Chapter 3

Taking Humans Mental States into account while executing Shared Plans

Contents

3.1 Motivations

When collaborating with humans, it is primordial for the robot to not consider humans as obstacles or tools impacting the environment. As humans are social creatures, the robot must take into account their comfort and so, their point of view. Several works already allow robots to estimate humans perspective and beliefs concerning its environment. In order to improve human-robot Joint Action, the robot must be able to take these information into account when taking decision on how to act or what to communicate. Even if several works have been done on how to integrate humans perspective in dialogue or use it to help the understanding of humans behavior, there is still a gap when it came to use it during Shared Plan execution. This work aims to start filling this gap by extending the robot knowledge on humans mental states to the joint task and using it to better communicate during Shared Plan execution. It has been the subject of a publication into the HRI conference [Devin 2016].

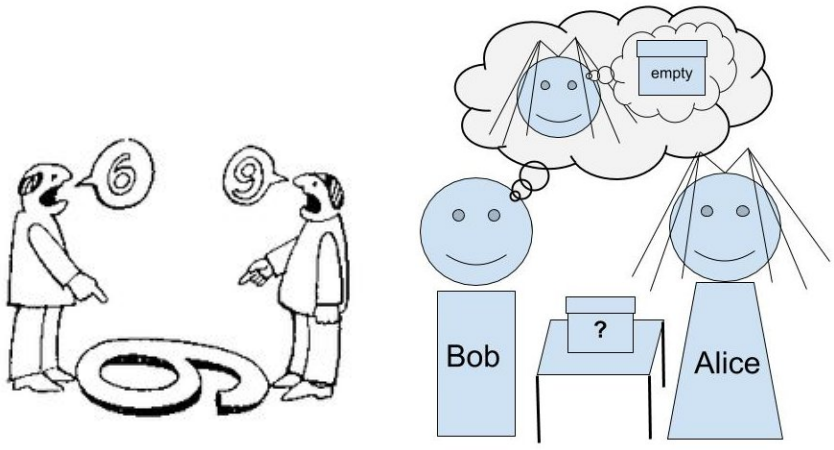
3.2 Theory of Mind

3.2.1 Social Sciences literature

Theory of the Mind (ToM) refers to the ability humans have to recognize and attribute mental states not only to themselves but to other people, to understand that feelings and beliefs we have may be different than others and to take others mental states into account. ToM has been deeply studied in psychology, notably in the developmental domain [Baron-Cohen 1985, Premack 1978]. [Verbrugge 2008] defines what is called "order" of ToM:

> "To have a first-order ToM is to assume that someone's beliefs, thoughts and desires influence one's behavior. A first-order thought could be: 'He does not know that his book is on the table'. In second-order ToM it is also recognized that to predict others' behavior, the desires and beliefs that they have of one's self and the predictions of oneself by others must be taken into account. So, for example, you can realize that what someone expects you to do will affect his behavior. For example, '(I know) he does not know that I know his book is on the table' would be part of my second-order ToM. To have a third-order ToM is to assume others to have a second-order ToM, etc."

There is an infinitesimal numbers of orders, however, studied shown that orders above the second one do not help in cooperative tasks [De Weerd 2014] and above the third one do not help for competitive games [De Weerd 2014].



(a) Perceptual perspective taking: two individuals can have a different representation of their environment considering their locations.

(b) Conceptual perspective taking: here Bob attributes to Alice a belief concerning the box. He thinks Alice thinks the box is empty.

Figure 3.1: Illustration of perceptual and conceptual perspective taking.

ToM includes the notion of perspective taking: the capacity for a person to reason by taking the point of view of someone else. Studied in literature

[Tversky 1999, Flavell 1992], perspective taking is crucial during humans interaction and studies have demonstrate that individuals who lack of this ability have difficulties in their daily social interactions [Frick 2014]. Two levels of perspective taking are defined in [Flavell 1977]: perceptual and conceptual perspective taking. Perceptual perspective taking design the capacity of a person to understand that others have a different perception of the world (fig 3.1(a)). Conceptual perspective taking designs the capacity of a person to attribute beliefs and feelings to others (fig 3.1(b)).

