Bayesian Theory of Mind: Modeling Joint Belief-Desire Attribution

Chris L. Baker (clbaker@mit.edu)
Rebecca R. Saxe (saxe@mit.edu)
Joshua B. Tenenbaum (jbt@mit.edu)
Department of Brain and Cognitive Sciences, MIT
Cambridge, MA 02139

Abstract

We present a computational framework for understanding *Theory of Mind (ToM)*: the human capacity for reasoning about agents' mental states such as beliefs and desires. Our Bayesian model of ToM (or BToM) expresses the predictive model of belief- and desire-dependent action at the heart of ToM as a *partially observable Markov decision process (POMDP)*, and reconstructs an agent's joint belief state and reward function using Bayesian inference, conditioned on observations of the agent's behavior in some environmental context. We test BToM by showing participants sequences of agents moving in simple spatial scenarios and asking for joint inferences about the agents' desires and beliefs about unobserved aspects of the environment. BToM performs substantially better than two simpler variants: one in which desires are inferred without reference to an agent's beliefs, and another in which beliefs are inferred without reference to the agent's dynamic observations in the environment.

Keywords: Theory of mind; Social cognition; Action understanding; Bayesian inference; Partially Observable Markov Decision Processes

Introduction

Central to human social behavior is a *theory of mind (ToM)*, the capacity to explain and predict people's observable actions in terms of unobservable mental states such as beliefs and desires. Consider the case of Harold, who leaves his dorm room one Sunday morning for the campus library. When he reaches to open the library's front door he will find that it is locked – closed on Sunday. How can we explain his behavior? It seems plausible that he wants to get a book, that he believes the book he wants is at the library, and that he also believes (falsely, it turns out) that the library is open on Sunday.

Such mental state explanations for behavior go well beyond the observable data, leading to an inference problem that is fundamentally ill-posed. Many different combinations of beliefs and desires could explain the same behavior, with inferences about the strengths of beliefs and desires trading off against each other, and relative probabilities modulated heavily by context. Perhaps Harold is almost positive that the library will be closed, but he needs a certain book so badly that he still is willing to go all the way across campus on the off chance it will be open. This explanation seems more probable if Harold shows up to find the library locked on Saturday at midnight, as opposed to noon on Tuesday. If he arrives after hours already holding a book with a due date of tomorrow, it is plausible that he knows the library is closed and is seeking not to get a new book, but merely to return a book checked out previously to the night drop box.

Several authors have recently proposed models for how people infer others' goals or preferences as a kind of Bayesian inverse planning or inverse decision theory (Baker, Saxe, & Tenenbaum, 2009; Feldman & Tremoulet, 2008; Lucas, Griffiths, Xu, & Fawcett, 2009; Bergen, Evans, & Tenenbaum, 2010; Yoshida, Dolan, & Friston, 2008; Ullman et al., 2010). These models adapt tools from control theory, econometrics and game theory to formalize the *principle of rational action* at the heart of children and adults' concept of intentional agency (Gergely, Nádasdy, Csibra, & Biró, 1995; Dennett, 1987): all else being equal, agents are expected to choose actions that achieve their desires as effectively and efficiently as possible, i.e., to maximize their expected utility. Goals or preferences are then inferred based on which objective or utility function the observed actions maximize most directly.

ToM transcends knowledge of intentional agents' goals and preferences by incorporating representational mental states such as subjective beliefs about the world (Perner, 1991). In particular, the ability to reason about *false* beliefs has been used to distinguish ToM from non-representational theories of intentional action (Wimmer & Perner, 1983; Onishi & Baillargeon, 2005). Our goal in this paper is to model human ToM within a Bayesian framework. Inspired by models of inverse planning, we cast Bayesian ToM (BToM) as a problem of inverse planning and inference, representing an agent's planning and inference about the world as a partially observable Markov decision process (POMDP), and inverting this forward model using Bayesian inference. Critically, this model includes representations of both the agent's desires (as a utility function), and the agent's own subjective beliefs about the environment (as a probability distribution), which may be uncertain and may differ from reality. We test the predictions of this model quantitatively in an experiment where people must simultaneously judge beliefs and desires for agents moving in simple spatial environments under incomplete or imperfect knowledge.

