Theory of Mind Based Assistive Communication in Complex Human Robot Cooperation

Moritz Bühler, Jürgen Adamy, Thomas Weisswange

2022

Preprint:

This is an accepted article published in IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2022) - Workshop on Human Theory of Machines and Machine Theory of Mind for Human-Agent Teams (TOM4HAT).
The final authenticated version is available online at:
https://doi.org/10.48550/arXiv.2109.01355

Theory of Mind-based Assistive Communication in Complex Human-Robot Cooperation

Moritz C. Buehler[1,2,3], Jürgen Adamy[1] and Thomas H. Weisswange[2]

***Abstract*— When cooperating with a human, a robot should not only care about its environment and task but also develop an understanding of the partner's reasoning. To support its human partner in complex tasks, the robot should share information. However simply communicating everything it knows might annoy and distract humans as not all information might be relevant or novel in a given situation. To decide when and what type of information should be communicated, one needs to take into account knowledge and situation awareness of the partner. We previously proposed the concept of Theory of Mind-based Communication which selects information sharing actions based on evaluation of relevance and an estimation of human beliefs. Here, we integrate this into a communication assistant in a cooperative human-robot setting and evaluate performance benefits. Human belief is estimated based on task progression and gaze positions. Using this belief estimate, the system predicts the most likely next actions and evaluates the impact on future rewards. If a simulated belief update from a robot communication action will lead to a future plan with higher expected reward, while including an explicit cost of communication, the system will choose to assist the human with information. We designed a task that is challenging for the human and generates situations where humans could potentially profit from additional information communicated by the robot. Naive human participants performed the task with the robot to generate data which was then used to evaluate the influence of the human-centric communication concept on a range of performance measures. Compared to the condition without information exchange, assisted participants could recover from unawareness much earlier. The approach respects the costs of communication and balances interruptions better than other approaches. Belief inference enables an assistance concept that does not patronize but enables humans to make good decision themselves.**

I. INTRODUCTION

In complex environments, such as driving scenarios, navigation tasks or emergency rescue coordination, many different types of information are important to handle a situation appropriately. To maintain situation awareness for environment and task, information has to be perceived, combined and processed, and finally used to anticipate the evolution of the current situation [1].

With the improvement of technical systems like robots and AI, more interest shifts towards interacting with and assisting humans in complex environments where many aspects need to be taken into account to achieve and maintain situation awareness. There exist many specific assistance approaches designed for single and limited use cases. Some approaches take a very direct approach, telling the human what to do which might prevent immediate errors effectively. However, to help the human to regain awareness and to enable good future decisions it might be better to analyze the problem and provide necessary pieces of information to the human.

On the other side, interrupting a human is not always necessary as not all information is relevant in every situation. Every interaction with a human requires attention and can distract from other important tasks and annoy the human.

For a more general support to a human partner, a robot assistant will have to reason when and what type of information to share, to flexibly support a human partner according to her needs. To enable such human-centric support in different situations, an explicit understanding of the task and human mental processes, a Theory of Mind (ToM), will be helpful to flexibly detect and evaluate current communication needs and effects.

ToM describes the inference of others' mental states such as beliefs, desires or intentions [2]. The development of a Theory of Mind is an important capability for humans and a basis for human interaction and communication [3].

We previously introduced an artificial Theory of Mind that allows us to develop an understanding of the human from interactions online [4]. This can be combined with a task model to evaluate relevance of information and balance possible communication costs and benefits in an explicit way. With this information, a concept of Theory of Mind-based communication was formaly proposed for abstract POMDPs in [5]. This concept considers what humans might already know and what they need to know to handle the current situation appropriately, leading to the decision when and what type of information should be shared to support the receiver's awareness. Here, we evaluate this communication concept on data from human participants performaing a challenging real-world robot collaboration task. An extended version of this work can be found in [6].