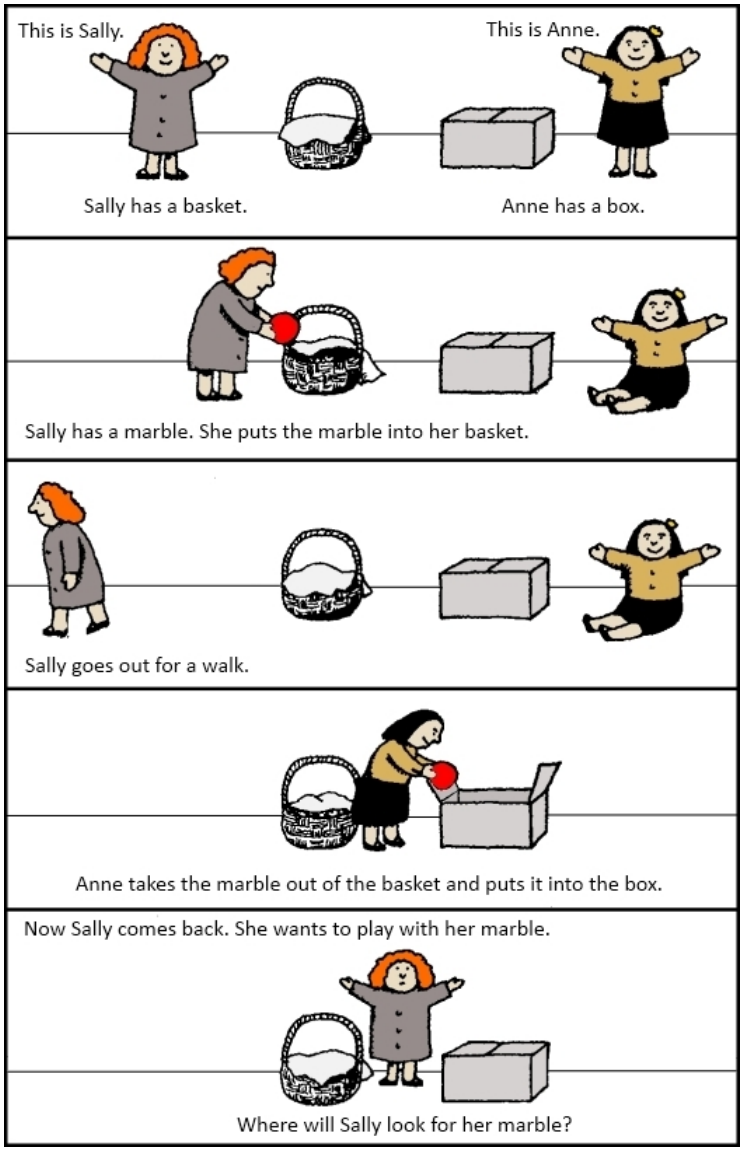


Figure 3.2: The Sally and Anne test: it allows to check the capacity of someone to attribute a false-belief to another person. Illustration from the work of Axel Scheffler.

To check if an individual has ToM capacities, several tests have been developed in psychology. One of the most famous is the Sally and Anne test (fin 3.2). This test allows to check the capacity of someone to attribute a false-belief to another person and have been reused in robotics to validate perspective taking systems.

3.2.2 Robotics background

One of the pioneer work in robotics Theory of Mind is [Scassellati 2002]. Scassellati presents two models from social sciences (Leslie [Leslie 1984] et Baron-Cohen [Baron-Cohen 1997]) and proposes a model on how to implement ToM in robotics. However, the implementation of this model did not go further than perception level.

Then, several works have been done in order to endow robots with perspective taking abilities. Using ACT-R architecture [Anderson 2004], the team of Hiatt and Trafton models mechanisms used during the Sally and Anne test and constructs a model that learns to deal with false belief to pass this test [Hiatt 2010]. They extend this work to second-order in [Hiatt 2015] and to spatial reasoning in [Hiatt 2004]. The Sally and Anne test has also been passed in [Milliez 2014] where the robot constructs a semantic representation of the world from its partners point of view. In [Berlin 2006], authors present a way to record different beliefs of other agents and so to have a memory of perspective taking. Finally, [Johnson 2005b] presents a system which computes perspective taking perspective taking based on forward and inverse visual models.

Perspective taking abilities have been used in robotics for several purposes. It has been used in [Hiatt 2011] to deal with uncertainty in humans behavior and in [Ros 2010] to solve ambiguous references to an object. One important application of perspective taking is action recognition. [Johnson 2005a] takes the visual point of view of humans to improve action recognition, Dynamic Bayesian Networks (DBN) are used in [Baker 2014] or inverse reinforcement learning in [Nagai 2015]. The human perspective is also used in [Breazeal 2006] to learn a task from a situation that can be ambiguous from the robot point of view and in [Gray 2014] to choose actions with the adequate effects in order to manipulate humans mental models. Finally, [Görür 2017] uses perspective taking to infer humans intention and adapt robot decision.

Concerning Shared Plans, they can be found by using perspective taking in order to add communication actions [Guitton 2012]. Then, the human perspective is used to share the plan with a level of details depending of human knowledge [Milliez 2016]. However, there is no works for now concerning the management of Shared Plans execution taking into account the human point of view.

3.3 Estimating Humans Mental States

- reminder representation, estimate operator
- world state: observable and non-observable

- goal
- Shared Plan, Actions

3.4 Mental States for Shared Plans execution

- solve db operator
- weak achievement goal
- human should act
- preventing mistakes
- robot action signalling
- inaction and uncertainty

3.5 Results

3.5.1 Tasks

- clean the table
- inventory

3.5.2 Interesting scenario

Clean the table ICSR paper

3.5.3 Experiment and results

- in simulation
- human follows hatp plan, human leaves
- criteria
- results
- discussion

Bibliography

[Anderson 2004] John R Anderson, Daniel Bothell, Michael D Byrne, Scott Douglass, Christian Lebiere and Yulin Qin. *An integrated theory of the mind.* Psychological review, vol. 111, no. 4, page 1036, 2004. (Cited in page 4.)

[Baker 2014] Chris L Baker and Joshua B Tenenbaum. *Modeling human plan recognition using bayesian theory of mind.* Plan, activity, and intent recognition: Theory and practice, pages 177–204, 2014. (Cited in page 4.)

[Baron-Cohen 1985] Simon Baron-Cohen, Alan M Leslie and Uta Frith. *Does the autistic child have a "theory of mind"?* Cognition, vol. 21, no. 1, pages 37–46, 1985. (Cited in page 2.)

[Baron-Cohen 1997] Simon Baron-Cohen. Mindblindness: An essay on autism and theory of mind. MIT press, 1997. (Cited in page 4.)

[Berlin 2006] Matt Berlin, Jesse Gray, Andrea Lockerd Thomaz and Cynthia Breazeal. *Perspective taking: An organizing principle for learning in human-robot interaction.* In AAAI, volume 2, pages 1444–1450, 2006. (Cited in page 4.)

[Breazeal 2006] Cynthia Breazeal, Matt Berlin, Andrew Brooks, Jesse Gray and Andrea L Thomaz. *Using perspective taking to learn from ambiguous demonstrations.* Robotics and autonomous systems, vol. 54, no. 5, pages 385–393, 2006. (Cited in page 4.)

[De Weerd 2014] Harmen De Weerd, Rineke Verbrugge and Bart Verheij. *Theory of Mind in the Mod Game: An Agent-Based Model of Strategic Reasoning.* In ECSI, pages 128–136, 2014. (Cited in page 2.)

[Devin 2016] Sandra Devin and Rachid Alami. *An implemented theory of mind to improve human-robot shared plans execution.* In Human-Robot Interaction (HRI), 2016 11th ACM/IEEE International Conference on, pages 319–326. IEEE, 2016. (Cited in page 1.)

[Flavell 1977] John H Flavell. *The development of knowledge about visual perception.* In Nebraska symposium on motivation. University of Nebraska Press, 1977. (Cited in page 3.)

[Flavell 1992] John H Flavell, H Beilin and P Pufall. *Perspectives on perspective-taking.* Piaget's theory: Prospects and possibilities, pages 107–139, 1992. (Cited in page 3.)

[Frick 2014] Andrea Frick, Wenke Möhring and Nora S Newcombe. *Picturing perspectives: development of perspective-taking abilities in 4-to 8-year-olds.* Frontiers in psychology, vol. 5, page 386, 2014. (Cited in page 3.)

[Görür 2017] O Can Görür, Benjamin Rosman, Guy Hoffman and Şahin Albayrak. *Toward Integrating Theory of Mind into Adaptive Decision-Making of Social Robots to Understand Human Intention.* 2017. (Cited in page 4.)

[Gray 2014] Jesse Gray and Cynthia Breazeal. *Manipulating mental states through physical action.* International Journal of Social Robotics, vol. 6, no. 3, pages 315–327, 2014. (Cited in page 4.)

[Guitton 2012] Julien Guitton, Matthieu Warnier and Rachid Alami. *Belief Management for HRI Planning.* European Conf. on Artificial Intelligence, 2012. (Cited in page 4.)

[Hiatt 2004] Laura M Hiatt, J Gregory Trafton, Anthony M Harrison and Alan C Schultz. *A Cognitive Model for Spatial Perspective taking.* In ICCM, pages 354–355, 2004. (Cited in page 4.)

[Hiatt 2010] Laura M Hiatt and J Gregory Trafton. *A cognitive model of theory of mind.* In Proceedings of the 10th International Conference on Cognitive Modeling, pages 91–96. Citeseer, 2010. (Cited in page 4.)

[Hiatt 2011] Laura M Hiatt, Anthony M Harrison and J Gregory Trafton. *Accommodating human variability in human-robot teams through theory of mind.* In IJCAI Proceedings-International Joint Conference on Artificial Intelligence, volume 22, page 2066, 2011. (Cited in page 4.)

[Hiatt 2015] Laura M Hiatt and J Gregory Trafton. *Understanding second-order theory of mind.* In Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts, pages 167–168. ACM, 2015. (Cited in page 4.)

[Johnson 2005a] Matthew Johnson and Yiannis Demiris. *Perceptual perspective taking and action recognition.* International Journal of Advanced Robotic Systems, vol. 2, no. 4, page 32, 2005. (Cited in page 4.)

[Johnson 2005b] MR Johnson and YK Demiris. *Perspective taking through simulation.* 2005. (Cited in page 4.)

[Leslie 1984] Alan M Leslie. *Spatiotemporal continuity and the perception of causality in infants.* Perception, vol. 13, no. 3, pages 287–305, 1984. (Cited in page 4.)

[Milliez 2014] Grégoire Milliez, Matthieu Warnier, Aurélie Clodic and Rachid Alami. *A framework for endowing an interactive robot with reasoning capabilities about perspective-taking and belief management.* In Robot and Human Interactive Communication, 2014 RO-MAN: The 23rd IEEE International Symposium on, pages 1103–1109. IEEE, 2014. (Cited in page 4.)

[Milliez 2016] Grégoire Milliez, Raphaël Lallement, Michelangelo Fiore and Rachid Alami. *Using human knowledge awareness to adapt collaborative plan generation, explanation and monitoring.* In HRI ACM/IEEE, 2016. (Cited in page 4.)

[Nagai 2015] Takayuki Nagai, Kasumi Abe, Tomoaki Nakamura, Natsuki Oka and Takashi Omori. *Probabilistic modeling of mental models of others.* In Robot and Human Interactive Communication (RO-MAN), 2015 24th IEEE International Symposium on, pages 89–94. IEEE, 2015. (Cited in page 4.)

[Premack 1978] David Premack and Guy Woodruff. *Does the chimpanzee have a theory of mind?* Behavioral and brain sciences, vol. 1, no. 04, pages 515–526, 1978. (Cited in page 2.)

[Ros 2010] Raquel Ros, Emrah Akin Sisbot, Rachid Alami, Jasmin Steinwender, Katharina Hamann and Felix Warneken. *Solving ambiguities with perspective taking.* In Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction, pages 181–182. IEEE Press, 2010. (Cited in page 4.)

[Scassellati 2002] Brian Scassellati. *Theory of mind for a humanoid robot.* Autonomous Robots, vol. 12, no. 1, pages 13–24, 2002. (Cited in page 4.)

[Tversky 1999] Barbara Tversky, Paul Lee and Scott Mainwaring. *Why do speakers mix perspectives?* Spatial cognition and computation, vol. 1, no. 4, pages 399–412, 1999. (Cited in page 3.)

[Verbrugge 2008] Rineke Verbrugge and Lisette Mol. *Learning to apply theory of mind.* Journal of Logic, Language and Information, vol. 17, no. 4, pages 489–511, 2008. (Cited in page 2.)