Important precursors to our work are several computational models (Goodman et al., 2006; Bello & Cassimatis, 2006; Goodman, Baker, & Tenenbaum, 2009) and informal theoretical proposals by developmental psychologists (Wellman, 1990; Gopnik & Meltzoff, 1997; Gergely & Csibra, 2003). Goodman et al. (2006) model how belief and desire inferences interact in the classic "false belief" task used to assess ToM reasoning in children (Wimmer & Perner, 1983). This model instantiates the schema shown in Fig. 1(a) as a causal Bayesian network with several psychologically interpretable, but task-dependent parameters. Goodman et al. (2009) model adult inferences of an agent's knowledge of the causal structure of a simple device ("Bob's box") based on observing the agent interacting with the device. To our knowledge, our work here is the first attempt to explain people's joint inferences about agents' beliefs and desires by explicitly inverting POMDPs – and the first model capable of reasoning about the graded strengths and interactions between agents' beliefs and desires, along with the origins of agents' beliefs via environmentally constrained perceptual observations.

Computational Framework

This section describes Bayesian Theory of Mind (BToM): a theory-based Bayesian framework (Tenenbaum, Griffiths, & Kemp, 2006) that characterizes ToM in terms of Bayesian inference over a formal, probabilistic version of the schema in Fig. 1(a). BToM represents an ideal *observer* using a theory of mind to understand the actions of an individual *agent* within some environmental context. This ideal-observer analysis of ToM asks how closely human judgments approach the ideal limit, but also what mental representations are necessary to explain human judgments under hypothetically unbounded computational resources. We will first describe BToM in general, but informal terms before progressing to the mathematical details involved in modeling our experimental domain.

Informal sketch

For concreteness, we use as a running example a simple spatial context (such as a college campus or urban landscape) defined by buildings and perceptually distinct objects, with agents' actions corresponding to movement, although in general BToM can be defined over arbitrary state and action spaces (for example, a card game where the state describes players' hands and actions include draw or fold). The observer's representation of the world is composed of the *environment state* and the *agent state* (Fig. 1(a)). In a spatial context, the state of the environment represents its physical configuration, e.g., the location of buildings and objects, and the state of the agent specifies its objective, external properties, such as its physical location in space.

The observer's theory of the agent's mind includes representations of the agent's subjective desires and beliefs, and the principles by which desires and beliefs are related to actions and the environment. Similar to previous models, the content of the agent's desire consists of objects or events in the world. The agent's degree of desire is represented in terms of the subjective reward received for taking actions in certain states, e.g., acting to attain a goal while in close proximity to the goal object. The agent can also act to change its own state or the environment state at a certain cost, e.g., navigating to reach a goal may incur a small cost at each step.

The main novel component of the current model is the inclusion of a representation of beliefs. Like desires, beliefs are defined by both their content and the strength or degree with which they are held. The content of a belief is a representation corresponding to a possible world. For instance, if the agent is unsure about the location of a particular object, its belief contents are worlds in which the object is in different locations. The agent's degree of belief reflects the subjective probability it assigns to each possible world.

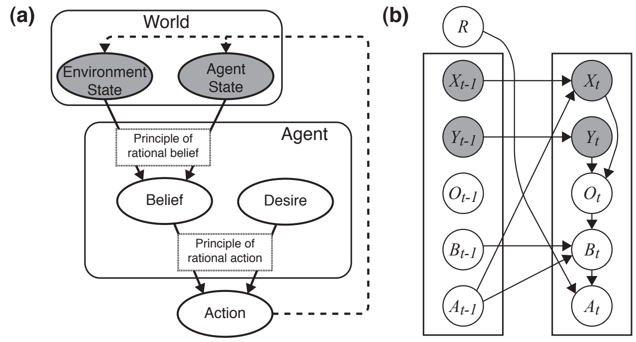


Figure 1: Causal structure of theory of mind. Grey shaded nodes are assumed to be observed (for the observer; not necessarily for the agent, as described in the main text). **(a)** Schematic model of theory of mind. Traditional accounts of ToM (e.g. Dennett, 1987; Wellman, 1990; Gopnik & Meltzoff, 1997) have proposed informal versions of this schema, characterizing the content and causal relations of ToM in commonsense terms, e.g., "seeing is believing" for the principle of rational belief. **(b)** Observer's grounding of the theory as a dynamic Bayes net (DBN). The DBN encodes the observer's joint distribution over an agent's beliefs $B_{1:T}$ and desires R over time, given the agent's physical state sequence $x_{1:T}$ in environment y.

