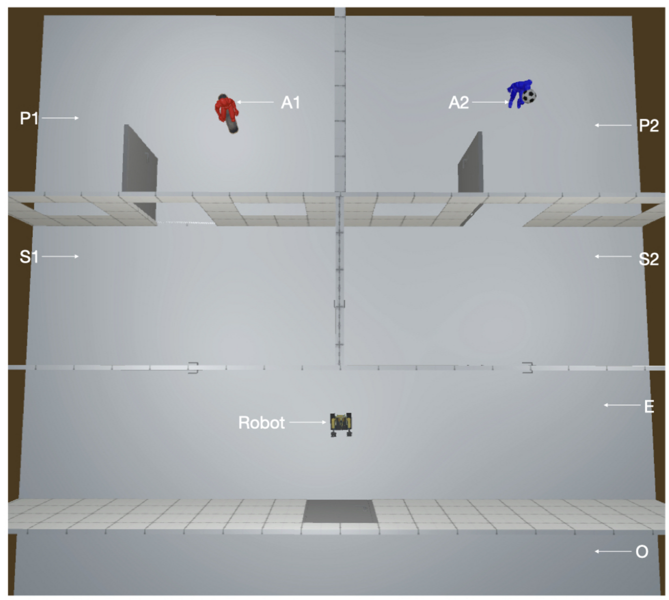


Fig. 1. Simulated environment created with Unity Engine, including the agents, the robot and the items. The agent $\mathcal{A}_1$ is skating in P_1 (playground 1), while $\mathcal{A}_2$ is playing with the ball in P_2 (playground 2). The robot is at the entrance E, waiting to help the human if needed. More information regarding the environment is provided in Section IV.

These adaptations are challenging and necessitate that the robot be fully aware of its users and surroundings. To achieve this goal, we have shifted our focus to psychology, exploring the cognitive ability known as Theory of Mind (ToM). ToM involves the human capacity to infer the mental states of others, including their beliefs, desires, and intentions [6]. By delving deeply into this cognitive process, ToM endows humans with the ability to predict the actions of others based on their visual perceptions.

As an extension of our recent work [7], we have developed a cognitive model to bestow social robots with ToM in an assistive task. The environment, which was created by using the *Unity Engine*, involves a robot organising rooms in a complex setting with two toys and two human agents who can interact with them (overview of the environment in Figure 1). Within the simulation, the robot adopts various strategies to assist the agents in playing with their preferred items according to their beliefs. To achieve this, we aim to implement the model proposed by Baker *et al.* the Bayesian Theory of Mind (BToM) [8]. BToM is a probabilistic model that leverages Bayesian inferences and partially observable Markov decision processes (POMDPs) to predict people's beliefs and desires. By using POMDPs to represent the agent's planning and inference about the world, BToM reasons about the agent's intention to accomplish a specific task based on its beliefs.

In this project, we are interested in predicting agents' mental states in "false beliefs understanding" (*FBu*) through their interactions within the environment. Researchers have predominantly explored *FBu* understanding to investigate ToM capabilities in human psychology, particularly with infants. This exploration aims to analyse infants' ability to discern moments when individuals hold beliefs that contradict reality [9]–[12]. A foundational experiment, known as "Sally-Anne" test (*SA*) as been widely used not only in the field of psychology research but also in robotics, with various adaptations [7]. The experiment introduces two characters, Sally and Anne, wherein Sally has a basket, and Anne has a box. Sally puts a marble in her basket and leaves the room. While away, Anne moves the marble from Sally's basket to her box. Subsequently, participants are asked, "Where will Sally look for her marble when she returns?". A significant part of people who have developed ToM will answer "Sally's basket".

In this context, we are keen on implementing this process within our scenario, involving an agent adopting false beliefs during interactions. The cognitive model is subsequently employed to comprehend the agents' actions by predicting their beliefs and preferences. We demonstrate the model's performance in various conditions, where agents exhibit specific behaviours. As a results, we endorse the model's efficacy in inferring and tracking the agents' mental states.

II. Related works

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 indicated that children undergo cognitive development, enabling them to understand others' mental states, including beliefs, desires, emotions, and intentions [9]–[14]. As mentioned earlier, the experiment widely used as measurement for *FBu* understanding is the *SA* test. Some experiments involve the active participation of children interacting directly with the environment [15]–[17]. The interaction of children serves as the basis for evaluating *FBu*. Other experiments focus on participants' ability to understand *FBu* and manipulate it in specific game scenarios [18].

The application of ToM principles has gained significant interest in advancing autonomous and intelligent systems in computer science and robotics. Traditionally, ToM research has been associated with psychology. However, leveraging Bayesian Networks (BN) [19], widely used graphical models in data analysis, has enabled the development of ToM-capable agents. This technique has empowered psychologists and researchers to better understand how beliefs influence decision-making in children when faced with *FBu* situations [20], [21].

