AToM: An Analogical Theory of Mind

Irina Rabkina[1]

[1] Northwestern University, Evanston, IL 60208, USA
irabkina@u.northwestern.edu

Abstract. Theory of Mind (ToM) is what gives adults the ability to predict other people's beliefs, desires, and related actions, and has been heavily studied in psychology. When ToM has not yet developed, as in young children, social interaction is difficult. Cognitive systems that interact with people on a regular basis would benefit from having a ToM. In this research summary, I propose a computational model of ToM, Analogical Theory of Mind (AToM), based on Bach's [2012, 2014] theoretical Structure-Mapping model of ToM. Completed work demonstrates how ToM might be learned under this model. Future steps include a full implementation and test of AToM.

Keywords: Analogy, Structure Mapping, Theory of Mind

1 Introduction

Humans are inherently social creatures. In fact, it has been suggested that our need for social interaction is responsible for our large brains and incredible language abilities [e.g. Reader and Laland, 2002]. If artificial intelligence systems are to be integrated into our society, then they must share the social capabilities available to us.

Theory of Mind (ToM) is one example of a capability necessary for social interaction. ToM, sometimes referred to as mind reading, is the ability to predict others' desires, beliefs, and other mental states even when they may be different from our own. While some evidence of ToM exists in other highly social animals, such as dolphins and apes [e.g. Krupenye et al. 2016], the extent to which we use and rely on ToM seems to be uniquely human.

Several theories of how ToM is developed and used by humans exist. The philosopher Theodore Bach [2011, 2014] proposed one such theory, based in the Structure-Mapping Theory of analogy [SMT, Gentner, 1983]. This research summary describes a computational cognitive model of ToM, Analogical Theory of Mind (AToM), which is based on Bach's theory. Previous work, which shows how processes which play a role in ToM development can be used to train AToM, is presented. Finally, future directions are discussed.

2 Analogical Theory of Mind (AToM)

AToM is based on the Structure-Mapping Theory of ToM proposed by Bach [2011, 2014]. It is built on top of the Structure-Mapping Engine [SME, Forbus et al. 2016], a computational model of SMT [Gentner, 1983]; the SAGE model of analogical generalization [McLure et al. 2010]; and the MAC/FAC model of analogical retrieval [Forbus et al. 1995]. AToM assumes a long term memory (LTM) of predicate calculus cases that can be retrieved via MAC/FAC. These cases represent memories of life experiences.

When a situation which requires ToM reasoning is encountered, AToM retrieves a relevant case from LTM using MAC/FAC (see Fig. 1). If the retrieved case is a generalized schema, it is applied via analogical mapping as if it were a rule. If the retrieved case is a single event, an interim generalization is created in working memory [Kandaswamy et al. 2014]. While standard interim generalizations are created via SAGE, a slightly different process is involved for AToM's generalizations. Candidate inferences from the retrieved case are projected onto the probe case and, where necessary, portions of the probe case are re-represented. This interim generalization is used for ToM reasoning. AToM then asks for feedback in natural language [using EA-NLU, Tomai and Forbus, 2009]. This is analogous to a person receiving feedback on their reasoning by interacting with others. If the reasoning was correct, AToM uses SAGE to generalize the original probe with the retrieved case, and stores the new generalized case in LTM. Otherwise, it uses MAC/FAC to find a better match (again, given the feedback) and generalizes with the new match. In this way, schemas become more and more generalized, and ToM abilities continue to improve.

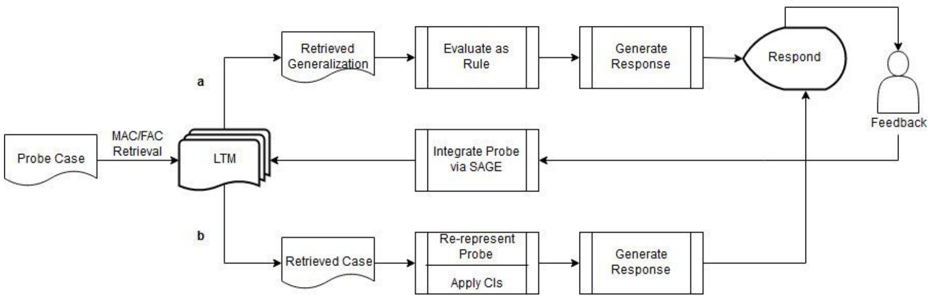

Fig. 1. Process diagram of AToM. Path **a** shows the process when a generalization is retrieved. Path **b** shows the process when a single example case is retrieved.

While AToM is based on Bach's Structure-mapping Theory of ToM [2011, 2014], it differs from the theory in several crucial ways. I will discuss the two biggest differences here. The first major change is to what Bach refers to as the *base representation*, or the case from which reasoning occurs. He suggests that the base representation is formed by re-representing the probe case from the third person into the first person, and adding facts that represent mental state, which are generated by a separate decision-making system. While the interim generalization generated by AToM is analogous to Bach's base representation, the re-representation process is based on a specific retrieved case. The mental state facts, then, are also projected as candidate inferences from the retrieved case, rather than being generated by a separate system.

Another important difference between AToM and Bach's theory lies in the integration of the probe case to LTM. Bach posits that schemas for ToM reasoning are abstracted from simulations. This abstraction happens during construction of the base representation and the comparison between it and the original probe. In AToM, the schemas are instead formed by generalizing the original probe with the retrieved case. In this way AToM builds up its LTM directly from its experiences.

3 Progress to Date

I have completed two computational models of processes involved in ToM learning. These models were used to simulate psychological studies and show results consistent with human data. These results suggest that AToM is a plausible model of ToM reasoning.

3.1 Pretense

Pretend play is ubiquitous throughout childhood. Psychologists believe that it plays a large role in social development in general, and ToM development in particular [Weisberg, 2015]. The mechanisms by which pretense aids with development, however, is an open question. We [Rabkina and Forbus, in prep] suggest that pretense is an analogical process which drives the development of some aspects of analogical reasoning. Because AToM, per Bach, argues that ToM is also analogical, it follows that development of analogical processes will aid ToM development.

Our model of pretense suggests that pretend play relies heavily on analysis of candidate inferences. In the model, when a pretend scenario is encountered, a schema of its real-life equivalent is retrieved. The two are compared via SME, and candidate inferences are projected from the schema to the pretend scenario. Pretend play is successful when the child is able to accept the proper candidate inferences and transform the pretend scenario accordingly. The model successfully replicates the patterns of behavior, including success and failure in pretense, observed in two psychological studies [Fein, 1975; Onishi et al. 2007].

The process by which interim generalizations are formed in AToM is very similar to how they are formed in the pretense model: candidate inferences from the retrieved case must be evaluated and applied to the probe. Thus, it is reasonable that practicing this skill via pretense would improve ToM abilities.

3.2 ToM Training Study

While the pretend play study suggests one mechanism by which ToM might be learned in the wild, psychologists have been able to teach children some aspects of ToM in short intervention sessions. For example, Hoyos et al. [2015] used the repetition break paradigm, described below, to teach children false belief tasks.

In this study, children heard three vignettes. These vignettes were all of the same form: the child is presented with a container (e.g. a crayon box) and asked what they believe is inside. The contents of the container are then revealed. In two of the vignettes, the contents of the box are as expected (e.g. crayons in the crayon box); in the third, they are surprising (e.g. grass in the crayon box).This format is referred to as repetition-break. After the reveal, a new character is introduced, and the child is asked what the character believes is inside the box. In the case where the contents of the box are surprising, the child is expected to answer with the false belief (e.g. the character thinks there are crayons in the box, even though there is actually grass).

From just hearing the three vignettes, children improved significantly on several false belief tasks. Importantly, children who heard vignettes that were highly alignable, that is had high structural similarity, outperformed children who heard vignettes that did not align [Hoyos et al. 2015].

While this alone provides evidence for the role of structure-mapping in ToM development, AToM provides a mechanism by which it may actually happen. In fact, a version of AToM [Rabkina et al. 2017] accurately modeled this task. The model included only a simplified version of the learning steps of AToM: retrieval and integration, along with a reasoning step. Using a simplified-English version of the vignettes and tests used by Hoyos et al. [2015], it replicated the pattern of learning achieved by the children in the study. That is, the model learned false belief tasks from both sets of vignettes, but learned more of them from the vignettes which were highly alignable. Furthermore, the model provided several predictions about ToM in humans.

4 Future Directions

The experiments described above provide evidence that AToM is a plausible mechanism for ToM. However, ToM covers a broad range of phenomena, and a complete model of ToM should be able to model human performance on a variety of tasks. I am currently in the process of identifying additional tasks for testing AToM that would provide a base of evidence that AToM can explain the breadth of ToM reasoning and development in both children and adults.

There are also several areas in which AToM can be improved as a model. For example, the repetition-break study [Hoyos et al. 2015] and our model of it [Rabkina et al. 2017] suggest that surprise plays a role in learning ToM. Incorporating a model of surprise into AToM is a future goal. Furthermore, candidate inference evaluation is important to both AToM and our pretense model [Rabkina and Forbus, in prep]. Developing a cognitively plausible mechanism for these evaluations is also future work.

5 Acknowledgements

This research was supported by the Socio-Cognitive Architectures for Adaptable Autonomous Systems Program of the Office of Naval Research, N00014-13-1-0470.

References

[Bach, 2011] Theodore Bach. Structure-mapping: Directions from simulation to theory. *Philosophical Psychology*, 24(1): 23-51. 2011.

[Bach, 2014] Theodore Bach. A Unified Account of General Learning Mechanisms and Theory-of-Mind Development. *Mind & Language*, 29(3): 351-381. 2014.

[Gentner, 1983] Dedre Gentner. Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7, 155–170. 1983.

[Fein, 1975] Greta G. Fein. A transformational analysis of pretending. *Developmental Psychology* 11.3:291. 1975.

[Forbus et al. 1995] Kenneth D. Forbus, Dedre Gentner, and Keith Law. MAC/FAC: A Model of Similarity-based Retrieval. *Cognitive Science*, 19:141-205. 1995.

[Forbus et al. 2016] Kenneth D. Forbus, Ronald W. Ferguson, Andrew Lovett, and Dedre Gentner. Extending SME to handle large-scale cognitive modeling. *Cognitive Science*, DOI: 10.1111/cogs.12377, pp 1-50. 2016.

[Hoyos et al. 2015] Christian Hoyos, William S. Horton, and Dedre Gentner. Analogical comparison aids false belief understanding in preschoolers. In *Proceedings of the Cognitive Science Society*. 2015.

[Kandaswamy et al. 2014] Subu Kandaswamy, Kenneth Forbus, and Dedre Gentner. Modeling Learning via Progressive Alignment using Interim Generalizations. *Proceedings of the Cognitive Science Society*. 2014.

[Krupenye et al. 2016] Christopher Krupenye, Fumihiro Kano, Satoshi Hirata, Josep Call, and Michael Tomasello. Great apes anticipate that other individuals will act according to false beliefs. Science 354, no. 6308: 110-114. 2016.

[McLure et al. 2010] Matthew McLure, Scott Friedman, and Kenneth Forbus. Learning concepts from sketches via analogical generalization and near-misses. *Proceedings of the 32nd Annual Conference of the Cognitive Science Society (CogSci).* Portland, OR. 2010.

[Onishi et al. 2007] Kristine H. Onishi, Renée Baillargeon, and Alan M. Leslie. 15-month-old infants detect violations in pretend scenarios. *Acta Psychologica* 124.1: 106-128. 200.7.

[Rabkina and Forbus, in prep] Irina Rabkina and Kenneth D. Forbus. An Analogical Model of Pretense. In prep.

[Rabkina et al. 2017] Irina Rabkina, Clifton McFate, Kenneth D. Forbus, and Christian Hoyos. Towards a Computational Analogical Theory of Mind. *Proceedings of the 39th Annual Meeting of the Cognitive Science Society*. London, England. 2017.

[Reader and Laland, 2002] Simon M. Reader and Kevin N. Laland. Social intelligence, innovation, and enhanced brain size in primates. *Proceedings of the National Academy of Sciences*. 99.7:4436-4441. 2002.

[Tomai and Forbus, 2009] Emmett Tomai and Kenneth Forbus. EA NLU: Practical Language Understanding for Cognitive Modeling. *Proceedings of the 22nd International Florida Artificial Intelligence Research Society Conference*. Sanibel Island, Florida. 2009.

[Weisberg, 2015] Deena S. Weisberg. Pretend play. *Wiley Interdisciplinary Reviews: Cognitive Science* 6.3: 249-261. 2015.