The principles governing the relation between the world and the agent's beliefs, desires and actions can be naturally expressed within partially observable Markov decision processes (POMDPs). POMDPs capture the causal relation between beliefs and the world via the principle of rational belief, which formalizes how the agent's belief is affected by observations in terms of Bayesian belief updating. Given an observation, the agent updates its degree of belief in a particular world based on the likelihood of receiving that observation in that world. In a spatial setting, observations depend on the agent's line-of-sight visual access to features of the environment. POMDPs represent how beliefs and desires cause actions via the principle of rational action, or rational planning. Intuitively, rational POMDP planning provides a predictive model of an agent optimizing the tradeoff between exploring the environment to discover the greatest rewards, and exploiting known rewards to minimize costs incurred.

On observing an agent's behavior within an environment, the beliefs and desires that caused the agent to generate this behavior are inferred using Bayesian inference. The observer maintains a hypothesis space of joint beliefs and desires, which represent the agent's initial beliefs about the environment state and the agent's static desires for different goals. For each hypothesis, the observer evaluates the likelihood of generating the observed behavior given the hypothesized belief and desire. The observer integrates this likelihood with the prior over mental states to infer the agent's joint belief and desire.

As an example of how this works, consider Fig. 2. The "college campus" environment is characterized by the campus size, the location and size of buildings, and the location of

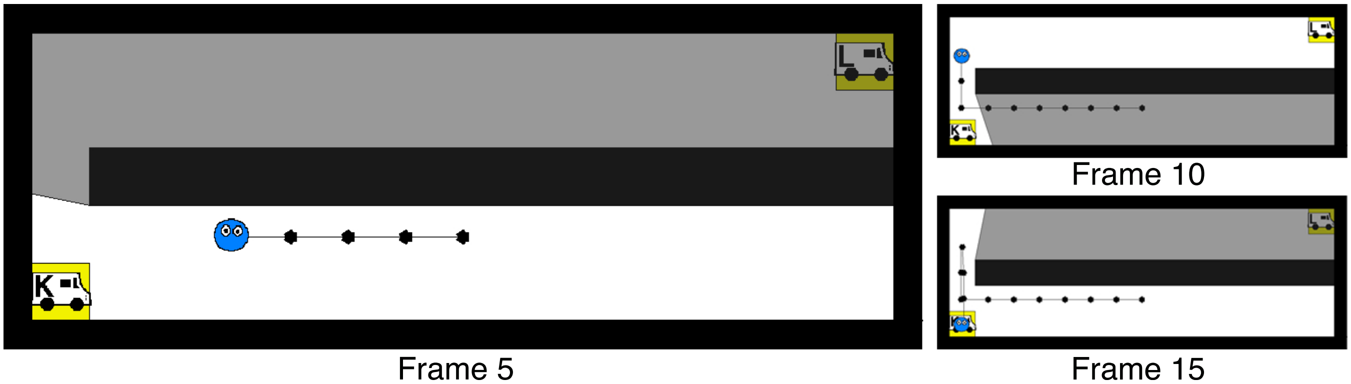


Figure 2: Example experimental stimulus. The small blue sprite represents the location of the agent, and the black trail with arrows superimposed records the agent's movement history. The two yellow cells in opposite corners of the environment represent spots where trucks can park, and each contains a different truck. The shaded grey area of each frame represents the area that is outside of the agent's current view.

several different goal objects, here "food trucks". The agent is a hungry graduate student, leaving his office and walking around campus in search of satisfying lunch food. There are three trucks that visit campus: Korean (K), Lebanese (L) and Mexican (M), but only two parking spots where trucks are allowed to park, highlighted with a yellow background in Fig. 2. The student's field of view is represented by the unshaded region of the environment.

In Fig. 2, the student can initially only see where K (but not L) is parked. Because the student can see K, they know that the spot behind the building either holds L, M, or is empty. By frame 10, the student has passed K, indicating that they either want L or M (or both), and believe that their desired truck is likely to be behind the building (or else they would have gone straight to K under the principle of rational action). After frame 10, the agent discovers that L is behind the building and turns back to K. Obviously, the agent prefers K to L, but more subtly, it also seems likely that the agent wants M more than either K or L, despite M being absent from the scene! BToM captures this inference by resolving the desire for L or M over K in favor of M after the agent rejects L. In other words, BToM infers the best explanation for the observed behavior – the only consistent desire that could lead the agent to act the way it did.

Formal modeling

In the food-truck domain, the agent occupies a discrete state space $\mathcal{X}$ of points in a 2D grid. The environment state $\mathcal{Y}$ is the set of possible assignments of the K, L and M trucks to parking spots. Possible actions include North, South, East, West, Stay, and Eat. Valid actions yield the intended transition with probability $1-\epsilon$ and do nothing otherwise; invalid actions (e.g., moving into walls) have no effect on the state.

The agent's visual observations are represented by the *isovist* from the agent's location: a polygonal region containing all points of the environment within a 360-degree field of view (Davis & Benedikt, 1979; Morariu, Prasad, & Davis, 2007). Example isovists from different locations in one environment are shown in Fig. 2. The observation distribution $P(o|x,y)$ encodes which environments in $\mathcal{Y}$ are consistent with the contents of the isovist from location x. We model observation noise with the simple assumption that ambiguous observations can occur with probability ν, as if the agent failed to notice something that should otherwise be visible.

The observer represents the agent's belief as a probability distribution over $\mathcal{Y}$; for $y \in \mathcal{Y}$, $b(y)$ denotes the agent's degree of belief that y is the true state of the environment. Bayesian belief updating at time t is a deterministic function of the prior belief b_{t-1}, the observation o_t, and the world state $\langle x_t, y\rangle$. The agent's updated degree of belief in environment y satisfies $b_t(y) \propto P(o_t|x_t,y)b_{t-1}(y)$.

The agent's reward function $R(x,y,a)$ encodes the subjective utility the agent derives from taking action a from the state $\langle x_t, y\rangle$. Each action is assumed to incur a cost of 1. Rewards result from taking the "Eat" action while at a food truck; the magnitude of the reward depends on the strength of the agent's desire to eat at that particular truck. Once the student has eaten, all rewards and costs cease, implying that rational agents should optimize the tradeoff between the number of actions taken and the reward obtained.

The agent's POMDP is defined by the state space, the action space, the world dynamics, the observation model, and the reward function. We approximate the optimal value function of the POMDP for each hypothesized reward function using a point-based value iteration algorithm over a uniform discretization of the belief space. The agent's policy is stochastic, given by the softmax of the lookahead state-action value function Q^{LA} (Hauskrecht, 2000): $P(a|b,x,y) \propto \exp(\beta Q^{LA}(b,x,y,a))$. The β parameter establishes the degree of determinism with which the agent executes its policy, capturing the intuition that agents tend to, but do not always follow the optimal policy.

Our approach to joint belief and desire inference is closely related the model of belief filtering in Zettlemoyer, Milch, and Kaelbling (2009), restricted to the case of one agent reasoning about the beliefs of another. Fig. 1(b) shows the observer's dynamic Bayes net (DBN) model of an agent's desires, states, observations, beliefs and actions over time. The observer's belief and reward inferences are given by the joint posterior marginal over the agent's beliefs and rewards at time t, given the state sequence up until $T \geq t$: $P(b_t, r|x_{1:T}, y)$. This computation is analogous to the forward-backward algorithm in hidden Markov models, and provides the basis for model predictions of people's joint belief and desire inferences in our experiment.

To perform inference over the multidimensional, continuous space of beliefs and rewards, we uniformly discretize the hypothesis spaces of beliefs and reward functions with grid resolutions of 7. The range of reward values was calibrated to the spatial scale of our environments, taking values $-20, 0, \ldots, 100$ for each truck. Model predictions were based on the student's expected reward value for each truck (K, L, M) and the expected degree-of-belief in each possible world for each trial.

Alternative models

To test whether the full representational capacity of our model is necessary to understand people's mental state attributions, we formulate two alternative models as special cases of our joint inference model. Each alternative model "lesions" a central component of the full model's representation of beliefs, and tests whether it is possible to explain people's inferences about agents' desires in our experiment without appeal to a full-fledged theory of mind.

Our first alternative model is called TrueBel. This model assumes that the state is fully observable to the agent, i.e., that the agent knows the location of every truck, and plans to go directly to the truck that will provide the maximal reward while incurring the least cost. We hypothesized that this model would correlate moderately well with people's desire judgments, because of the statistical association between desired objects and actions.

Our second alternative model is called NoObs. In this model, the agent has an initial belief about the state of the environment, but there is no belief updating – the initially sampled belief remains fixed throughout the trial. We hypothesized that this model might fit people's belief and desire inferences in situations where the agent appeared to move toward the same truck throughout the entire trial, but that for actions that required belief updating or exploration to explain, for instance, when the agent began by exploring the world, then changed direction based on its observation of the world state, NoObs would fit poorly.

Experiment

Fig. 4 illustrates our experimental design. Truck labels were randomized in each trial of the experiment, but we will describe the experiment and results using the canonical, unscrambled ordering Korean (K), Lebanese (L), Mexican (M).

The experiment followed a $3 \times 5 \times 2 \times 3 \times 2$ design. These factors can be divided into 30 ($3 \times 5 \times 2$) unique paths and 6 (3×2) unique environmental contexts. There were 3 different starting points in the environment: "Left", "Middle", or "Right"; all shown in Fig. 4. These starting points were crossed with 5 different trajectories: "Check-Left, go to K"; "Check-Left, go to L/M"; "Check-Right, go to K"; "Check-Right, go to L/M"; and "No-check, go straight to K". Four of these trajectories are shown in Fig. 4. Each path was shown with 2 different judgment points, or frames at which the animation paused and subjects gave ratings based on the information shown so far. Judgment points were either at the moment the student became able to see the parking spot that was initially occluded ("Middle"; e.g., frame 10 in Fig. 2), or at the end of the path once the student had eaten ("Ending"; e.g., frame 15 in Fig. 2). All potential paths were crossed with 6 environmental contexts, generated by combining 3 different building configurations: "O", "C" and "backwards C", (all shown in Fig. 4) with 2 different goal configurations: "One truck" or "Two trucks" present; both shown in Fig. 4.

After all possible trials from this design were generated, all invalid trials (in which the student's path intersected with a building), and all "Ending" trials in which the path did not finish at a truck were removed. This left 78 total trials. Of these, 5 trials had a special status. These were trials in the "O" environment with paths in which the student began at the Right starting point, and then followed a Check-Left trajectory. These paths had no rational interpretation under the BToM model, because the Check-Right trajectory was always a more efficient choice, no matter what the student's initial belief or desire. These "irrational" trials are analyzed separately in the Results section.

Several factors were counterbalanced or randomized. Stimulus trials were presented in pseudo-random order. Each trial randomly scrambled the truck labels, and randomly reflected the display vertically and horizontally so that subjects would remain engaged with the task and not lapse into a repetitive strategy. Each trial randomly displayed the agent in 1 of 10 colors, and sampled a random male or female name without replacement. This ensured that subjects did not generalize information about one student's beliefs or desires to students in subsequent trials.

The experimental task involved rating the student's degree of belief in each possible world (Lebanese truck behind the building (L); Mexican truck behind the building (M); or nothing behind the building (N)), and rating how much the student liked each truck. All ratings were on a 7-point scale. Belief ratings were made retrospectively, meaning that subjects were asked to rate what the student thought was in the occluded parking spot before they set off along their path, basing their inference on the information from the rest of the student's path. The rating task counterbalanced the side of the monitor on which the "likes" and "believes" questions were displayed.

Subjects first completed a familiarization stage that explained all details of our displays and the scenarios they depicted. To ensure that subjects understood what the students could and couldn't see, the familiarization explained the visualization of the student's isovist, which was updated along each step of the student's path. The isovist was displayed during the testing stage of the experiment as well (Fig. 2).

Participants were 17 members of the MIT subject pool, 6 female, and 11 male. One male subject did not understand the instructions and was excluded from the analysis.

Results & Discussion

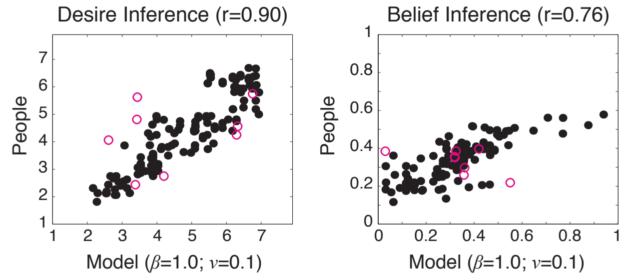
Debriefing of subjects suggested that many were confused by the "Middle" judgment point trials; this was also reflected by greater variability in people's judgments within these trials. Because of this, our analyses only include trials from the "Ending" judgment point condition, which accounted for 54 out of the 78 total trials.

We begin by analyzing the overall fit between people's judgments and our three models, and then turn to a more detailed look at several representative scenarios. Two parameters β and ν were fit for the BToM model; only the determinism parameter β is relevant for the TrueBel and NoObs models. Parameter fits are not meant to be precise; we report the best values found among several drawn from a coarse grid.

BToM predicts people's judgments about agents' desires relatively well, and less well but still reasonably for judgments about agents' initial beliefs (Fig. 3). In Fig. 3, data from the "irrational" trials are plotted with magenta circles, and account for most of the largest outliers. TrueBel and NoObs fit significantly worse for desire judgments and provide no reasonable account of belief judgments. TrueBel's belief predictions are based on the actual state of the world in each trial; the poor correlation with people's judgments demonstrates that people did not simply refer to the true world state in their belief attributions. The NoObs model in principle can infer agents' beliefs, but without a theory of how beliefs are updated from observations it must posit highly implausible initial beliefs that correlate poorly with subjects' judgments over the whole set of experimental conditions.

Fig. 4 shows several revealing comparisons of human judgments and model predictions in specific cases. When the agent follows a long path to an unseen goal (A1) it is suggestive of a strong initial belief that a more desirable truck is present behind the wall. In contrast, going straight to a nearby observed truck says only that this truck is likely to be desired more than the others (A2). When the agent goes out of its way to check an unseen parking spot, sees the second truck there, and returns to the previously seen truck, it suggests a strong desire for the one truck not present (compare B1 to B2). Finally, the relative strengths of inferences about desires and initial beliefs are modulated by how far the agent must travel to observe the unseen parking spot (compare C1 to C2, and C3 to C4). In each of these cases people reflect the same qualitative trends predicted by the model.

The finding that people's inferences about agents' desires are more robust than inferences about beliefs, and more consistent with the model's predictions, is intriguingly consistent with classic asymmetries between these two kinds of mental state attributions in the ToM literature. Intentional actions are the joint consequence of an agent's beliefs and desires, but inferences from actions back to beliefs will frequently be more difficult and indirect than inferences about desires.



	BToM	TrueBel	NoObs
Desire	0.90	0.67	0.61
Belief	0.76	0.11	0.39

Figure 3: Scatter plots show overall correlations between BToM model predictions and human judgments about desires and beliefs in our experiment. Each dot corresponds to the mean judgment of subjects in one experimental condition. Magenta circles correspond to trials which had no rational interpretation in terms of POMDP planning. The table shows correlations with human judgments for BToM and two simpler variants, which do not represent beliefs (TrueBel) or do not update beliefs based on observations (NoObs).

Actions often point with salient perceptual cues directly toward an agent's goal or desired state. When a person wants to take a drink, her hand moves clearly *toward* the glass on the table. In contrast, no motion so directly indicates what she believes to be inside the glass. Infants as young as five months can infer agents' goals from their actions (Gergely & Csibra, 2003), while inferences about representational beliefs seem to be present only in rudimentary forms by age one and a half, and in more robust forms only by age 4 (Onishi & Baillargeon, 2005).

Conclusion & Future Work

Our experiment showed that human ToM inferences come surprisingly close to those of an ideal rational model, performing Bayesian inference over beliefs and desires simultaneously. By comparing with two alternative models we showed that it was necessary to perform joint inference about agents' beliefs and desires, and to explicitly model the agent's observational process, as part of modeling people's theory of mind judgments. Crucially, it was also necessary to represent initial uncertainty over both the agent's beliefs and desires.

We have not attempted to distinguish here between agents' general desires and their specific goals or intentions at particular moments of action. In previous work we showed that inferences about which object is most likely to be an agent's instantaneous goal were well explained using a similar Bayesian inverse planning framework (Baker et al., 2009). However, goals are not always about objects. In the present experiments, it feels intuitive to describe agents as attempting to maximize their overall expected utility by adopting a combination of object- and information-seeking goals (or goals intended to update the agent's beliefs). For instance, in Fig. 4, B1 it looks as if the agent initially had a goal of finding out which truck was parked on the other side of the wall, and then after failing to find their preferred truck (M) there, set a goal of returning to the previously observed second-favorite truck

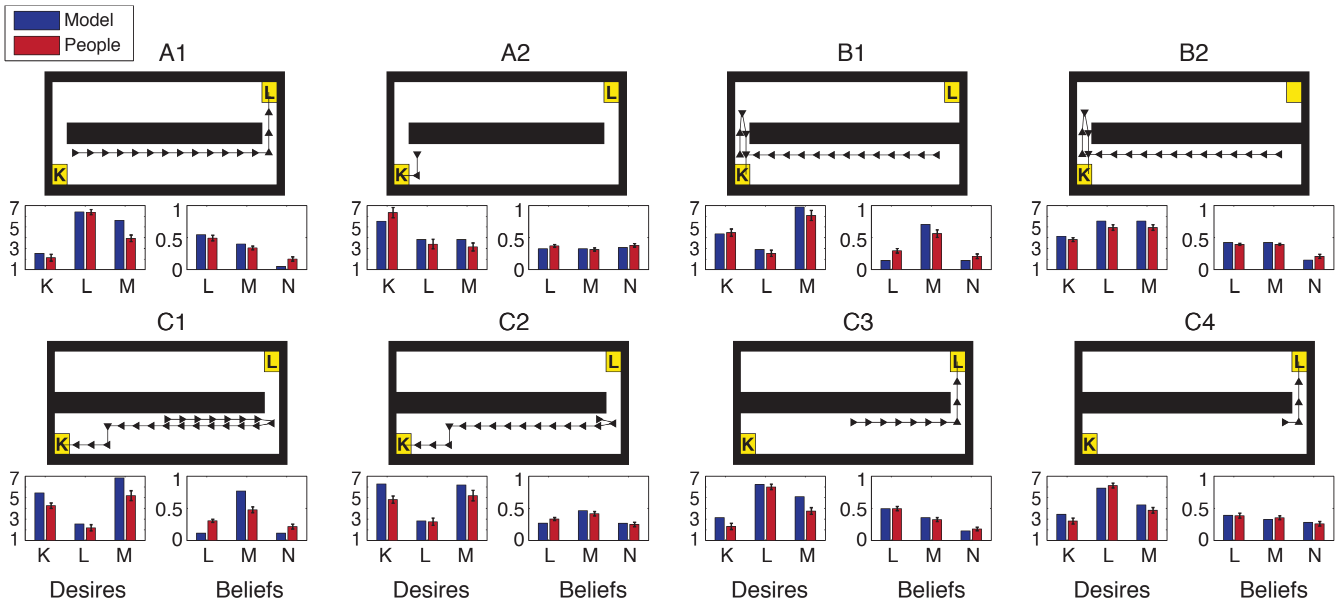


Figure 4: Eight representative scenarios from the experiment, showing the agent's path, BToM model predictions for the agent's desires (for trucks K, L or M, on a scale of 1 to 7) and beliefs about the unseen parking spot (for trucks L, M or no truck (N), normalized to a probability scale from 0 to 1), and mean human judgments for these same mental states. