Michelson, J., Sanyal, D., Ainooson, J., Yang, Y., & Kunda, M. (2021). "Social cognition paradigms ex machinas."
Proceedings of the AAAI Fall Symposium on Computational Theory of Mind for Human-Machine Teams.

Social Cognition Paradigms *ex Machinas*

Joel Michelson, Deepayan Sanyal, James Ainooson, Yuan Yang, Maithilee Kunda
Vanderbilt University Department of Computer Science
Nashville TN, USA
{joel.p.michelson, deepayan.sanyal, james.ainooson, yuan.yang, mkunda}@vanderbilt.edu

Abstract

In this paper, we discuss the creative design of task paradigms invented to study the social and theory-of-mind skills utilized by humans and animals as well as the potential applications of these paradigms in artificial intelligence research. We first present a detailed review of 21 tasks from the cognitive literature. Next, we provide a description of our process for translating these tasks into AI-suitable environments, along with a detailed example using the competitive feeding task paradigm. Finally, we discuss how a battery of these tasks would be useful for building, training, and evaluating future artificial models of social intelligence.

1 Introduction

In the late 1980s, ecologist James Gould performed a series of experiments to better understand honeybees' navigational abilities (Gould, Gould et al. 1988), but these experiments also ended up posing fascinating questions about bees' social reasoning abilities. In the first experiment, Gould captured several foraging bees and carried them to a boat with flowers in the middle of a lake. These foragers then returned to their hive and indicated the flowers' location by dancing, but failed to inspire any recruits to fly in that direction. Later, foragers were shown a new location of flowers, in the same boat but now moved close to the opposite shoreline. This time, their recruitment was successful.

Why did the bee recruits decide to "believe" the foragers the second time, but not the first? While this example has a lot to do with mental maps, navigation, and memory, it also involves bees reasoning about the beliefs of other bees in relation to their own in a pretty sophisticated way.

"But wait!" the skeptical reader exclaims. "What if this wasn't about beliefs, but something more basic? What if the foragers simply smelled like lakewater, or gave off some other basic cue, and recruits merely avoided following that smell/cue?" This very question was asked by experimenters, and they conducted a followup experiment in which the entire hive was transported to a field with flower stations at analogous relative positions.[1]

[1]How gullible are bees to this kind of house-swap? As the paper amusingly notes, "Bees readily accept a new site as the home locale if the most prominent landmarks are roughly equivalent, and substitutions of grass for water and vice versa are not the most outrageous exchanges bees will tolerate" (Gould 1990, p.100).

Now, while foragers enjoyed good food at both field sites and danced roughly equally for both, recruits (presumably who had "not yet been out to note that the lake has mysteriously dried up overnight" (Gould 1990, p.100)) still preferred the shoreline-analogous location.

This example illustrates two important points motivating our work. First: social and theory-of-mind (ToM) abilities (i.e., reasoning about the mental states of the self and others (Bird and Viding 2014)) are essential for intelligence in a wide variety of contexts faced by a wide variety of species.

Second: studying these abilities requires *extremely* careful task designs. It can be easy to design tasks that look like ToM tasks but that can be solved using simple perceptual cues. The comparative cognition (nonhuman animal) literature is rife with debates about ToM tasks and what they purportedly measure versus what they actually measure (e.g. Penn and Povinelli 2007; Penn, Holyoak, and Povinelli 2008).

In artificial intelligence (AI), social and ToM skills are receiving increasing attention due to their essential role in settings involving cooperation and competition, including in multi-agent settings as well as for human-machine teams. And, while AI research has begun to pull inspiration from the rich literature on biological social cognition, we propose that there is much to be learned on both sides by bridging research across cognitive and computational approaches.

In particular, AI research is often driven forward by having concrete challenge tasks in a specific domain (e.g., chess, Go, ImageNet ILSVRC). We observe that, in the current AI literature, social and ToM tasks are often studied in isolation, with different AI systems built to tackle one or a small set of related tasks, like the ToMNet system (Rabinowitz et al. 2018). On the other hand, collections of tasks in other areas of AI have served to catalyze interesting lines of ensuing research, like ALE (Bellemare et al. 2013) and the Animal-AI Testbed (Crosby et al. 2020) for various single-agent scenarios, or Arena (Song et al. 2020) and MARLÖ (Perez-Liebana et al. 2019) for multi-agent scenarios.

In this paper, we present our initial steps towards creating a new ToM-Testbed for AI research, inspired by the human and animal ToM literature. We envision the ToM-Testbed as containing a large suite of ToM tasks implemented in a uni-

form gridworld environment like those commonly used in multi-agent research. While our ToM-Testbed is still under construction, the contributions of this paper include:

- A detailed review of 21 tasks (with multiple variants per task) from the human and non-human animal ToM literature that are candidates for inclusion in our ToM-Testbed.
- A description of our process of translating tasks designed for humans and animals to gridworld environments, with a detailed example using the competitive feeding task.
- A discussion of specific ways in which the ToM-Testbed could be leveraged in computational experiments to study ToM abilities, learning, transfer, and more.

Our eventual goal is to be able to answer questions about ToM tasks and models that were previously inaccessible. For example, which tasks are readily solvable by off-the-shelf machine learning models? Does success at one set of tasks by an artificially intelligent model seem to imply success at another? If so, do models' performances replicate findings in human childhood development and the rest of the animal kingdom?

2 Background: Populations of Interest

Before diving into our review of specific tasks, we first present a high-level overview of where significant pockets of research on social and ToM reasoning are to be found: 1) typical child development; 2) atypical development (e.g., autism); 3) non-human animals; and 4) artificial agents.

2.1 Typical child development

Tasks involving ToM have been vital for understanding child development. In 1983, (Wimmer and Perner 1983) designed what became known as the Sally Anne task, a test of the ability to attribute false beliefs to other people, that can be given to children. Although false belief (FB) tests are popular and useful predictors of multiple aspects of social skills' development, other aspects of social cognition are examined independently. In Beaudoin's et al. review of developmental ToM measures, skills are divided into seven categories, each referring to the inference of and reasoning about others' emotions, desires, intentions, percepts, knowledge, beliefs, and non-literal communication (Beaudoin et al. 2020).

