MUTUAL MODELLING ABILITY FOR A HUMANOID ROBOT
HOW CAN IT IMPROVE MY LEARNING AS WE SOLVE A PROBLEM TOGETHER?

A PREPRINT

Utku Norman ⓘ
CHILI Lab
EPFL
Lausanne, Switzerland
utku.norman@epfl.ch

Barbara Bruno ⓘ
CHILI Lab
EPFL
Lausanne, Switzerland
barbara.bruno@epfl.ch

Pierre Dillenbourg
CHILI Lab
EPFL
Lausanne, Switzerland
pierre.dillenbourg@epfl.ch

February 17, 2021

1 Background and Research questions

Robot-mediated learning activities are often designed as collaborative exercises where two or more children work together to achieve the activity objectives. Witnessing such activities allows for making an interesting observation: although miscommunications and misunderstandings occur frequently, humans, even at young ages, are very good at understanding each other. Humans, unlike robots, are highly skilled in detecting and addressing misunderstandings. We represent whether the others understood what we said or did by using the complex cognitive ability of *mutual modelling*[1], i.e. the reciprocal ability to construct a *mental representation* of the other, by attributing beliefs, desires and other mental states to the other [10]. This ability is critical in order for humans to comprehend each other and react appropriately in their interactions. Thus, the main goal of my PhD is to equip a robot with mutual modelling ability, and use this ability in an educational activity in order to improve the quality of the interactions between the robot and a learner and (hopefully) the learning outcomes.

1.1 What research question are you trying to address?

We target the following research question: "How can we equip a robot with mutual modelling ability, so that it can build and maintain a second-order mental representation of a human in the context of a collaborative learning activity, and use this ability to improve the quality of interactions and (hopefully) the learning outcomes?".

1.2 What learning theory or theoretical background do you plan to use?

Collaboration happens within a collaborative activity, that is a situation in which individuals work together as group members in order to solve a problem [11], where the members in the group build a *shared understanding* of the problem at hand [12]. The case where collaboration is in an educational situation is termed as *collaborative learning*: individuals learn or attempt-to-learn something together, where their interactions are anticipated to trigger processes that would induce learning, even though there is no guarantee that these kind of interactions will happen [9]. These productive interactions among learners can be 'designed' as they are shaped by the activity and the environment, while it is the required *effort* to construct the shared understanding together that results in collaborative learning [13]. We believe that a robot equipped with mutual modelling ability, while participating in a collaborative learning activity, can act as an adaptive, intelligent component to complement the designed activity environment, and can help trigger learning mechanisms by supporting the effort to build a shared understanding.

---

[1]Similar terms include: theory of mind [1, 2, 3] (for robots as in [4]), mentalising (e.g. of humans about an agent [5]), mindreading (as in the HRI workshop [6]), adopting the 'intentional stance' [7], and social cognition. For instance, social cognition can be preferred over theory of mind, as the latter may bring in undesired theoretical commitments [8]. We prefer mutual modelling as it highlights the mutuality and modelling aspects, while linking work in collaborative learning as in [9] to HRI as in [10].

Theories of individual cognitive development could give insight into the learning processes that can occur in collaborative learning situations. Linking cognitive development with social interaction goes back to the works of Mead (posthumously in [14]), Piaget (e.g. [15]), and Vygotsky [16, 17]. These bring about Vygotsky's idea that "what children can do with others today, they can do alone tomorrow" [16]: socio-cultural theories, as built on Vygotsky's ideas [16, 17], present cognitive development as a result of social and cultural experience, that is socially guided and constructed through interactions between individuals [18]. The causal direction of an individual's cognitive development for Vygotsky is from "outside in", with the social experience initiating the development, whereas in Piaget's account development proceeds with the child adapting and hence it is more from the inside out [19]. Alternatively, the causality may rather be bi-directional, or reciprocal: development "progresses in both a circular and spiral fashion", with the social interaction improving the individual, who can in return take part in more sophisticated interactions, and so on [20]. Following this line of thought, an account for the individual development is by *socio-cognitive conflict* [21, 20], and its regulation [22]. One of the mechanisms that can bring about development in a conflict is *appropriation*, which allows the learner to consider the experienced social interaction to reexamine his/her own understanding [23]. In our opinion, a robot with mutual modelling ability can deliberately create a socio-cognitive conflict, so that the learner can be nudged towards an effort for instance in order to appropriate the interaction, towards building a shared understanding.