II. RELATED WORK

Our human centric communication concept can be divided into two parts, human understanding and the evaluation of communication influence. By interpreting observed human behavior, we estimate information aspects the human might have missed. The evaluation of their importance for the task leads to the decision when and what type of information should be shared. The question how communication should be designed is not in our focus here, explicit as well

[1] Control Methods & Robotics Lab, Technical University Darmstadt, Germany
[2] Honda Research Institute Europe GmbH, Germany
[3] now at Bosch Sensortec GmbH, Germany

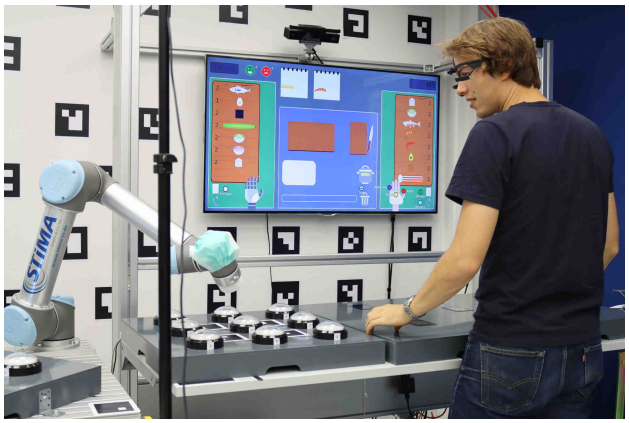

Fig. 1. Human-robot cooperative task setup.

as implicit communication forms have been considered in previous human-robot interaction works [7], [8], [9].

To support a human in one specific decision, classical assistance approaches compare human behavior to some predefined or optimal reference. When a deviation is detected, a correcting action is proposed. This concept was for example applied to healthcare and automotive contexts to prevent human errors [10], [11].

More advanced methods require an explicit understanding of human behavior. One principled approach to understand hidden causes of human behavior such as intentions and beliefs is to collect human trajectories and search for explanations in retrospective. Inverse Reinforcement Learning [12] estimates reward functions that can explain human actions. Baker et al. [13] extend the estimation of human intentions to also account for human state uncertainty. They introduce a perception model and develop a Bayesian observer approach to estimate human desires and beliefs. In repetitive or static settings, such estimated rewards or beliefs can be used to predict human behavior [14], [15].

During interaction human mental states may change and it can be necessary to infer them continuously. In some situations a single hidden state is relevant to coordinate behaviors. Unhelkar et al., [16] estimate the intended next human action based on observations to resolve collisions. Nikolaidis et al., [17] infer human's willingness to adapt to a robot's plan based on their reactions when moving objects together.

The relevance of different information types can change during a task as well. Task relevance is a key aspect in critical situations or in multi agent exploration settings, where others beliefs are known (for example as they have access to non-overlapping parts of the observation space). Here, relevance can be evaluated by the influence of information sharing on other agents belief and expected behavior. This can be used to decide when and what type of information is shared between agents [18], [19], [20]. The relevance can further depend on static preferences of a receiver [21] or the reliability and exclusiveness of the communication channel [22].

When collocated, a robot can observe the human and interpret action and information gathering and the task relevance of information should be combined with an understanding of the human. With the overlap in perception, the estimation of other's knowledge becomes important to avoid sharing information that the other already is aware of.

III. THEORY OF MIND-BASED COMMUNICATION

We will shortly review the general concept of Theory of Mind-based communication (ToM-Com) that was introduced in [5].

To account for a human partner that may not represent every important environmental information correctly, we model the human as an actor in a Partially Observable Markov Decision Process (POMDP) [23]. While not having access to the true state s at all times, it only receives observations o_H (e.g. through visual perception), that may reveal some aspects of the true state according to an observation function $O_H(s, o_H) = p(o_H|s)$. The human can maintain an internal representation of the environment, as a probability distribution over states, $b_H = p(s)$ – the belief. Upon receiving observations, it can update the belief as a Bayesian observer according to the observation likelihood,

$$p(s|o_H) \sim p(o_H|s)p(s). \tag{1}$$

State transitions can be influenced through an action a_H.[1]

$$T(s, a_H, s') = \sum_{a_R} p(s'|s, a_H, a_R)p(a_R). \tag{2}$$

Note that the robot's actions a_R will not be explicitly modeled in our case, but included in the definition of the state transition function

The cooperative task goal is encoded with a reward function $R_H(s, a_H, s')$. The human decision making is modeled to be approximately rational following a softmax policy

$$p(a_H|b_H) \sim \exp(\tau Q_H(b_H, a_H)), \tag{3}$$

where $Q_H(b_H, a_H)$ describes the action value function that is based on the current human belief b_H. This stochastic policy specifically accounts for unmodelled effects in human decision making or action execution.

This model is used to infer the human belief at a given time from observing human decisions and information gathering actions. As such second level inference – inference of human state inference – is computationally hard for larger state spaces, we constrain the representation of the inferred belief. We approximate the belief by factorizing distributions according to the independent state aspects s_i, corresponding e.g. to the sub-state of different task locations, and use a Dirichlet distribution for each factor, $p(b_H) \approx \prod_i Dir(b_{H,i}\alpha_i)$.

By observing the human we can predict belief changes according to the human perception model eq. (1) and state transition eq. (2). Human actions provide feedback from human decision making eq. (3) and are used to update the belief according to $p(b_H|a_H) \sim p(a_H|b_H)p(b_H)$. For these

[1]We use $<>'$ to denote the state of a variable in the next time step

steps we sample the prior human belief estimate, apply the updates, and remap the resulting distribution back to the Dirichlet distributions by moment matching [24].

To decide what and when to communicate, the robot evaluates relevance by estimating communication benefits on the task and comparing it to the cost of communication, optimizing the reward function $R_R = R_H(s, a_H, s') + R_{\text{comm}}(s, a_{\text{comm}})$. The decision for communication actions a_{comm} results from a second POMDP which includes the human belief as part of the robot's state space, $s_R = (s, b_H, a_H)$. The robot needs to estimate communication effects on human beliefs and how this in consequence may change human behavior (eq. (3)). We assume full task knowledge in that the robot can observe the true state and the human action.

The transition function for the communication POMDP combines the human model eqs. (1)-(3), and the influence of communication on the human belief,

$$\begin{aligned} p(s'_R|s_R, a_R) &= p(s', b'_H, a'_H|s, b_H, a_H, a_R) \\ &= \sum_{o'_H, b_{H-}} \Big(\underbrace{p(b'_H|a_H, b_{H-}, o'_H)}_{\text{belief update}} \underbrace{p(o'_H|s')}_{\text{human perception}} \\ &\quad \underbrace{p(s'|s, a'_H)}_{\text{state trans.}} \underbrace{p(a'_H|b_{H-})}_{\text{human decision}} \underbrace{p(b_{H-}|a_R, b_H)}_{\text{comm. effect}} \Big), \quad (4) \end{aligned}$$

where b_{H-} is the intermediate human belief after communication. An information sharing action updates the human belief corresponding to the communicated information and might help the human to chose better actions.

IV. EVALUATION

To evaluate the ToM-Com assistance concept in a real-world setting, we developed a compex human-robot cooperation task and recorded interaction data with human participants.

A. Human-robot cooperation task

Many tasks for human-robot interaction are relatively easy to solve for human participants, since they are designed with a main focus on the robot (see e.g. the tasks in [16], [17]). To evaluate quantitative effects of human support, however, we require a task that challenges the human. Support of a human in dynamic situations is only possible if human awareness problems can be detected from available observations during interaction time (and not only retrospectively), and if a problem in human awareness will influence their behavior over multiple subsequent decisions.

We developed a cooperative meal preparation task where a robot and a human have to assemble different types of Sushi together (Fig. 1).