Numerous theories about the ontogenetic development of ToM have been proposed. Nativist theories maintain that children's learning is largely independent of environment, that evolution essentially hardwires social skills into their brains (Leslie 1994). 'Theory theory' focuses on the idea of a 'conceptual revolution' in which children learn to formulate scientific theories (Gopnik and Wellman 1994). Simulation theory is a view that highlights the importance of pretend play in children, under the assumption that the capacity for pretence is the mechanism that allows for ToM (Harris 1992). The executive function hypothesis overlaps with simulation theory, and focuses on the importance of ToM as a component of flexible planning and goal-directed behavior (Carlson and Moses 2001; Hughes 1998; Russell 1997).

ToM tests for human children frequently involve storytelling. Many types of measurement are used, such as verbal question answering, making choices of pictures or objects, making actions within a setting, or eye-tracking.

2.2 Atypical child development

ToM is also a cornerstone of studying various trajectories of atypical child development. For example, ToM has long been shown to develop and present in atypical ways in autism. Many (though not all) children on the autism spectrum show difficulties in false belief tasks (Baron-Cohen et al. 1985) and other areas of ToM (Happé and Frith 1995), though the sources and full effects of these differences are still not well understood.

As another example, deaf children who are born to hearing parents have been observed to show ToM deficits similar to those shown by children with autism, but similar deficits were *not* seen in deaf children born to deaf parents, who presumably had the benefits of rich parent-based language exposure from early infancy (Peterson and Siegal 1999).

Research on ToM in atypical development can not only provide clues as to ingredients and dependencies that support ToM in typical development, but also can highlight how intelligent agents can develop compensatory strategies in the absence of some of these ingredients or dependencies.

Much of this research has also yielded debates about specific ToM tasks, their design, and what they measure.

2.3 Non-human Animals

There is longstanding and vigorous debate about the higher-level social and ToM reasoning capabilities of nonhuman animals, usually studied as a function of different species. Even nonhuman primates like chimpanzees show only limited ToM abilities relative to what even young typically developing human children can do. Even so, the gulfs in social and ToM abilities among different nonhuman animal species are vast, with nonhuman primates and a handful of other species (corvids, i.e., jays and crows, dolphins, domesticated dogs, etc.) showing quite sophisticated abilities relative to other animal species.

In animals, ToM tasks are even more specific and difficult to interpret than they are in humans. As such, their design is generally incredibly strict, with researchers inventing increasingly ingenious controls to avoid null and alternate hypotheses (Penn, Holyoak, and Povinelli 2008), like the Clever Hans effect.[2] Furthermore, at times, animals produce puzzling results in which they succeed at one puzzle but fail at something that seems (to us) to be much simpler.

Animals' tests are restricted in form, as we cannot verbally explain rules, stories, etc. to the subjects. Only certain kinds of responses can be measured for the same reasons.

2.4 Artificial Agents

Recently, social skills have been the focus of much attention by AI researchers, so some of the ideas from human and animal tests have been adapted for machine use.

[2]Clever Hans was a horse who was seemingly able to solve difficult problems of arithmetic, but later found to be reliant upon his trainer's involuntary body language cues.

Rabinowitz et al. developed ToMNet, a supervised learning system designed to predict the actions, beliefs, preferences, and percepts of agents moving around a gridworld environment (Rabinowitz et al. 2018). In their study of ToMNet's abilities, they design computational variants of the Sally Anne test, a paradigm designed to test for false-belief attribution in humans and animals (Wimmer and Perner 1983).

Hernandez-leal et al. provide a thorough review of research in the field of multi-agent deep reinforcement learning (Hernandez-Leal, Kartal, and Taylor 2019), including a discussion of algorithms that perform some version of theory-of-mind reasoning in the context of adversarial games.

From a slightly different angle, the work described in this paper is heavily inspired by the Animal AI Testbed (Crosby et al. 2020), which is a suite of first-person navigation challenges implemented in a three-dimensional environment modeled after studies of animal intelligence, and in which many of the tasks have specifically been deployed with animals in other research. Although it does not currently include tests of social reasoning, its repertoire of 900 sub-tasks are all variants of one of 12 animal cognition paradigms or 16 classes of environment used to teach fundamental skills like basic exploration.

3 Review of ToM Tasks

We have conducted a detailed review of numerous widely-used ToM tasks from the cognitive literature, mostly drawing from studies of non-human animals (though several of these tasks have been used in human studies as well).

We present this review as a resource for AI researchers to get a sense of the kinds of tasks used in the cognitive literature and to understand which of these tasks are more or less difficult for our various animal brethren. These tasks are also highly informative from the perspective of task designs, i.e., in identifying when and how tasks might be solved using alternative (e.g., non-ToM) methods, and how task variants can be combined to pinpoint the extent to which an individual intelligent agent truly demonstrates certain capabilities.

We present an overview of studies we reviewed in Table 1, and we also present narrative descriptions of all tasks in the Appendix. We have classified these tasks into the kinds of ToM reasoning they entail: preferences, perception, intent, knowledge, beliefs, deception, and other.

4 Translating real-world tasks $\rightarrow$ AI tasks

We first describe some desiderata that we are using to guide our selection of real-world (referring here to tasks designed for humans and animals) ToM tasks and our implementation of them for AI frameworks. Then, we describe the process of implementation for one example task.

4.1 Desiderata

Just as there are almost always alternate explanations for observed animal behaviors, performing well at a given example of these tasks does not necessarily indicate success at any specific skill. We do not expect to be able to build a system for testing models with outputs as convenient as "has ToM" or even "can infer preferences of another agent". Like the paradigms' use for studying animals, any particular test is subject to interpretation and criticism. For this reason, our aim in translating these problems is to do so with a purely task-based perspective.

That said, a task for an artificially intelligent agent should be as informative and meaningful as possible. For each task variant, our goal is to create an artificial environment as close as possible to the original design. In human and animal tests, intelligent controls are implemented to help ensure the results report on their intended subject. By providing multiple variations of each task, with every reasonable control and dependent variable available, we hope to provide the most perspectives into models' reasoning processes as possible. In fact, the computational nature of the tasks and models offers novel methods of observation that are infeasible in the real world. Agents' belief states may be quantified objectively (Rabinowitz et al. 2018), their perceptions reported with perfect accuracy, and their ontological training process can be understood deeply, a stark contrast with any test performed on humans and (especially) animals.

Due to tasks' dependence on precursor knowledge, we attempt to retain commonalities between them when sensible. Tasks should involve the simplest percepts and controls possible so that a completely naïve model does not need to overcome too many unrelated learning hurdles to succeed. In Table 2, we hypothesize commonalities of selected skills that may be necessary for the completion of various tasks.

Task	Memory	Desires$_A$	Sees$_A$	FB$_A$
YummyYucky		x		
Two-action	x	x		
Knower-guesser	x		x	
Sally Anne	x	x	x	x

Table 2: Hypothetical commonalities between selected tasks. While Memory refers to a requirement for the subject to have memory, Desires$_A$, Sees$_A$, and FB$_A$ refer to the subject's inference of (i.e. attribution of) desires, vision, and false beliefs (FBs) in other agents present in the task's setting.

One of the benefits of using a visual task environment is the possibility of these tests (in their reimagined formulation) being run in human or animal studies for comparison. The ability to run these same benchmarks with human subjects of similar populations to previous studies will provide insight into both the adequacy of our specific implementations as well as additional data reproducing findings on the real-world versions of these tasks.

Finally, the testbed should be able to run quickly and in parallel for optimal reinforcement learning training. The battery of tasks on the testbed should be extensible, to accommodate the frequent development of new tasks by comparative and developmental psychologists. The system should be open source to enable rapid advances in social capabilities of AI.

Task	Reference	Variant	Species	n	Training	Control	Test
Preference							
Yummy-yucky	(Repacholi and Gopnik 1997)		14, 18 month-old children	159	2	1	1
Multiple desires	(Bennett and Galpert 1993)	Successive	5, 8 year-old children	20; 15	0	0	2
		Simultaneous	Children of various ages	75	0	0	4
		Study 3	5, 7 year-old children	25; 25	0	0	2
Perception							
Picture identification	(Masangkay et al. 1974)	Picture task	2-3;6 year-old children	16; 9	2	2	4
Appearance-reality	(Flavell et al. 1986)	Testing	Children of various ages	>16	0	3	5
		Teaching	3 year-olds	16	0	0	1
Intent							
Two-action	(Akins and Zentall 1996)		Japanese quail	12	1	1	2
Distinguishing Intentionality	(Call and Tomasello 1998)		2 and 3 year-old children	8; 8	1	0	2
			Chimpanzees, orangutans	5; 5	1	0	2
Rational imitation	(Meltzoff 1988)	Head-touching	14-month-old children	36	0	0	2
	(Gergely, Bekkering, and Király 2002)		Preverbal infants	27	0	1	2
Accidental Transgression	(Killen et al. 2011)	Experiment 1	3-8-year-old children	162	0	2	2
		Experiment 2	3-8-year-old children	46	0	2	1
Knowledge							
Competitive Feeding	(Hare, Call, and Tomasello 2001)	Did	Chimpanzees	9	0	1	2
		Who	Chimpanzees	8	0	1	1
		Which	Chimpanzees	9	0	1	2
Knower-Guesser	(Udell, Dorey, and Wynne 2011)	Begging	Wolves and Dogs	60	1	1	4
		Bucket training	Wolves and Dogs	8; 12	0	0	2
Goggles	(Karg et al. 2015)	Gaze following	Chimpanzees	25	1	1	2
		Competitive	Chimpanzees	19	1	1	3
See-know task	(Pillow 1989)	Experiment 1	3, 4 year-olds	16; 16	0	2	2
		Experiment 2	3 year-olds	12	0	4	4
Belief							
Sally Anne	(Wimmer and Perner 1983)		Children of various ages	36	0	2	2
		Exploration	Kindergarten children	92	0	2	3
	(Baron-Cohen et al. 1985)		Human children	61	0	3	1
	(Krupenye et al. 2016)	Experiment 1	Great apes	43	1	0	2
		Experiment 2	Great apes	44	1	0	2
	(Southgate, Senju, and Csibra 2007)	FB 1	2 year-old children	20	2	1	1
		FB 2	2 year-old children	20	2	1	1
Ignorance vs. FB	(Hogrefe, Wimmer, and Perner 1986)	Experiment 1	3,4 and 5 year-old children	20; 20; 20	0	2	2
		Experiment 2	3,4 and 5 year-old children	24; 24; 24	0	2	2
		Experiment 3	3;6 year-old children	22	0	2	2
		Experiment 4	3 and 4 year-old children	18; 18	0	2	1
		Experiment 5	3-4 year olds	36	0	0	2
		Experiment 6	4, 5 and 6 year-olds	12; 12; 12	0	2	3
Inhibitory FB	(Leslie and Polizzi 1998)	Negative desire	4 year-olds	16	0	3	2
		Opposite behavior	4 year-olds	16	0	3	2
Deception							
Penny/marble Hiding	(Gratch 1964)		2-7 year-old children	106	1	0	1
Box-locking Sabotage	(Sodian and Frith 1992)	One box	Human children	87	1	1	4
		Two boxes	Human chilcren	88	1	1	4
Back-and-forth Foraging	(Schmelz, Call, and Tomasello 2011)		Chimpanzees	12	2	1	2
Unseen competitors	(Bugnyar, Reber, and Buckner 2016)		Ravens	10	0	2	1
Other							
Mirror self-recognition	(Gallup 1970)	Mirror exposure	Chimpanzees	4	1	0	1
		Marking	Chimpanzees, monkeys	4; 4	1	1	1
Role-reversal	(Povinelli, Nelson, and Boysen 1992)		Chimpanzees	4	1	0	2

Table 1: Social Cognition Paradigms. *Training*, *control*, and *test* refer to the numbers of individually defined pretraining, control, and test conditions. We consider measured exposure to most task elements to be training, but certain elements assumed to be understood (such as familiarization with the Likert scale) are not included. In several cases, pretraining involves the prior tasks in the experiment; these do not count as additional training tasks. Likewise, if a training task is used as a control condition, we only count it once as a pretraining task.

4.2 Passive observers and active participants

While some tasks involve observing and then answering questions, plenty require an agent's participation in the given setting. This requirement is not surprising, given that many social skills exhibit themselves through cooperation and competition with peers. The distinction between observational and interaction-based tasks is not always clear, and many tests may be imagined in either light.

Our selection of a two-dimensional environment allows for multiple kinds of input types (e.g. egocentric and allocentric worldviews, three-channel and 'rich' image formats), and should be amenable to both supervised and reinforcement-learning models. To maintain consistency across task implementations, we ensure that a human subject should be able to participate in all tasks with similar inputs and controls. Although certain tasks make use of objects in different ways, generally objects' representations are retained across tasks (e.g. 'food' is a green circle).

4.3 Precursor knowledge

All of the experiments for humans and animals require a wealth of precursor 'common sense' knowledge, such as object detection, memory, navigation, etc.

Many of these experiments are intended to be performed on a subject lacking certain prior experience, but exactly what experience they can be allowed to have is not clear. In Rabinowitz' et al. implementation of Sally Anne, they account for agents having novel goal preferences by training a multitude of agents with random preferential permutations (Rabinowitz et al. 2018). That way, ToMNet's training involves repeatedly learning other agents' preferences from recent memory (ontogenetically). But can we expect the same of novel objects, like translucent glass? In our task-based approach, we may generally approach these sorts of problems by providing a multitude of training variants, e.g. a training set that includes translucent glass and one that does not.

4.4 Implementation

Due to its ease of use, runtime speed, and imagistic representations, we opt to adapt MarlGrid (Ndousse et al. 2021), a multi-agent fork of MiniGrid (Chevalier-Boisvert, Willems, and Pal 2018), an open-source implementation of a gridworld for reinforcement learning in a setting that is compatible with the popular OpenAI Gym (Brockman et al. 2016).

4.5 Task Selection

For our initial set of tasks to consider for implementation, we select those with the most apparent translations to gridworld environments. This set includes several tasks that use verbal or image-based storytelling, as many of these stories can be expressed with observable events.

As mentioned in Section 1, (Beaudoin et al. 2020) divides ToM tasks into seven categories: emotions, desires, intentions, percepts, knowledge, beliefs, and mentalistic understanding of non-literal communication. Although emotional understanding is a valuable aspect of social intelligence, we opt to omit it from our tasks due to the bounty of existing work in the field of affective computing and the difficulty of translating the complexity of emotions to simplified systems. Likewise, non-literal communication is a complex concept that cannot be easily translated to toy environments due to its dependence on natural language understanding, so we omit that category as well.

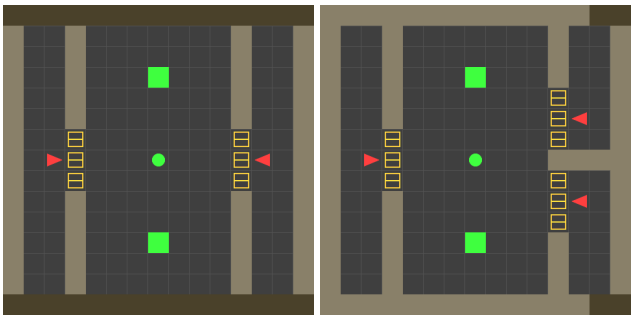

Figure 1: General setups for experiments CF1 and CF2. Individual test conditions and probe trials differ only in the sequence of changes to the environment, including the ordering of doors opening, the opacity of the dominant's door, and the food's conditional re-positioning. During the simulated 'baiting', the food object (green circle) moves to overlap one of the green squares, where it is no longer visible to either agent. Agents and doors are depicted as red triangles and yellow-barred boxes, respectively. This specific setup differs slightly from (Hare, Call, and Tomasello 2001) in that the subordinate cannot see the food after this stage, and must remember its location.

4.6 Detailed Example: Competitive Feeding

The competitive feeding paradigm is a test for specific ToM skills, like attribution of seeing and knowing, to conspecifics in social hierarchies.

The competitive feeding paradigm requires two animals, one of whom—the subject—is subordinate to the other in an existing social hierarchy. The two animals are kept on opposite sides of a central enclosure, separated from the enclosure by barriers (a top-down view of this task is shown in Figure 1). A researcher first places food on the subordinate's side of one barrier (baiting), and later moves it to the subordinate's side of another barrier. Finally, both animals are released, with the subordinate having a slight head start. Two conditions are varied: whether the dominant's door is open or closed during the first baiting, as well as during the second baiting; and whether the subject can see the dominant during the baiting.

This test attempts to distinguish ToM in animals by showing that the subject attempts to get more food when it can see the dominant, and knows the dominant does not know the location of food; i.e. whether the dominant's door is closed during the second baiting event. In other words, the subject must generate a ToM of the dominant agent to accurately predict whether the dominant will attempt to reach food at its first location or its second.

In (Hare, Call, and Tomasello 2001), three experiments are performed, each with its own set of testing conditions. During tests, the dominant's door opens only once the subordinate touches the floor of the central cage, or after thirty seconds, giving the subordinate a head start towards the food. Probe trials are randomly interspersed, in which food is placed in the open and both animals are released simultaneously. The purpose of these probe trials is to make sure the subordinate animal does not gain confidence in its ability to reach food before its competitor.

Experiment CF1a-d The goal of experiment 1 is to test for the attribution of sight, or the answer to the question "did she see it hidden or moved?". Four testing variants are used, with names referring to the dominant's condition: Uninformed, in which the subordinate has vision of the dominant, but the dominant is unable to see the baiting; Control uninformed (competitor informed), in which both subjects may see one another and the dominant observes the baiting; Misinformed, in which subjects see one another as food is placed, but then the dominant's door is closed and the food is moved to a new location; and Control misinformed (competitor informed), in which food is moved as in Misinformed but in view of the dominant.

Experiment CF2a-c Experiment 2 requires the subject to distinguish *who* saw it hidden, of multiple potential competitors. Two dominant competitors are placed in cages opposing the subject, but only one witnesses the baiting. The two test conditions are each simply releasing one of each of the competitors: the one who witnessed or the one who did not witness the baiting. It is made apparent to the subject which competitor will be released before it is released.

Experiment CF3a-d Experiment 3 uses multiple food objects to study whether the subject can understand *which* piece is seen hidden by a competitor. Now, there are three food locations, and two pieces of food are placed during the baiting. The same four conditions as experiment 1 are used, except the dominant always sees the first baiting, but only conditionally witnesses the second baiting or the movement of one piece of food.

Experiment CF4a-i (Penn and Povinelli 2007) argue the competitive feeding paradigm does not distinguish theory of mind from non-mentalist problem solving. In this version, there are n (e.g., 5) lanes, each with a food bucket with hidden contents. After initial exposures, nine separate conditions are presented randomly to the subjects, eliminating the possibility of solution via a single, simple strategy.

Further details The subject should be pretrained until familiar with several concepts, including that food is hidden under similarly-colored tiles; that competitors have similar perceptions, actions, and goal-driven intentions (they will always attempt to take the food if they see it); and that doors have three distinct states that sometimes change spontaneously: open, closed (opaque), and closed (transparent). **Precursors CF0**, then, are designed to integrate all three of these concepts in randomly generated settings.

5 Discussion: Challenges and Promise

Despite its theoretical and demonstrated usefulness in both the natural world and in artificial settings, the cognitive requirements—fulfilling both necessity and sufficiency—of ToM are not well understood. These abilities are rare in the natural world, so logic dictates they must be either very difficult to create or are only useful in niche circumstances. By implementing a cognitive model that demonstrates these abilities, and by testing that model in a variety of environments, we may learn what is necessary and sufficient for the model's success, and which environmental conditions encourage agents to train and make use of such a model, even at significant cost to the agent.

We hypothesize that ToM's presence alongside other advanced cognitive abilities in the human repertoire is no coincidence; many of the abilities we consider uniquely human (e.g. compositionality, etc.) have roots in the same core mental constructs. Given the recent success of (Rabinowitz et al. 2018) at training an agent to correctly answer questions regarding other agents' false beliefs, we believe a similar implementation will provide an excellent starting point for further development.

While developmental psychology has produced evidence of somewhat-regular sequential orderings, or stages, in which skills often emerge (Piaget 1976), the understanding of skills' intrinsic dependencies in the field of artificial intelligence is fairly underdeveloped. Transfer and curriculum learning are already massive fields of study, but—perhaps due to AI's relatively more easily accessible nature—these studies tend to aim to capture the admittedly more alluring concept of skills themselves rather than rote task performance.

One potential direction for our ToM-Testbed will be to organize tasks according to a ladder or graph of dependencies, based on findings from the human and animal literature. Then, we can examine these dependencies in the context of transfer learning and curriculum learning. For instance, to what extent does training on precursor tasks result in more efficient or more robust higher-level ToM abilities?

6 Conclusion

In this work we addressed the immense potential in leveraging the diverse tasks invented by biologists and psychologists to study ToM in animals for AI research. The development of these tasks initially required careful planning to overcome the many alternate intelligent and unintelligent explanations of animal behavior. We examine 21 tasks from the cognitive literature, including many more sub-tasks, for their eligibility in a battery of tests for the training and evaluation of artificial agents. We present a brief description of the setup and goal of each task examined, found in the Appendix. After discussing the desirable properties of a ToM-Testbed, we examine the process of translating one task, competitive feeding, to a simplified multi-agent gridworld environment. Finally, we discuss how the endeavor of understanding ToM skills may present a challenging frontier, but also the promise of helping us—and our bots—better understand each other.

Appendix: ToM Task Descriptions

Preferences

Yummy-yucky The Yummy-yucky task is designed to tell whether a subject is able to attribute preferences to an experimenter (Repacholi and Gopnik 1997). First, the subject's preferences are established by allowing them time with two bowls of different foods. An experimenter tries one food and then the other, and makes expressions of either disgust or happiness. The experimenter requests food from the subject by placing their palm halfway between the bowls.

Multiple desires (Bennett and Galpert 1993) test for children's ability to attribute multiple desires to another being. In these tests, children are told a story, and are then asked to answer specific questions, either verbally or by choosing a picture of what might happen next. Three variants are tested: successive desires, simultaneous and contradictory desires, and scenarios involving false beliefs.

Perception

Picture identification The picture identification task is a simple task of perspective-taking ability (Masangkay et al. 1974). A subject is shown a flat occluder with a picture on both sides. The occluder is rotated so that one side faces the subject, and one faces the experimenter. The subject is then asked questions such as what it is able to see, what the experimenter is able to see, and whether the experimenter can see the picture on the subject's side of the occluder.

Appearance-reality The appearance-reality task is a general framework for distinguishing whether a subject has the ability to understand that objects' appearances and true natures sometimes differ (Flavell et al. 1986). For example, a red car held behind a green pane of glass might appear black. Experimenters question subjects about their perceptual experience and reality (e.g. "What color is the car really?") under different circumstances, such as when the car is only partially occluded by the green pane.

Intent

Two-action The two-action test is a general framework for differentiating imitation and emulation in animals (Akins and Zentall 1996). Experimenters demonstrate one of two methods by which a subject may achieve a reward. The subject's behaviors are then recorded and compared with control subjects' behaviors without demonstration.

Distinguishing intentions from accidents In this task, subjects choose one box out of three based on a mark placed by the experimenter. In each trial, two boxes out of the three are marked, one *intentionally* and one (by way of observed performance) *accidentally*. Subjects are rewarded for choosing the box which the experimenter marked intentionally (Call and Tomasello 1998).

Rational Imitation The rational imitation task is similar to the two-action test, but in this variant the demonstrator is sometimes shown to have a reason for performing a task in an inconvenient way, e.g. using their head to flip a light switch because their hands are occupied (Meltzoff 1988; Gergely, Bekkering, and Király 2002). Now we may compare whether a subject truly understands the demonstrator's goal-oriented behavior, as the subject might reason that they are able to use their hands to complete a task rather than directly imitating the demonstrator.

Accidental/Moral transgression In this task, a subject is presented with a story involving either an accidental or a moral transgression (Killen et al. 2011). An accidental transgression would be a mistake made by a character due to their lack of understanding, i.e. a false belief. For example, a character might throw a bag in the trash without knowing it contains another character's prized possession. A moral transgression would have the same outcome in the story but is performed intentionally by the character. The subject is asked numerous questions about the scenario similar to those in the Sally–Anne task, but including additional questions about whether characters should be punished and why.

Knowledge

Competitive Feeding The competitive feeding paradigm requires two animals, one of whom (the subject) is considered subordinate to the other in their social hierarchy (Hare, Call, and Tomasello 2001). The two animals are kept on opposite sides of an enclosure with two barriers. A researcher first places food on the subordinate's side of one barrier, and later moves it to the subordinate's side of another barrier. Finally, both animals are released, with the subordinate having a slight head start.

Knower-Guesser The knower-guesser paradigm is a commonly used method for determining whether animals can attribute concepts such as 'seeing' and 'knowing' (Udell, Dorey, and Wynne 2011). It also allows a nonverbal subject to directly participate in an interspecific exercise rather than simply observing. Two human experimenters, the Knower and the Guesser, are presented to an animal subject. The Guesser leaves the room (or has their gaze somehow occluded), while the Knower places food in some location that is not visible to the subject. The Guesser returns, and then both Knower and Guesser point to places where they think the food is located. The subject may then search one container for food and keep it as a reward.

Goggles Because a subject of a knower-guesser test might employ a number of non-mentalistic strategies, in the goggles/visor test the difference between knower and guesser may only be correctly inferred by generalizing from first-person experience with an occluding object (Karg et al. 2015). In this variant, the subject spends some time with an object that appears opaque from a distance, like goggles or a wire screen. This object may be either opaque or translucent. Then, a test like knower-guesser is performed, with the experimenter's vision being occluded by the object. In the test, the object is always opaque, so the experimenter is truly blind and the subject has no chance of seeing their eyes.

See-Know This task is similar to the Knower-Guesser task, with the exception that the evaluation is verbal in nature (Pillow 1989). In one version of the task, the subject is either the Knower or the Guesser, while the other role is played by a puppet. The Knower then observes the process of hiding a toy in a box and the subject is quizzed on whether

they or the puppet know the color of the toy in the box. In the other variant, there are two puppets which play both roles, and the subject is asked to attribute knowledge about their knowledge and percept to either of them.

Beliefs

Sally Anne The Sally Anne test, also referred to as the standard FB test or the change-of-location FB test, is a commonly used test for the attribution of false beliefs (Wimmer and Perner 1983). The prototypical Sally Anne test, first used in 1985 by (Baron-Cohen et al. 1985), involves the use of puppets to tell a short story. Sally places her toy in one location, and then leaves the room. Next, unbeknownst to Sally, Anne moves the toy to a new location. Finally, Sally returns to look for her toy. Several control questions establish that the subject understands the basic story elements such as the characters' names and the toy's location. The subject is then asked: "Where will Sally look for her toy?"

Ignorance and False Belief While false belief tasks require some form of advanced understanding of somebody else's mental state, understanding ignorance of certain knowledge in others is likely an easier task. The tasks for testing ignorance in others follow a similar pattern as False Belief tests, with the modification that the subject is directly asked if the other participant is aware of the location of the manipulated object (Hogrefe, Wimmer, and Perner 1986).

Inhibitory FB The inhibitory false-belief test is intended to explain successful performance at the Sally Anne test (Leslie and Polizzi 1998). In addition to having true or false beliefs, characters might have positive or negative desires. In the negative desire condition, the Sally character wants to look in a container where the hidden object is *not* located. In the Opposite behavior condition, Sally is introduced as an odd person who always does things she does not want.

Deception

Penny-hiding The penny-hiding game is a simple test of deception (Gratch 1964). During training, an experimenter repeatedly allows the subject to guess which of their closed hands hides a penny. Rather than leave the results to chance, in some number of trials both hands contain a penny, and in other trials neither does. For the test, the roles are reversed: the subject is asked to hide a penny in one of their hands, and the interviewer guesses which. The subject's hands are visible during the hiding, so they may only hide the penny's location by repeatedly passing it between their hands or imitating the same action. Subjects are graded by the experimenter based on their apparent use of deceptive strategies.

Box-locking The box-locking test examines subjects' abilities in settings that allow for sabotage and verbal deception (Sodian and Frith 1992). First, puppets are introduced to the subject along with rewards for 'success': the friendly seal shares what it finds, but the thieving wolf takes everything for itself. In both settings, a reward is hidden in a box, and the subject is tasked with making sure the seal is able to find the reward, but the wolf is not. In the sabotage setting, the child is able to use a key to lock the box, physically preventing a puppet character from opening it. In the deception setting, the child is asked by the puppet character whether the box is locked, and the child may lie to prevent it from attempting to open the box. A minor variation involves the use of two boxes, so the child may lock either one, or may lie about the location of the reward.

Back-and-forth foraging This task studies whether subjects are able to ascribe their own reward preferences to competitors (Schmelz, Call, and Tomasello 2011). Two rewards are hidden under boards on a platform, as viewed by the subject. One of the boards has a hole below it so that that board is flat after the reward is put under it, while the other one is slanted. The platform is then presented first to the competitor, who has to choose one of the reward items. Then, the platform is presented to the subject, who has to decide which reward to pick. In a control condition, the subjects do not display strong preference towards either reward in the absence of the competitor. The social condition tests whether the subject determines that the competitor would go for the reward under the slanted board and choose the other.

Caching food from unseen competitors While changes in food caching behavior of scrubjays in the presence of competitors has been well-documented, most experiments allow the subject direct access to the conspecific's gaze. This task seeks to test whether ravens can use the fact that unseen competitors have visual access to their food-caching behavior and alter their behavior based on that (Bugnyar, Reber, and Buckner 2016). The subject is put in a room and audio recordings of other scrub-jays are played from behind a closed window which has a peephole in it. The subjects are made aware to the presence of the peephole by the experimenter. The task tests whether the subjects infer the presence of competitors in the adjacent room by the sound and alter their behavior based on the assumption that the competitor can see them through the peephole.

Other

Mirror self-recognition Mirror self-recognition tests are generally tests for bodily self-awareness (Gallup 1970). A subject is given exposure time with mirrors, with increasing proximity. In some observational variants, experimenters observe the subject and rate its behaviors as social (e.g. trying to communicate with the mirrored self), versus self-directed (e.g. using the mirror to help clean its own teeth). Often researchers will paint a marking on a subject's body in a location that would otherwise be undetectable to them, such as their forehead. The subject may then touch the mark after being exposed to it via the mirror.

Role-reversal The role-reversal test is similar to the penny-hiding test, but in a cooperative scenario instead of competitive (Povinelli, Nelson, and Boysen 1992). Using the same apparatus as for the knower-guesser test, a subject is trained to take on one of two roles: the informant, or the operator. The other role is performed by an experimenter. The informant is able to see where food is hidden, and may communicate that information to the operator. The operator then pulls a lever and shares the food with their partner. Success is determined by the operator's correct choice of food location. After subjects learn their roles successfully, they are given the alternate role and tested for their ability to complete the new task without training.

Acknowledgements

This work was supported in part by the Neurodiversity Inspired Science and Engineering (NISE) NSF program grant DGE 19-22697 (K. Stassun, PI). We also extend thanks to our anonymous reviewers for their helpful criticisms, comments, and suggestions.

References

Akins, C. K.; and Zentall, T. R. 1996. Imitative learning in male Japanese quail (Coturnix japonica) using the two-action method. *Journal of Comparative Psychology* 110(3): 316.

Baron-Cohen, S.; Leslie, A. M.; Frith, U.; et al. 1985. Does the autistic child have a "theory of mind". *Cognition* 21(1): 37–46.

Beaudoin, C.; Leblanc, É.; Gagner, C.; and Beauchamp, M. H. 2020. Systematic review and inventory of theory of mind measures for young children. *Frontiers in psychology* 10: 2905.

Bellemare, M. G.; Naddaf, Y.; Veness, J.; and Bowling, M. 2013. The arcade learning environment: An evaluation platform for general agents. *Journal of Artificial Intelligence Research* 47: 253–279.

Bennett, M.; and Galpert, L. 1993. Children's understanding of multiple desires. *International Journal of Behavioral Development* 16(1): 15–33.

Bird, G.; and Viding, E. 2014. The self to other model of empathy: providing a new framework for understanding empathy impairments in psychopathy, autism, and alexithymia. *Neuroscience & Biobehavioral Reviews* 47: 520–532.

Brockman, G.; Cheung, V.; Pettersson, L.; Schneider, J.; Schulman, J.; Tang, J.; and Zaremba, W. 2016. Openai gym. *arXiv preprint arXiv:1606.01540* .

Bugnyar, T.; Reber, S. A.; and Buckner, C. 2016. Ravens attribute visual access to unseen competitors. *Nature Communications* 7(1): 1–6.

Call, J.; and Tomasello, M. 1998. Distinguishing intentional from accidental actions in orangutans (Pongo pygmaeus), chimpanzees (Pan troglodytes) and human children (Homo sapiens). *Journal of Comparative Psychology* 112(2): 192.

Carlson, S. M.; and Moses, L. J. 2001. Individual differences in inhibitory control and children's theory of mind. *Child development* 72(4): 1032–1053.

Chevalier-Boisvert, M.; Willems, L.; and Pal, S. 2018. Minimalistic Gridworld Environment for OpenAI Gym. https://github.com/maximecb/gym-minigrid.

Crosby, M.; Beyret, B.; Shanahan, M.; Hernández-Orallo, J.; Cheke, L.; and Halina, M. 2020. The Animal-AI Testbed and Competition. In Escalante, H. J.; and Hadsell, R., eds., *Proceedings of the NeurIPS 2019 Competition and Demonstration Track*, volume 123 of *Proceedings of Machine Learning Research*, 164–176. PMLR.

Flavell, J. H.; Green, F. L.; Flavell, E. R.; Watson, M. W.; and Campione, J. C. 1986. Development of knowledge about the appearance-reality distinction. *Monographs of the society for research in child development* i–87.

Gallup, G. G. 1970. Chimpanzees: self-recognition. *Science* 167(3914): 86–87.

Gergely, G.; Bekkering, H.; and Király, I. 2002. Rational imitation in preverbal infants. *Nature* 415(6873): 755–755.

Gopnik, A.; and Wellman, H. M. 1994. The theory theory. In *An earlier version of this chapter was presented at the Society for Research in Child Development Meeting, 1991.* Cambridge University Press.

Gould, J. L. 1990. Honey bee cognition. *Cognition* 37(1-2): 83–103.

Gould, J. L.; Gould, C. G.; et al. 1988. *The honey bee.* Scientific American Library.

Gratch, G. 1964. Response alternation in children: A developmental study of orientations to uncertainty. *Vita humana* 49–60.

Happé, F.; and Frith, U. 1995. Theory of mind in autism. In *Learning and cognition in autism*, 177–197. Springer.

Hare, B.; Call, J.; and Tomasello, M. 2001. Do chimpanzees know what conspecifics know? *Animal behaviour* 61(1): 139–151.

Harris, P. L. 1992. From simulation to folk psychology: the case for development. *Mind & Language* .

Hernandez-Leal, P.; Kartal, B.; and Taylor, M. E. 2019. A survey and critique of multiagent deep reinforcement learning. *Autonomous Agents and Multi-Agent Systems* 33(6): 750–797.

Hogrefe, G.-J.; Wimmer, H.; and Perner, J. 1986. Ignorance versus false belief: A developmental lag in attribution of epistemic states. *Child development* 567–582.

Hughes, C. 1998. Finding your marbles: Does preschoolers' strategic behavior predict later understanding of mind? *Developmental psychology* 34(6): 1326.

Karg, K.; Schmelz, M.; Call, J.; and Tomasello, M. 2015. The goggles experiment: Can chimpanzees use self-experience to infer what a competitor can see? *Animal Behaviour* 105: 211–221.

Killen, M.; Mulvey, K. L.; Richardson, C.; Jampol, N.; and Woodward, A. 2011. The accidental transgressor: Morally-relevant theory of mind. *Cognition* 119(2): 197–215.

Krupenye, C.; Kano, F.; Hirata, S.; Call, J.; and Tomasello, M. 2016. Great apes anticipate that other individuals will act according to false beliefs. *Science* 354(6308): 110–114.

Leslie, A. M. 1994. Pretending and believing: Issues in the theory of ToMM. *Cognition* 50(1-3): 211–238.

Leslie, A. M.; and Polizzi, P. 1998. Inhibitory processing in the false belief task: Two conjectures. *Developmental science* 1(2): 247–253.

Masangkay, Z. S.; McCluskey, K. A.; McIntyre, C. W.; Sims-Knight, J.; Vaughn, B. E.; and Flavell, J. H. 1974. The early development of inferences about the visual percepts of others. *Child development* 357–366.

Meltzoff, A. N. 1988. Infant imitation after a 1-week delay: long-term memory for novel acts and multiple stimuli. *Developmental psychology* 24(4): 470.

Ndousse, K. K.; Eck, D.; Levine, S.; and Jaques, N. 2021. Emergent Social Learning via Multi-agent Reinforcement Learning. In *International Conference on Machine Learning*, 7991–8004. PMLR.

Penn, D. C.; Holyoak, K. J.; and Povinelli, D. J. 2008. Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences* 31(2): 109–130.

Penn, D. C.; and Povinelli, D. J. 2007. On the lack of evidence that non-human animals possess anything remotely resembling a 'theory of mind'. *Philosophical Transactions of the Royal Society B: Biological Sciences* 362(1480): 731–744.

Perez-Liebana, D.; Hofmann, K.; Mohanty, S. P.; Kuno, N.; Kramer, A.; Devlin, S.; Gaina, R. D.; and Ionita, D. 2019. The Multi-Agent Reinforcement Learning in MalmÖ (MARLÖ) Competition. *arXiv preprint arXiv:1901.08129* .

Peterson, C. C.; and Siegal, M. 1999. Representing inner worlds: Theory of mind in autistic, deaf, and normal hearing children. *Psychological science* 10(2): 126–129.

Piaget, J. 1976. Piaget's theory. In *Piaget and his school*, 11–23. Springer.

Pillow, B. H. 1989. Early understanding of perception as a source of knowledge. *Journal of experimental child psychology* 47(1): 116–129.

Povinelli, D. J.; Nelson, K. E.; and Boysen, S. T. 1992. Comprehension of role reversal in chimpanzees: Evidence of empathy? *Animal behaviour* 43(4): 633–640.

Rabinowitz, N.; Perbet, F.; Song, F.; Zhang, C.; Eslami, S. A.; and Botvinick, M. 2018. Machine theory of mind. In *International conference on machine learning*, 4218–4227. PMLR.

Repacholi, B. M.; and Gopnik, A. 1997. Early reasoning about desires: evidence from 14-and 18-month-olds. *Developmental psychology* 33(1): 12.

Russell, J. 1997. How executive disorders can bring about an inadequate'theory of mind'. In *Autism as an executive disorder*, 256—-304. Oxford University Press.

Schmelz, M.; Call, J.; and Tomasello, M. 2011. Chimpanzees know that others make inferences. *Proceedings of the National Academy of Sciences* 108(7): 3077–3079.

Sodian, B.; and Frith, U. 1992. Deception and sabotage in autistic, retarded and normal children. *Journal of child psychology and psychiatry* 33(3): 591–605.

Song, Y.; Wojcicki, A.; Lukasiewicz, T.; Wang, J.; Aryan, A.; Xu, Z.; Xu, M.; Ding, Z.; and Wu, L. 2020. Arena: A general evaluation platform and building toolkit for multi-agent intelligence. *Proceedings of the AAAI Conference on Artificial Intelligence* 34(05): 7253–7260.

Southgate, V.; Senju, A.; and Csibra, G. 2007. Action anticipation through attribution of false belief by 2-year-olds. *Psychological science* 18(7): 587–592.

Udell, M. A.; Dorey, N. R.; and Wynne, C. D. 2011. Can your dog read your mind? Understanding the causes of canine perspective taking. *Learning & Behavior* 39(4): 289–302.

Wimmer, H.; and Perner, J. 1983. Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13(1): 103–128.