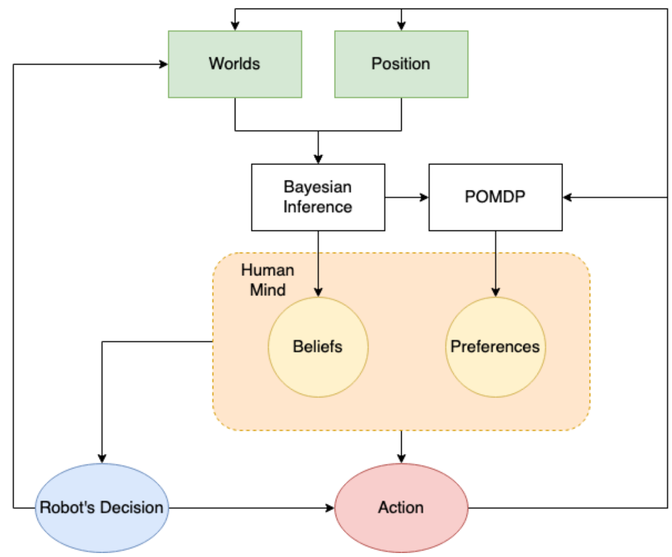


Fig. 2. The cognitive model incorporates the principles of BToM. This modified version involves the robot in the loop, taking the human's mental state as input and making decisions accordingly, such as bringing items and explaining their locations.

After this breakthrough, other researchers delved into the world of probabilistic graphical models (PGMs) to create ToM for social robots. For instance, Vinanzi *et al.* [22] developed a BN that can predict a robot's action based on the beliefs and actions of a human informant. Another remarkable example is Baker *et al.*'s BToM [8], [23], which employs Bayesian and POMDPs to represent how individuals infer others' goals or preferences. The authors compared their model to the cognitive process through an interactive experiment involving an agent's movement in a grid-world environment. Given the promising outcomes, numerous studies have incorporated this model into their work. For instance, [7] utilises a modified version of BToM to ascertain the cognitive states of users involved in HRI, while [24] learns individual preferences of drivers.

There are other models found in literature that utilised neural networks as interpreters for human mental states. In fact, Oguntola *et al.* [25] have created an interpretable modular neural framework to model the intentions of observed entities. This framework was put to the test in a rescue task within a Minecraft's 2D grid world.

III. Proposed approach

Inspired by the BToM model proposed by Baker et al. [8], [23], our cognitive model predicts the mental states of a dynamic agent for the decision-making of an autonomous robot providing support. Figure 2 presents an overview of the model, with the addition of the robot in the loop, influencing the human during action execution. The robot is capable of moving items, delivering them to specific places, and providing explanations to users, such as the location of preferred items and their availability. It can also request users to relinquish items they are playing with so that others can use them. These actions are rooted in the robot's understanding of human minds as observed through their interactions with the world. The robot's model predictions follow three steps: (1) given the world's disposition and the agent's position, the robot's BToM model determines beliefs through Bayesian Inference; (2) based on the agent's beliefs and actions, the robotic cognitive model infers the agent's preferences using rational plans designed as POMDPs; (3) the prediction of the agent's intentions is defined as a combination of beliefs and preferences. Further details are provided in the subsequent paragraphs.

The Bayesian inference enables the robot-embedded cognitive model to update the agent's beliefs regarding location and the world's disposition. In contrast to the method in [8], [23], we create a dynamic environment where the items can move at different locations, making it more challenging for agents to accurately perceive their surroundings. The update of the agent's beliefs is based on prior knowledge, the current state of the world and the agent's observation. It represents the agent's beliefs as a probability distribution over the set of possible worlds $W = \{w^1, \ldots, w^n\}$. This take into account the observation space O, including all the possible observations the agent can make $\{o^1, \ldots, o^n\}$ where o^n represent the observation of the world w^n. The action space A includes motion actions $\{$*Up, Down, Right, Left and Stay*$\}$ and interaction with items $\{$*Take and Drop*$\}$.

To update the agent's belief about the next world w^i is true at time t, we consider the prior knowledge, the current state of the world, and the agent's observation. We denote the belief that the next world w^i is true at time t knowing the prior world w^j at $t-1$ as $b_t(w_t^i)$, where $i, j \in \{1, \ldots, n\}$. Hence, we update the belief of w_t^i at t, regarding the prior belief $b_{t-1}(w_{t-1}^j)$ of the world w^j, the likelihood $P\left(o_t^i \mid x_t, w_t^i\right)$ of observing the agent state x_t and the world w^i at t, and the probability $P\left(x_t, w_t^i \mid x_{t-1}, w_{t-1}^j, a_{t-1}\right)$ of observing the agent move from position x_{t-1} to x_t and the world changing from w_{t-1}^j to w_t^i given the action a_{t-1}, with $a \in A$. We defined a transition function $f(w_t^i, w_{t-1}^j)$ to represent unexpected change in the world (i.e., when $i \neq j$) and emphasise the belief when the agent has *FBu*. Using the Bayes rules, we can update the Bayesian belief $b_t(w_t^i)$ as followed:

$$b_t(w_t^i) \propto P\left(o_t^i \mid x_t, w_t^i\right) \cdot P\left(x_t, w_t^i \mid x_{t-1}, w_{t-1}^j, a_{t-1}\right)$$
$$b_{t-1}(w_{t-1}^j) \cdot f(w_t^i, w_{t-1}^j)$$

with:

$$f(w_t^i, w_{t-1}^j) = \begin{cases} \frac{r}{len(PW)} & \text{, if } i \neq j \\ \frac{1}{len(PW)} & \text{, else} \end{cases}$$

where r is a numerical factor determined to highlight the importance of *FBu*, and $len(PW)$ represents the size of all possible worlds the agent can observe when moving from x_{t-1} to x_t. After conducting various tests, we decided to set the factor to $r = 8$, as it appeared to be the most appropriate value to represent the change in the agent's belief.

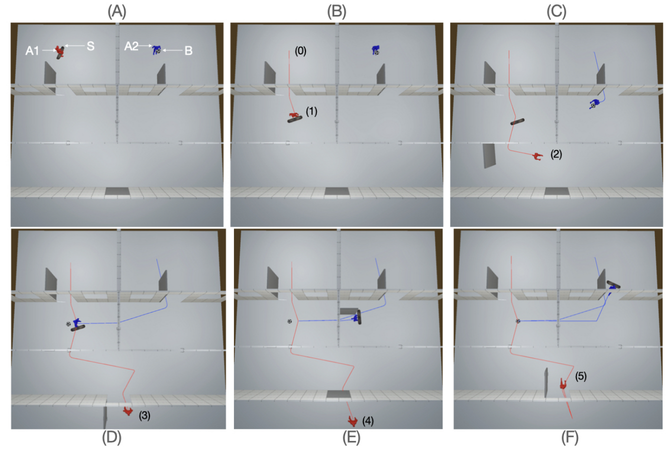


Fig. 3. Start of the scenario where $\mathcal{A}_1$ is using $\mathcal{S}$ and $\mathcal{A}_2$ is using $\mathcal{B}$. The paths of $\mathcal{A}_1$ and $\mathcal{A}_2$ are represented by the red and blue lines, respectively. The steps taken by $\mathcal{A}_1$ serve as reference for Section V.

As mentioned earlier, POMDPs represent the possible plans of an agent regarding their preferences and beliefs. This part of the robotic cognitive model uses a *Boltzmann Value iteration* algorithm [26], [27] to learn the possible policies that the agents follow regarding their preferences. The agent's POMDP state space includes the agents' position and beliefs, and the reward function $R(x, w^i, a)$ generates the utility to be in a state x according to a certain world w^i and an action a. Each action has a cost of 1, except the "playing" action, which is valued as the reward of the agents to play with the item according to their preferences. When the agents are in goal states, the reward appears as the distance between them and their preferences. In the end, the convergence of the algorithm generates the Q-Values for each state-action pair describing the likelihood of the agents performing action regarding their beliefs b, positions x, and the world's setting w^i: $P\left(a \mid b, x, w^i\right)$.

In the last step of our process, we calculate the joint beliefs and preferences to determine the agent's intention. Essentially, the agent's desire to achieve a specific goal through internal motivation, such as their beliefs and desires, can be defined as their intention [28]. The agent's intention in this context reflects their reasons for interacting with their preferred items based on their beliefs. The goal of this computation is to deduce the agent's beliefs and desires regarding their observed actions. We represent the joint probability at a given time step $t \leq T$, by considering the sequence of actions $\{a_0, ..., a_{T-1}\}$ executed by the agents regarding their positions along a path $x_0, ..., x_{T-1}$ starting from time 0 and going up to $T-1$:

$$P\left(b_t, d_t \mid x_{0:T-1}, a_{0:T-1}\right) \propto P\left(b_t, d_t \mid x_{0:t-1}, a_{0:t-1}\right)$$
$$P\left(x_{t:T-1}, a_{t:T-1} \mid b_t, d_t\right)$$

This give the joint probabilities of both mental states and the computation is similar to the forward-backward algorithm (FB) in Hidden Markov Model [29], [30].

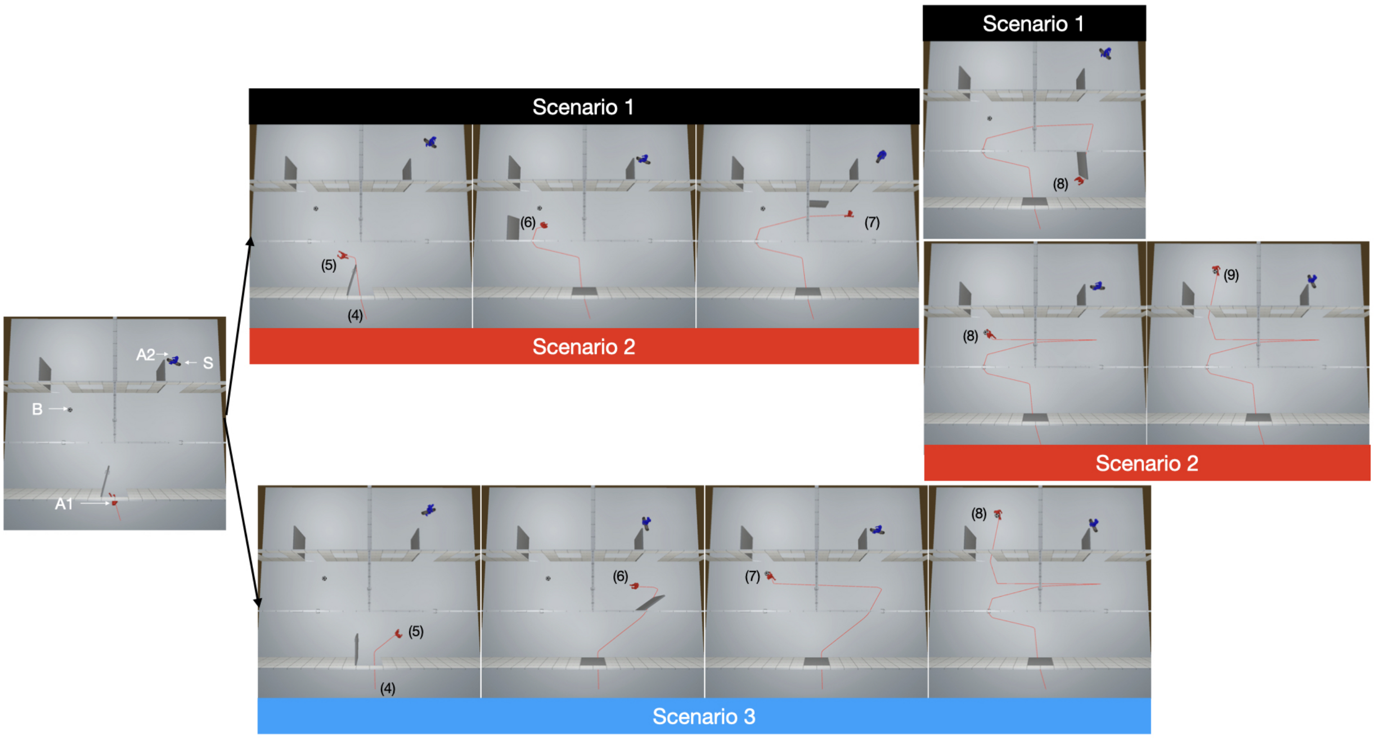
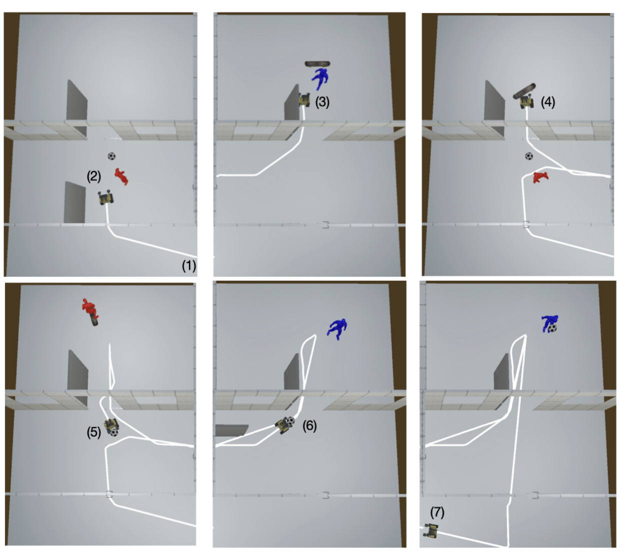
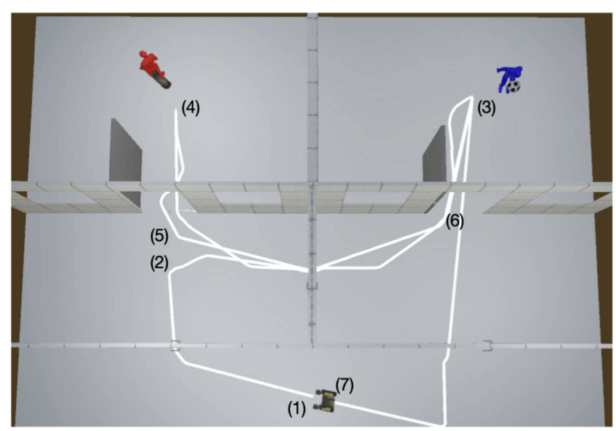


Fig. 4. Representation of the different scenario where $\mathcal{A}_2$ is playing with $\mathcal{S}$ and $\mathcal{A}_1$ is following three different paths according to its behaviour. The steps taken by $\mathcal{A}_1$ serve as reference for Section IV and V.



(a) The subsequent steps taken by the robot to deliver the skate to $\mathcal{A}_1$ and the ball to $\mathcal{A}_2$.



(b) The overall path followed by the robot to assist the agents.

Fig. 5. Here is an example of the robot's role in organising the environment while taking into consideration the mental states of the agents.

IV. Experiment

A. Incorporation of multiple scenarios

As explained in Section I, we use the scenario of a cleaning robot that organises a playground environment to assist two agents playing with toys. The environment consists of two playground areas, P_1 and P_2, where two agents, $\mathcal{A}_1$ and $\mathcal{A}_2$, can play with different toys, such as a ball $\mathcal{B}$ and a skate $\mathcal{S}$. Each playground has an assigned storage room, S_1 and S_2, respectively, that are interconnected and linked to the entrance room E, which connects to the external environment O. When the agents are located in S_1 or P_1, they have full observations about those locations and cannot observe what is happening in S_2, P_2, E, and O. The same applies when they are in the second part (S_2 and P_2) of the environment. When the agents are in the entrance room E or the outside environment O, they can only observe what is happening inside of those rooms. In other words, they can not observe any other rooms. By observing the agents' behaviours interacting with the toys and navigating within the environment, the robot's model infers their mental states regarding their preferences and beliefs about the environment (*i.e.*, their locations). These mental states serve as decisions for the robot when interacting with humans, *e.g.* acknowledging agents when their preferred toy is not available or avoiding them to have *FBu* by returning the toys to their original location.

To assess the robotic cognitive model, we analyse it in different scenarios where one of the agents, here $\mathcal{A}_2$, misleads the other agent $\mathcal{A}_1$ into developing false beliefs. Figure 3 and 4 depicts the scenario we describe in the following lines. In Figure 3, $\mathcal{A}_1$ and $\mathcal{A}_2$ are respectively playing with the skate and the ball in P_1 and P_2 (A). $\mathcal{A}_1$ decide to store the skate in S_1 (B) and take a walk outside (D and E). Meanwhile, $\mathcal{A}_2$ wants to play with $\mathcal{S}$ and decides to retrieve it on the opposite side of the room (C and D). However, $\mathcal{A}_2$ exchanges the location of S with B and takes $\mathcal{S}$ in the other playground P_2 to dupe the other agent (E and F). As a result, we have at this moment A_2 playing skate in P_2 and $\mathcal{B}$ in the storage room S_1. Then, the first agent returns to the entrance (F) and follows different paths relating to the conditions. Figure 4 depicts the different paths followed by $\mathcal{A}_1$ according to the conditions explained in detail below:

- ***Scenario 1***: $\mathcal{A}_1$, located outside of room (step 4), enters S_1 where the agent initially placed $\mathcal{S}$ in (step 5). Upon discovering that the ball is present instead, $\mathcal{A}_1$ decides to move to the other storage room, S_2 (step 6). Upon realising that $\mathcal{A}_2$ is playing with the skate (step 7), $\mathcal{A}_1$ decides to leave and returns to the entrance, E (step 8).
- ***Scenario 2***: In the second scenario, $\mathcal{A}_1$ follows the same pattern as the previous one (steps 4 to 7). However, instead of leaving for the entrance, the agent decides to return to S_1 (E2), takes $\mathcal{B}$ (step 8), and plays with it in area P_1 (step 9).
- ***Scenario 3***: In the third scenario, $\mathcal{A}_1$ is not interested in playing with $\mathcal{S}$ and prefers to play with $\mathcal{B}$. Consequently, the agent moves to its belief's location in S_2 (step 6), notices that the ball is not present (step 6), and goes to S_1 (step 7) to play with the ball in P_1 (step 8).

Using those various scenarios, we test the robot's model inference regarding the beliefs and the preferences in Section V.

B. Tasks performed by the robot

This work focuses on integrating an autonomous system with ToM as an assistive machine. The robot's task involves cleaning and organising the environment, all while interacting with agents and adjusting its behaviours based on their beliefs and preferences. More specifically, the robot uses the cognitive model in Figure 2 to represent each agent's mental states and take decisions. It considers the activity performs by $\mathcal{A}_1$ and $\mathcal{A}_2$ and performs the action such as *bring(item,location)* (bring an item to a location). For example, consider the scenario depicted in Figure 5, where $\mathcal{A}_1$ strongly desires to play with $\mathcal{S}$.

When entering storage room S_1, the robot can predict the desire of $\mathcal{A}_1$ and explain to the agent that $\mathcal{S}$ is currently occupied (steps 1 and 2). The robot can also act as an intermediary between the agents by requesting $\mathcal{A}_2$ to leave the skate for $\mathcal{A}_1$ and play with the ball instead (step 3). Subsequently, the robot delivers the skate to $\mathcal{A}_1$ so that the agent can play with it (step 4) and the ball to $\mathcal{A}_2$ (steps 5 and 6) before returning to its initial position (step 7). Importantly, the entire process is conducted with the approval of $\mathcal{A}_2$, as the robot is designed to interact in a positive manner with users. This interaction demonstrates the potential of using ToM to guide the robot's decisions in facilitating seamless user interactions.

V. Results

For the purpose of validating the performance, we analyse the inferences made by the robot-embedded cognitive model regarding the preferences and beliefs of $\mathcal{A}_1$. Results are depicted in Figures 6 and 7 representing how the model track the agent $\mathcal{A}_1$'s mental states based on the path it followed. These results are then provided to guide the simulated robot bringing the appropriate items to the agent according to its beliefs and preference.

A. Beliefs

For the beliefs, the robot's model tracks and infers the locations where $\mathcal{A}_1$ thinks $\mathcal{B}$ and $\mathcal{S}$ are located (P_1, P_2, S_1, S_2). We compared the ground truth and predictions regarding the agent's beliefs for each item in Figure 6. To determine the ground truth, we manually represented the expected beliefs according to the scenario, assuming that the agent starts with equiprobability for both items. The model's predictions are generally in accordance with the true beliefs with some difference when the agent's belief changes, e.g. the ball's belief between steps 8 and 9 in scenario *Scenario 2*. This is common occurence when using the forward-backwards algorithm, which requires time for the information to pass through and adjust to the actual value. We will explain in

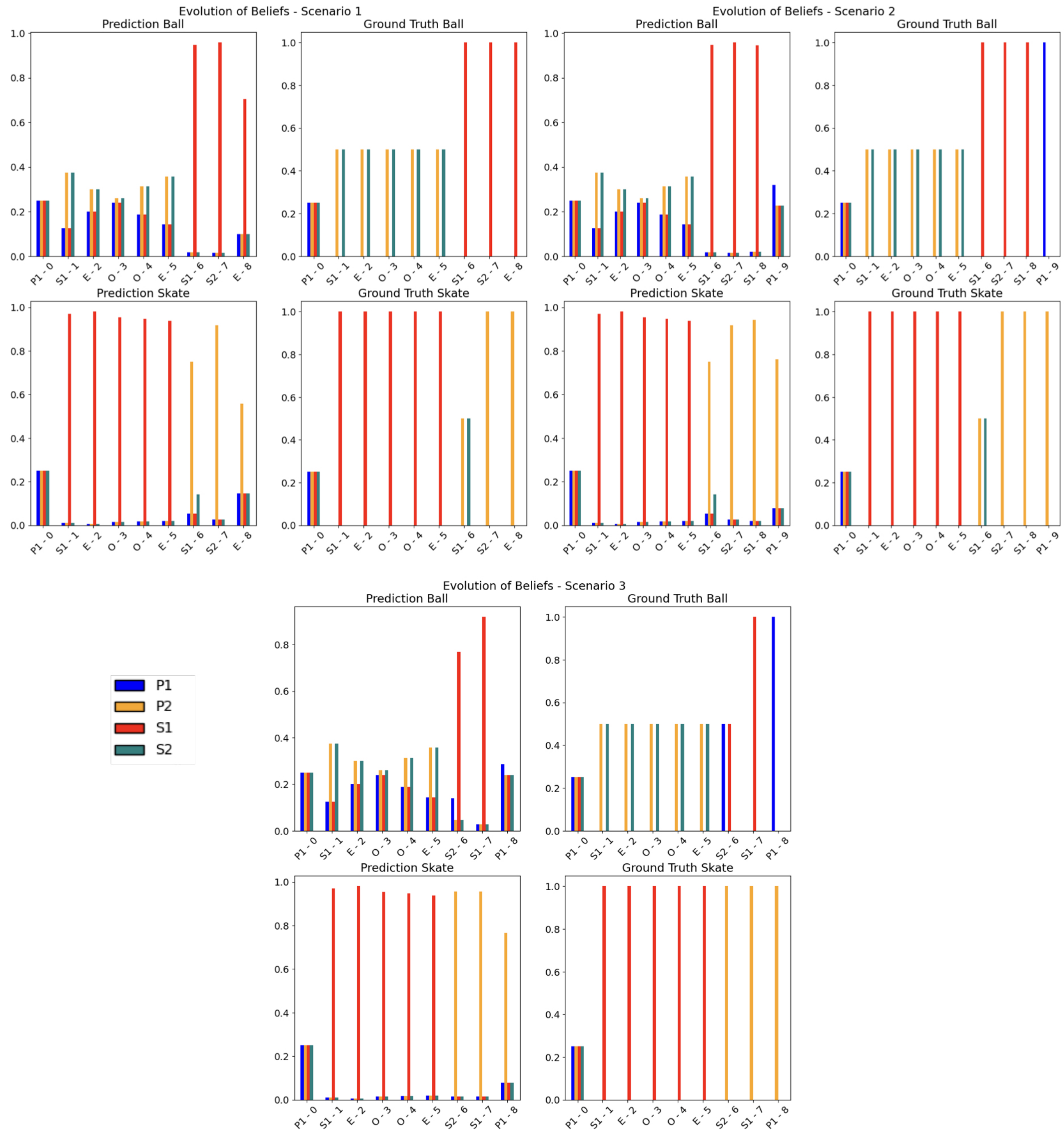


Fig. 6. The robot's model prediction and ground truth about $\mathcal{A}_1$ beliefs regarding the ball's and skate's location. The x-axis depict the position of $\mathcal{A}_1$ regarding the agent's position and the steps depicted in Figure 3 and 4. The y-axis represent the agent's beliefs as a probability distribution.

detail how the information about $\mathcal{A}_1$'s belief aligns with each scenario.

In the initial steps, which are common to all scenarios (step 0 to 5), the robotic cognitive model correctly implies that $\mathcal{S}$ is located in storage S_1 since $\mathcal{A}_1$ is directly interacting with the skate and knows its location. On the other hand, the agent has no precise knowledge regarding the ball's location, but it knows that this one is either present in S_2 or P_2. That follows our initial hypothesis since the agent could observe the objects present in S_1 and P_1 during the first steps (steps 0 and 2) but not in S_2 or P_2.

In scenarios *Scenario 1* and *Scenario 2*, when the agent goes back to visit the first storage, it finds the ball instead of the skate and updates its beliefs (step 6). When $\mathcal{A}_1$ decides to go to S_2 (step 7), the agent's beliefs match the world's actual state ($\mathcal{B}$ in S_1 and $\mathcal{S}$ in P_2). The agent then updates its beliefs in step 9, about its actions when taking the ball to play with it in P_1 (*Scenario 2*), or not modifying the belief when going back to the entrance (step 8 in *Scenario 1*).

In *Scenario 3*, we observe that the robot's model has already predicted a strong belief for the agent to retrieve $\mathcal{B}$ in S_1, despite the possibility of the object also being in P_1, when compared to the ground truth (step 6). This demonstrates a particularity of the forward-backward algorithm [31], which

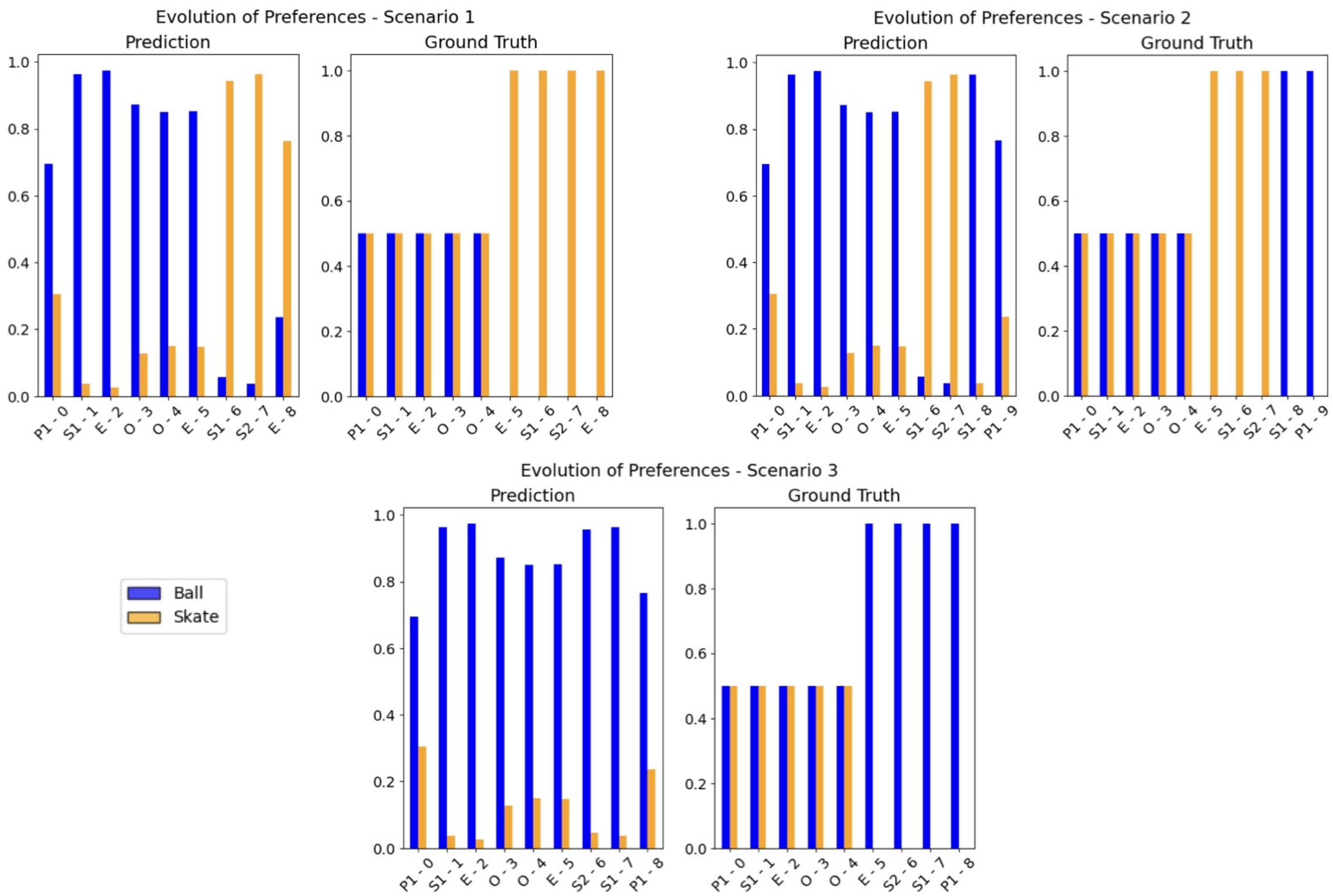


Fig. 7. The robotic cognitive model prediction about $\mathcal{A}_1$ preferences regarding the ball's (left bars) and skate's (right bars). The x-axis depict the position of $\mathcal{A}_1$ regarding the agent's position and the steps depicted in Figure 3 and 4. The y-axis represent the agent's preference as probabilities.

uses both past and future information to make predictions. Since the agent is moving to the next state in S_1 and retrieving $\mathcal{B}$ in S_1 (step 7), the robotic cognitive model anticipates this behaviour in the prior step.

B. Preferences

Similarly to the beliefs, we manually establish the ground truth for the preferences regarding the scenario. In all conditions, we assume that the agent has no preferences regarding both items since it goes outside (step 0 to 4). However, the robot's BToM model infers that the agent prefers the ball because it places the skate into the storage S_1. The reason is that we consider an important reward when the agent plays with one of the items, which decreases considerably when the agent disposes of them. Although we did not capture this specific behaviour, it was intentional, as we wanted to focus on tracking preferences during false belief situations.

This is particularly portrayed in step 5 when the agent returns to play with one of the toys. In *Scenario 1*, the agent prefers to play with $\mathcal{S}$, which is accurately captured by the model when $\mathcal{A}_1$ enters S_1 (step 6). Similarly, in *Scenario 2*, the cognitive model identifies the preference for the skate when entering S_1. Conversely, in *Scenario 3*, the preferences switch to the ball when the agent goes in S_2 (step 6). This implies that $\mathcal{A}_1$ will enter the storage rooms where it believes its preferred item is located. There is a symmetry between *Scenario 1* and *Scenario 3*, illustrating the agent's inclination to be drawn to its first choice. In *Scenario 3*, the agent can play with its favourite item ($\mathcal{B}$), whereas in *Scenario 1*, it faces a limitation due to the non-availability of the skate, currently in use by $\mathcal{A}_2$ (steps 7 and 8). Therefore, we did not change the preference in *Scenario 1* with the ground truth since the agent preferred to go to the entrance instead of playing with the ball. The model understands this difference between the two scenarios, which is consistent with the values depicted by the ground truth.

Alternatively, in *Scenario 2*, the robot's model captures the evolution of the agent's interest in items when $\mathcal{A}_1$ decides to play with the ball instead of the skate (steps 8 and 9). As a result, the model effectively captures the evolution of the preferences for $\mathcal{A}_1$ regarding its actions. Moreover, the model dynamically captured the preferences, a feature that has been previously noted in [8] but has yet to be fully demonstrated. However, further research is needed to verify the effectiveness of the robot's model in a wider range of situations.

VI. Conclusion

Our research paper proposes a cognitive model that integrates ToM into social robots, enabling them to predict mental states in a cleaning scenario where two human agents are playing with toys based on their preferences and beliefs. Building on our previous work [7], this updated robotic cognitive model incorporates the baseline of BToM to enhance its adaptability in complex false belief scenarios within dynamic environments. Our results have demonstrated its ability to infer and track an agent's beliefs and preferences through their interactions. Furthermore, we have implemented a simulated assistive scenario wherein a social robot supports agents with false beliefs. Our findings reinforce the robot's model performance in various conditions and highlight the potential of ToM in autonomous systems to aid in complex situations. They also emphasise the importance of allowing machine to adapt their behaviours to users when predicting their cognitive states and actively participate by interacting them, particularly when individual has beliefs that contradict the reality (*FBu*). This processing approach aligns with the procedural steps employed in certain psychological experiments to measure ToM ability for people [15]–[17].

Moving forward, we plan to test this robot-embedded cognitive model in a real-world HRI experiment to gather direct inputs from participants and enrich the decision-making process. This experiment will not only serve as additional validation for our model but also offer insights into people's perceptions when interacting with a robot equipped with such a cognitive process.

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.

We would like to acknowledge also the use of Grammarly in the preparation of this manuscript, for proofreading and typo correction.

References

[1] J.-Y. Sung, L. Guo, R. E. Grinter, and H. I. Christensen, ""my roomba is rambo": intimate home appliances," in *UbiComp 2007: Ubiquitous Computing: 9th International Conference, UbiComp 2007, Innsbruck, Austria, September 16-19, 2007. Proceedings 9*, pp. 145–162, Springer, 2007.

[2] N. Gasteiger, M. Hellou, and H. S. Ahn, "Deploying social robots in museum settings: A quasi-systematic review exploring purpose and acceptability," *International Journal of Advanced Robotic Systems*, vol. 18, no. 6, p. 17298814211066740, 2021.

[3] M. Hellou, J. Lim, N. Gasteiger, M. Jang, and H. S. Ahn, "Technical methods for social robots in museum settings: An overview of the literature," *International Journal of Social Robotics*, vol. 14, no. 8, pp. 1767–1786, 2022.

[4] 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, IEEE, 2019.

[5] N. Gasteiger, M. Hellou, and H. S. Ahn, "Factors for personalization and localization to optimize human–robot interaction: A literature review," *International Journal of Social Robotics*, vol. 15, no. 4, pp. 689–701, 2023.

[6] 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.

[7] M. Hellou, S. Vinanzi, and A. Cangelosi, "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)*, pp. 1893–1900, 2023.

[8] 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.

[9] 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.

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

[11] A. M. Leslie, "Pretending and believing: Issues in the theory of tomm," *Cognition*, vol. 50, no. 1-3, pp. 211–238, 1994.

[12] 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.

[13] K. Bartsch and H. Wellman, "Young children's attribution of action to beliefs and desires," *Child development*, pp. 946–964, 1989.

[14] A. N. Meltzoff, "Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children.," *Developmental psychology*, vol. 31, no. 5, p. 838, 1995.

[15] D. Buttelmann, M. Carpenter, and M. Tomasello, "Eighteen-month-old infants show false belief understanding in an active helping paradigm," *Cognition*, vol. 112, no. 2, pp. 337–342, 2009.

[16] 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, 2014.

[17] 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, 2018.

[18] K. E. Vanderbilt, D. Liu, and G. D. Heyman, "The development of distrust," *Child development*, vol. 82, no. 5, pp. 1372–1380, 2011.

[19] D. Heckerman, "Probabilistic similarity networks," *CoRR*, vol. abs/1911.06263, 2019.

[20] 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.

[21] 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.

[22] S. Vinanzi, M. Patacchiola, A. Chella, and A. Cangelosi, "Would a robot trust you? developmental robotics model of trust and theory of mind," *Philosophical Transactions of the Royal Society B*, vol. 374, no. 1771, p. 20180032, 2019.

[23] 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, no. 4, p. 0064, 2017.

[24] A. Darwish and H. J. Steinhauer, "Learning individual driver's mental models using pomdps and btom," in *Proceedings of the 6th International Digital Human Modeling Symposium*, pp. 51–60, 2020.

[25] I. Oguntola, D. Hughes, and K. Sycara, "Deep interpretable models of theory of mind," in *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*, pp. 657–664, IEEE, 2021.

[26] W. Zhang and N. L. Zhang, "Value iteration working with belief subset," in *Eighteenth National Conference on Artificial Intelligence*, (USA), p. 307–312, American Association for Artificial Intelligence, 2002.

[27] J. Pineau, G. Gordon, S. Thrun, *et al.*, "Point-based value iteration: An anytime algorithm for pomdps," in *Ijcai*, vol. 3, pp. 1025–1032, 2003.

[28] D. C. Dennett, *The intentional stance*. 1989.

[29] L. E. Baum *et al.*, "An inequality and associated maximization technique in statistical estimation for probabilistic functions of markov processes," *Inequalities*, vol. 3, no. 1, pp. 1–8, 1972.

[30] L. Zettlemoyer, B. Milch, and L. Kaelbling, "Multi-agent filtering with infinitely nested beliefs," *Advances in neural information processing systems*, vol. 21, 2008.

[31] L. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.