The task concept shares similarities with the video game Overcooked[2] previously proposed to evaluate human interaction effects [25]. Instead of simulating spatial conflicts, we added an explicit layer of physical human-robot cooperation. The actions of the task are mapped to physical buttons which

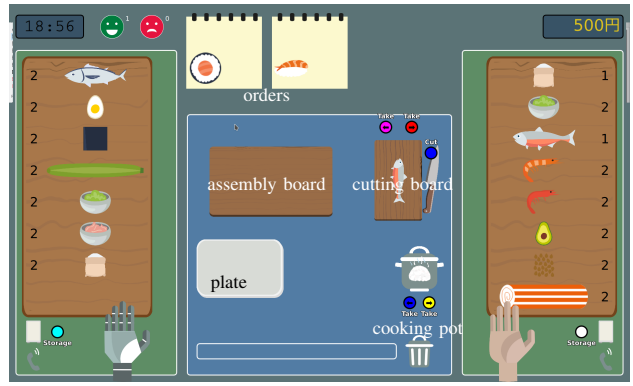

Fig. 2. Task environment visualization. Robot area on the left, human area on the right and shared processing locations in the center. Customer orders are presented at the top.

can be pressed by the agents. We use a UR5 robot, constructed for operating with humans in a common workspace. We use the area around the gaze fixation, measured by eye tracking glasses[3], to estimate an observation update of the human.

Both agents have access to exclusive actions but also to a shared action space (green and blue background respectively in Fig. 2). On average 15 subsequent actions are necessary until an order can be served, involving ingredients with partially similar look and receipes with opposing order of steps. This causes challenges for the human in memorizing different recipes, differentiating ingredients, planning actions, and coordinating with the robot. Coordination is necessary on the strategic level (which order to prepare first) as well as on the action level (who contributes which ingredient).

We assume that the human can be uncertain about the content of different locations, the recipes for the different sushi types, and the current robot action. The human belief model includes these aspects, however, for the recipes, we only represent 4-5 erroneous variations that where commonly found in a short pre-study.

During the experiment, the robot solves the task MDP for its best available action in a given time step, considering the human to have true believe and behave optimally. We set rewards of +10 for correct and -10 for incorrect suhsi delivery and a step penalty of -1.

We consider two types of communication signals aimed to provide information to correct an erroneous human belief, one to correct a false belief about required ingredients and a second to improve the belief on robot actions or order of steps, by drawing attention to the content of a specific location. A communication signal is assumed to update of the corresponding human belief aspects (recipe and order aspect for showing recipe, location contents for highlighting) to the true state of the respective aspect.

[2]https://www.team17.com/games/overcooked/

[3]Pupil Labs: https://pupil-labs.com/products/core/

B. Results

We asked 14 participants to cooperate with the robot on our sushi task to collect data which we then analyzed with respect to the potential of preventing errors through our communication concept. Participants started a session with task instructions and a familiarization phase until they signalled to feel comfortable with the task. Overall session duration per participant was limited to 1 hour.

The task challenged the participants and generated situations where they were not aware of important aspects. Out of 8515 recorded actions, 587 can be considered as errors, i.e. the selected action diverted from the optimal action in the given situation 481 of those were retrospectively classified by a human expert (the first author), considering subsequent human actions, to be caused by belief related awareness problems. The remaining errors (i.e., seemingly random actions that were immediately canceled) were probably caused by color mismatch or erroneous button presses. A false belief of one important aspect (as the current recipe) normally lead to a longer sequence of errors, e.g. where the participant works with wrong ingredients. Accordingly, the single action errors were clustered into 153 error sequences of actions with an expected common cause. When looking at the length of such error sequences, we find that almost 25% of the sequences contain only a single error. In such cases, it is very challenging for any assistance system to have a meaningful communication support after detecting an error, but we also encountered a number of longer sequences.

We hypothesized that human centric assistance (ToM-Com) can improve the performance of participants in the cooperative sushi task. As performance measures we considered true positive (potentially prevented errors divided by total errors) and false positive rate (unnecessary interruptions divided by number of optimal human actions). We let the system, for each step in the data, compute the human belief and select a communication action As an upperbound of the effect of the communication signal, all subsequent errors, labeled to have a common cause, were counted as being prevented.

We compared the communication decisions to a state of the art assistive communication concept and a Theory of Mind-based concept without communication planning. Typical human centric assistance approaches intervene after the human deviates from an expected behavior by proposing k good next action. For the analysis, we assumed that participants would follow the action proposal and that this deviation based concept (DEV) would prevent the k subsequent human error.

The ToM based alternative without communication planning decides to communicate whenever a false or uncertain human belief is detected (similar to previous work in [4]) based on a threshold θ_{ToM} for the probability of the correct belief. This should prevent human errors similar to ToM-Com (i.e. prevent subsequent errors with the same cause), but may lead to unnecessary communication, when the false belief aspect is not relevant for current decisions.

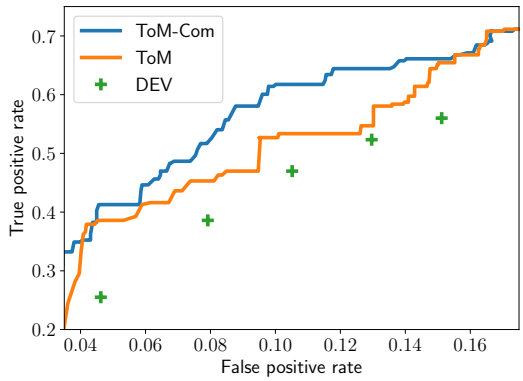

Fig. 3. ROC curves of human error prevention for the three communication concepts.

The different concepts, ToM-Com, DEV and ToM, were compared by varying the decision parameters of the concepts to create a receiving operator characteristic (ROC) (Fig. 3). For the ToM-Com concept we varied the cost of communitation between 0.5 and 2.5, for DEV, $k = [1:5]$ and the decision threshold θ_{ToM} was sampled between 0.4 and 0.99.

Comparing the different concepts, ToM-Com outperformed the other concepts. It could prevent more human errors while reducing unnecessary disturbances for the human partner. Communication with belief inference but without communication planning could achieve better results than the purely reactive deviation-based approach, but would sometimes lead to avoidable interruptions ("False Positives"). The best performance of any setup only reached slightly more than 70%, however most of the remaining cases were actually single error sequences, where prior interactions do often not provide sufficient hints for predictive assistance.

V. CONCLUSION

We showed a new human-centric communication concept to assist in collaborative robot task. Based on an understanding of the human behavior through the estimation of the human belief, the relevance of different types of information is evaluated leading to the decision when and what type of information to share to support a human partner. We designed a challenging Sushi task, and found a significant number of errors in a recording with human participants. We could show significant performance improvements of our Theory of Mind-based communication concept compared to conditions without communication planning and to reactive assistance. It could prevent more human errors, while avoiding unnecessary disturbances. We proposed a new class of assistance systems, which aim to treat the causes of human errors and not the symptoms, to empower the human for improved decision making. We believe that our generic human centric communication concept enables new types of intelligent decision support systems in different application fields. In future work we will test the ToM-Com system online with explicit commnication signals while participants are performing the task to validate the results in this paper.

REFERENCES

[1] M. R. Endsley, "Toward a theory of situation awareness in dynamic systems," *Human Factors*, vol. 37, no. 1, pp. 32–64, 1995.

[2] H. Wimmer and J. Perner, "Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception," *Cognition*, vol. 13, no. 1, pp. 103–128, 1983.

[3] F. G. Happé, "Communicative competence and theory of mind in autism: A test of relevance theory," *Cognition*, vol. 48, no. 2, pp. 101–119, 1993.

[4] M. C. Buehler and T. H. Weisswange, "Online inference of human belief for cooperative robots," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2018)*. IEEE, 2018, pp. 409–415.

[5] ——, "Theory of mind based communication for human agent cooperation," in *IEEE International Conference on Human-Machine Systems (ICHMS 2020)*. IEEE, 2020, pp. 1–6.

[6] M. C. Buehler, J. Adamy, and T. H. Weisswange, "Theory of Mind Based Assistive Communication in Complex Human Robot Cooperation," sep 2021. [Online]. Available: https://arxiv.org/abs/2109.01355v1

[7] C. Torrey, S. R. Fussell, and S. Kiesler, "How a robot should give advice," in *8th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2013)*. IEEE, 2013, pp. 275–282.

[8] A. Dragan and S. Srinivasa, "Integrating human observer inferences into robot motion planning," *Autonomous Robots*, vol. 37, no. 4, pp. 351–368, 2014.

[9] S. Reddy, S. Levine, and A. D. Dragan, "Assisted perception: optimizing observations to communicate state," *arXiv preprint arXiv:2008.02840*, 2020.

[10] J. Hoey, C. Boutilier, P. Poupart, P. Olivier, A. Monk, and A. Mihailidis, "People, sensors, decisions: Customizable and adaptive technologies for assistance in healthcare," *ACM Transactions on Interactive Intelligent Systems*, vol. 2, no. 4, pp. 1–36, 2013.

[11] J. C. McCall and M. M. Trivedi, "Driver behavior and situation aware brake assistance for intelligent vehicles," *Proceedings of the IEEE*, vol. 95, no. 2, pp. 374–387, 2007.

[12] A. Y. Ng, S. J. Russell, *et al.*, "Algorithms for inverse reinforcement learning." in *17th International Conference on Machine Learning (ICML 2000)*, 2000.

[13] 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, pp. 1–10, 2017.

[14] D. Sadigh, S. S. Sastry, S. A. Seshia, and A. Dragan, "Information gathering actions over human internal state," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2016)*. IEEE, 2016, pp. 66–73.

[15] E. Ferreira, G. Milliez, F. Lefèvre, and R. Alami, "Users' belief awareness in reinforcement learning-based situated human–robot dialogue management," in *Natural Language Dialog Systems and Intelligent Assistants*. Springer, 2015, pp. 73–86.

[16] V. V. Unhelkar, S. Li, and J. A. Shah, "Decision-making for bidirectional communication in sequential human-robot collaborative tasks," in *ACM/IEEE International Conference on Human-Robot Interaction (HRI 2020)*, 2020, pp. 329–341.

[17] S. Nikolaidis, M. Kwon, J. Forlizzi, and S. Srinivasa, "Planning with verbal communication for human-robot collaboration," *ACM Transactions on Human-Robot Interaction*, vol. 7, no. 3, pp. 1–21, 2018.

[18] C. V. Goldman and S. Zilberstein, "Optimizing information exchange in cooperative multi-agent systems," in *2nd International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2003)*, 2003, pp. 137–144.

[19] F. S. Melo, M. T. Spaan, and S. J. Witwicki, "Querypomdp: Pomdp-based communication in multiagent systems," in *European Workshop on Multi-Agent Systems*. Springer, 2011, pp. 189–204.

[20] M. Roth, R. Simmons, and M. Veloso, "What to communicate? execution-time decision in multi-agent pomdps," in *Distributed autonomous robotic systems 7*. Springer, 2006, pp. 177–186.

[21] R. Chitnis, L. P. Kaelbling, and T. Lozano-Pérez, "Learning what information to give in partially observed domains," in *2nd Conference on Robot Learning (CoRL 2018)*. PMLR, 2018, pp. 724–733.

[22] J. Renoux, T. S. Veiga, P. U. Lima, and M. T. Spaan, "A unified decision-theoretic model for information gathering and communication planning," in *29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN 2020)*. IEEE, 2020, pp. 67–74.

[23] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra, "Planning and acting in partially observable stochastic domains," *Artificial Intelligence*, vol. 101, no. 1-2, pp. 99–134, 1998.

[24] T. Minka, "Estimating a dirichlet distribution," 2000.

[25] J. Bishop, J. Burgess, C. Ramos, J. B. Driggs, T. Williams, C. C. Tossell, E. Phillips, T. H. Shaw, and E. J. de Visser, "Chaopt: a testbed

for evaluating human-autonomy team collaboration using the video game overcooked! 2," in *Systems and Information Engineering Design Symposium (SIEDS 2020)*. IEEE, 2020, pp. 1–6.