A major perspective for equipping robots with mutual modelling ability builds on research in psycholinguistics, i.e. the psychology of language [10]. As we communicate with each other, we produce a series of utterances that make up a *dialogue*: rather than an individual effort, a dialogue is a joint activity performed by two or more participants [24, 25], who try to reach a shared, mutual understanding of meanings of the utterances to a sufficient degree for the current purposes [26]. Building this shared understanding has been studied under the notion of *grounding* [27]. This notion does not immediately apply to collaborative learning, because what is 'shared' is at a different scale: psycholinguistics considers short dialogue episodes where a single referent is grounded, whereas in collaborative learning we have longer episodes to develop a shared conception of a domain [28]. Thus, the goal of collaborative learning can be seen as to trigger *conceptual change*, where the learners move from no conceptions or misconceptions to correct conceptions. As a computational account of grounding that can be used by an agent (and hence a robot) to track and perform grounding (and thus 'add' to the shared understanding, akin to to [26]), Traum presents a multi-stratal theory of action in dialogue, which includes a level of grounding acts: it is based on *dialogue acts* that generalise speech act theory to dialogue [29, 30], where speech acts go back to Austin, who observed utterances as performing actions [31]. We propose that a collaborative activity by design can incorporate actions that implement a relevant set of dialogue acts as affordances available to the learner and the robot: a robot equipped with mutual modelling ability can interact with a learner primarily through the activity[2] in order to build a shared understanding, by taking actions and keeping track of actions taken by the learner. Within the context of the activity, the robot can recursively represent, update, and reason with (higher-order) beliefs: these can include beliefs ascribed (by the robot) to the learner about the activity (first-order mental representation), as well as those ascribed to the learner about the robot's beliefs about the activity (second-order mental representation). With such a mental representation, the robot can infer the learner's misconceptions about the activity, misunderstandings and (dis)agreements between the learner and itself, and can have an opinion (that need not be complete, or even correct) as it participates in the co-construction of a shared understanding about the activity.

A family of computational methods for recursive reasoning in which agents can model other agents, as surveyed in [34], builds on Markov Decision Processes (MDPs): MDP is a general formal framework for sequential decision making under uncertainty that has been studied since the 50s [35]. It has been generalised to partially-observable contexts (POMDPs [36]), that could be used by a robot to represent the beliefs of a learner e.g. about the activity, since they are not directly observable by the robot. Other relevant extensions include: i) to multiagent contexts (e.g. Interactive-POMDPs [36]) with which a robot can explicitly account for the presence of multiple agents, which is exactly the case in collaborative activities; ii) to recursive, higher-order reasoning (finitely-nested I-POMDPs [37]), via which a robot can represent beliefs of the form "I believe that the learner believes that I believe ..." that would allow the robot to have a higher-order mental representation, and; iii) to communicative actions (Communicative I-POMDPs [38]) that could incorporate dialogue acts such as grounding acts performed by the interacting agents. The common use of MDPs is planning for optimal action, by assuming rationality and maximising the cumulative reward, i.e. the sum of (discounted) immediate rewards. We believe that in a learning activity, this framework could be re-purposed to plan not for completing the task in a shorter time or fewer steps, but rather (to try) to improve learning, by automatically selecting actions that e.g. could reveal or challenge the (inferred) misconceptions held by the learner, and induce a socio-cognitive conflict. Thus, via automated planning within an MDP-based framework, a robot can choose its actions

---

[2]The actions are possibly complemented by verbal interaction. Nevertheless, the idea here is not to rely on speech recognition in principle, especially given the limitations of the state-of-the-art tools for natural language understanding that can be utilised by a robot. Indeed, another approach would be to primarily focus on spoken dialogue: for instance, a model to ground the intentions of a human user, by incrementally processing the dialogue acts that are detected from a voice interface, has been developed and experimented with on a pick-and-place robot [32, 33]. Instead, we focus on an action understanding approach, where dialogue acts are incorporated into the task itself as task actions, so that certain aspects of the social interaction happen through the task.

from its state-based model, where the state contains its higher-order beliefs about the activity and the learner: it can consider the future consequences of an action, and exhibit adaptive/learning behaviour; instead of reflex-based behaviour with possibly fixed rules which might rather be inferior with regards to the quality of interactions, and learning.

2 Design

2.1 What is the context to your research?

We build on a collaborative problem solving activity for school children named JUSThink[3] [39, 40], which aims to improve their Computational Thinking (CT) skills by applying abstract and algorithmic reasoning to solve an unfamiliar problem on networks[4]. In particular, the activity exposes school children to the minimum-spanning-tree problem on graphs, that they are not expected to be familiar with.

In this revision of the JUSThink activity, a learner and a humanoid robot are given a network of gold mines, with possible positions for railway tracks, where each track if it is built connects one mine to another. They are asked to collect the gold by connecting the gold mines to each other, while spending as little as possible to build the tracks[5]. The learner and the robot as same-status peers collaboratively construct a solution by deciding together which tracks to build, and submit it as their solution. The cost of each track is visible to both the learner and the robot. The learner and the robot will take turns in suggesting which connection to pick and build via performing a suggest-a-pick action, complemented by other actions that have some communicative function[6]. A track will be built only if it is suggested by one and accepted by the other[7]. When the robot and the learner agree on a solution, they can submit it and receive feedback on whether it is a correct solution or not. They can submit a solution up to a fixed number of times (e.g. only two times), and then, the joint activity concludes. The goal of the robot is to automatically select which action to perform, while updating[8] and considering its beliefs, and construct a solution with the learner together.

2.1.1 Who are your learners?

The intended learners are English-speaking children aged 8 to 12 years old, that corresponds to middle-school children at international schools in Switzerland.

2.1.2 What are the learning objectives?

After completing this activity, the learner will be able to correctly choose a subset of connections on a given network, so that all nodes are connected to each other and the total cost on these connections is minimised (i.e. minimum-spanning-tree problem on graphs).

2.1.3 Where is the learning occurring? (home, school, elderly facility …)

The learning is in extracurricular exercise sessions with the robot and a learner solving the given problem together. This can happen in the lab, in school, or even at home.

---

[3]See the project website at `https://www.epfl.ch/labs/chili/index-html/research/animatas/justhink/`.

[4]Accordingly, recent studies emphasise a need to introduce CT earlier in schools [41].

[5]Note that the gold mines connected by railway tracks is the scenario that was adopted in [39, 40], where an alternative among many others could be cities being connected by electric transmission grids, and the connected cities lighting up and glowing. The exact scenario is to be decided and established e.g. in pre-experiments, keeping in mind possible familiarity and intuitiveness: for instance, connected cities to glow could be an intuitive cue to immediately see the connected, and distinguish from the disconnected.

[6]Performing suggest-a-pick action on a connection A-B can be verbalised in the form "Shall we pick A-B?", and could indicate that the performer believes "Picking A-B is correct" (and hence it is part of a correct solution). Another plausible action, guess-my-pick, can be verbalised as "What do you think I will pick now?", and would reveal your beliefs about my beliefs. What types of actions are planned to be incorporated into the activity, and how they relate to dialogue acts as well as the mental representation is currently in development, and a through discussion on them is beyond the scope of this paper.

[7]To illustrate how the actions on the task could relate to the dialogue acts, consider the suggest-a-pick action: it is akin to the `suggest` act as in [30], where one makes a proposal that could be taken up by the other, or not. For instance, given that a track can be built only if it is suggested by one and accepted by the other, we will have an `accept` act, by e.g. being 'suggested again', or otherwise a `reject` act, performed e.g. by suggesting another connection.

[8]The robot can update its mental representation by keeping track of which action is taken by who, by following simple rules based on two principles: i) your actions are cues for me to update my beliefs about your beliefs, of the form "I believe that you believe X." (i.e. my first order beliefs), and i) my actions are (giving you) cues, for me to update my beliefs about your beliefs about my beliefs, of the form "I believe that you believe that I believe X." (i.e. my second order beliefs). In the case of suggest-a-pick action for a connection A-B, X could be that "Picking A-B is correct" (i.e. part of a correct solution): then, if the child performs suggest-a-pick action, the robot can (rationally) form the belief "I believe that you believe picking A-B is correct.".

2.1.4 What robot do you use and why?

Our perception about the robot is influenced by the specific robot and its embodiment: thus, necessary features for the robot include a suitable type of embodiment. We prefer a humanoid form, assuming that it could be more likely for us humans to ascribe beliefs to a humanoid robot.

There are two alternative robots in consideration: QTrobot[9] and Reachy[10]. The JUSThink study in [39, 40] was performed with QTrobot, albeit in a different role in the activity. We have seen that the robot was positively perceived by the same target learner group (as also measured by a self-perception questionnaire, on e.g. items based on Godspeed, as reported in [39]). It is a quite stable and reliable platform that can operate for longer hours into consecutive experiments (with different learners), as it is fixed to a position with non-actuated legs: we have faced no serious problems due to e.g. heating through two weeks of experiments. The robot is designed to be quite safe, with a plastic cover and no sharp parts, and safety mechanisms to turn off motor torque if an external force is applied above a certain limit: although we have not used (and do not plan on using) it for physical interaction, this comes as an important feature as the robot is operating with and near children. In addition, QTrobot is simple to program and interface with (via ROS), and has a behaviour library of of modest size that contains facial expressions, and specifically gestures that can be expanded on quite easily. On the other hand, Reachy allows a much finer/precise control of the arms, that can be used to simulate very lifelike gestures of acting on the same screen as it works with a learner, as well as deictic gestures to point to and higlight certain parts of the problem. Given that behaviour libraries for Reachy are not readily available, we would like to have a complete study with the QTrobot as we develop on the libraries for Reachy in parallel, and then move to the Reachy as the robot to interact with the learner.

3 Assessment

3.1 What methods do you plan to use for the assessment? How do they map with the learning objectives?

To measure the learning effects, we will focus on assessment of the differential effects, by computing the learning gain from the pre-test and the post-test scores. The tests could consist of a new problem instance to be solved by the learner by him/herself which directly maps to the learning objective, as well as additional questions on the fundamental notions behind the minimum-spanning-tree problem that he/she will/has worked on, such as the existence, presence, or minimumness of a selection (see [39] for an explanation of these notions).

Quality of interactions as a construct will be measured through multiple ways: through analysing the i) interaction logs, as an objective account of what actually happened, in terms of which actions were taken in which situation, from which we can compute a measure of performance in the task based on e.g. their submitted solutions, and ii) self-perception questionnaire, that passes through the filter of self perception and assessment.

4 Acknowledgements

---

[9]By LuxAI, see `https://luxai.com/qtrobot-expressive-robot-for-research-and-teaching/`
[10]By Pollen Robotics, see `https://www.pollen-robotics.com/reachy/`

References

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

[2] Heinz Wimmer and Josef Perner. Beliefs about beliefs: Representation and constraircing function of wrong beliefs in young children's understanding of deception. *Cognition*, (13):103–128, 1983. doi: 10.1002/app.25226.

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

[4] Brian Michael Scassellati. Theory of Mind for a Humanoid Robot. *Autonomous Robots*, (12):13–24, 2002. doi: 10.1080/2040350X.2018.1469197.

[5] Chris L. Baker, Julian Jara-Ettinger, Rebecca Saxe, and Joshua B. Tenenbaum. Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, 1:0064, 2017. doi: 10.1038/s41562-017-0064.

[6] Agnieszka Wykowska, Jairo Perez-Osorio, and Stefan Kopp. Social Cognition for HRI: Exploring the Relationship between Mindreading and Social Attunement in Human-Robot Interaction. In *Proceedings of 2020 ACM International Conference on Human-Robot Interaction (HRI '20 Companion), March 23–26, 2020, Cambridge, United Kingdom*, New York, NY, USA, March 2020. ACM. doi: 10.1145/3371382.3374851.

[7] Daniel C. Dennett. *The Intentional Stance*. MIT Press, 1987.

[8] Rineke Verbrugge. Logic and Social Cognition. *Journal of Philosophical Logic*, 38(6):649–680, December 2009. doi: 10.1007/s10992-009-9115-9.

[9] Pierre Dillenbourg. What do you mean by collaborative learning? In Pierre Dillenbourg, editor, *Collaborative-learning: Cognitive and Computational Approaches*, pages 1–19. Oxford: Elsevier, 1999.

[10] S Lemaignan and P Dillenbourg. Mutual Modelling in Robotics: Inspirations for the Next Steps. In *HRI '15: Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, pages 303–310, 2015. doi: 10.1145/2696454.2696493.

[11] Jeremy Roschelle and Stephanie D. Teasley. The Construction of Shared Knowledge in Collaborative Problem Solving. *Computer Supported Collaborative Learning*, 128:69–97, 1995. doi: 10.1007/978-3-642-85098-1_5. Publisher: Springer.

[12] Daniel L Schwartz. The emergence of abstract representations in dyad problem solving. *Journal of the Learning Sciences*, 4(3):321–354, 1995. doi: 10.1207/s15327809jls0403_3. Publisher: Taylor & Francis.

[13] Pierre Dillenbourg, Sanna Järvelä, and Frank Fischer. The Evolution of Research on Computer-Supported Collaborative Learning. In Nicolas Balacheff, Sten Ludvigsen, Ton de Jong, Ard Lazonder, and Sally Barnes, editors, *Technology-Enhanced Learning: Principles and Products*, pages 3–19. Springer Netherlands, Dordrecht, 2009. doi: 10.1007/978-1-4020-9827-7_1.

[14] George H. Mead. *Mind, Self, and Society: from the Standpoint of a Social Behaviorist*. University of Chicago Press, 1st edition, 1934.

[15] Jean Piaget. *Genetic Epistemology*. W. W. Norton & Company, Inc, 1971.

[16] Lev S. Vygotsky. *Thought and Language*. The MIT Press, 1986. Note: Original work Myshlenie i rech published in 1934.

[17] Lev S. Vygotsky. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, 1978. Note: ca. 1930-1934.

[18] Mary Gauvain. Vygotsky's Sociocultural Theory. In Janette B. Benson, editor, *Encyclopedia of Infant and Early Childhood Development (Second Edition)*, pages 446–454. Elsevier, Oxford, January 2020. doi: 10.1016/B978-0-12-809324-5.23569-4.

[19] Rheta DeVries. Vygotsky, Piaget, and Education: a reciprocal assimilation of theories and educational practices. *New Ideas in Psychology*, 18(2):187–213, August 2000. doi: 10.1016/S0732-118X(00)00008-8.

[20] Willem Doise and Gabriel Mugny. *The social development of the intellect*, volume 10 of *International series in experimental social psychology*. Pergamon Press, 1984.

[21] Gabriel Mugny and Willem Doise. Socio-cognitive conflict and structure of individual and collective performances. *European journal of social psychology*, 8(2):181–192, 1978. doi: 10.1002/ejsp.2420080204. Publisher: Wiley Online Library.

[22] Fabrizio Butera, Nicolas Sommet, and Céline Darnon. Sociocognitive conflict regulation: How to make sense of diverging ideas. *Current Directions in Psychological Science*, 28(2):145–151, 2019. doi: 10.1177/0963721418813986.

[23] Barbara Rogoff. *Apprenticeship in Thinking: Cognitive Development in Social Context*. Oxford University Press, 1st edition, 1990.

[24] Herbert H. Clark and Deanna Wilkes-Gibbs. Referring as a collaborative process. *Cognition*, 22(1):1–39, 1986. doi: 10.1016/0010-0277(86)90010-7.

[25] Herbert H. Clark. *Using Language*. Cambridge University Press, 1996.

[26] Herbert H. Clark and Edward F. Schaefer. Contributing to discourse. *Cognitive Science*, 13(2):259–294, 1989. doi: 10.1016/0364-0213(89)90008-6.

[27] Herbert H. Clark and Susan E. Brennan. Grounding in communication. In *Perspectives on socially shared cognition*, pages 127–149. American Psychological Association, Washington, DC, US, 1991. doi: 10.1037/10096-006.

[28] Pierre Dillenbourg and David Traum. Sharing Solutions: Persistence and Grounding in Multimodal Collaborative Problem Solving. *Journal of the Learning Sciences*, 15(1):121–151, 2006. doi: 10.1207/s15327809jls1501_9. Publisher: Routledge.

[29] David R. Traum and Elizabeth A. Hinkelman. Conversation Acts in Task-Oriented Spoken Dialogue. *Computational Intelligence*, 8(3):575–599, 1992. doi: 10.1111/j.1467-8640.1992.tb00380.x.

[30] David R Traum. *A Computational Theory of Grounding in Natural Language Conversation*. Thesis, University of Rochester. Computer Science Department, 1994. URL https://apps.dtic.mil/sti/citations/ADA289894.

[31] John Langshaw Austin. *How to do things with words*. The William James Lectures delivered at Harvard University in 1955. Oxford University Press, 1962.

[32] Julian Hough and David Schlangen. A Model of Continuous Intention Grounding for HRI. In *The Role of Intentions in Human-Robot Interaction" workshop in conjunction with the 12th ACM / IEEE International Conference on Human-Robot Interaction (HRI 2017)*, Vienna, 2017. URL https://pub.uni-bielefeld.de/record/2909062.

[33] Julian Hough and David Schlangen. It's Not What You Do, It's How You Do It: Grounding Uncertainty for a Simple Robot. In *HRI '17: Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, pages 274–282, March 2017. doi: 10.1145/2909824.3020214.

[34] Stefano V. Albrecht and Peter Stone. Autonomous agents modelling other agents: A comprehensive survey and open problems. *Artificial Intelligence*, 258:66–95, May 2018. doi: 10.1016/j.artint.2018.01.002.

[35] Richard E Bellman. *Dynamic programming*. Princeton University Press, Princeton, New Jersey, 1957.

[36] Leslie Pack Kaelbling, Michael L Littman, and Anthony R Cassandra. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1):99–134, 1998. doi: 10.1016/S0004-3702(98)00023-X.

[37] Piotr J Gmytrasiewicz and Prashant Doshi. A framework for sequential planning in multi-agent settings. *Journal of Artificial Intelligence Research*, 24:49–79, 2005. doi: 10.1613/jair.1579.

[38] Piotr Gmytrasiewicz and Sarit Adhikari. Optimal Sequential Planning for Communicative Actions: A Bayesian Approach. In *AAMAS '19: Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, pages 1985–1987, Montreal QC, Canada, May 2019. International Foundation for Autonomous Agents and Multiagent Systems. URL https://dl.acm.org/doi/abs/10.5555/3306127.3331985.

[39] Jauwairia Nasir, Utku Norman, Barbara Bruno, and Pierre Dillenbourg. "You Tell, I Do, and We Swap until we Connect All the Gold Mines!". *ERCIM News*, 2020(120), 2020. URL https://ercim-news.ercim.eu/en120/special/you-tell-i-do-and-we-swap-until-we-connect-all-the-gold-mines.

[40] J. Nasir, U. Norman, B. Bruno, and P. Dillenbourg. When Positive Perception of the Robot Has No Effect on Learning. In *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pages 313–320, August 2020. doi: 10.1109/RO-MAN47096.2020.9223343.

[41] Divya Menon, Sowmya Bp, Margarida Romero, and Thierry Viéville. Going beyond digital literacy to develop computational thinking in K-12 education. *Smart Pedagogy of Digital Learning*, 2019. doi: 10.4324/9780429319501-2. Publisher: Taylor & Francis (Routledge).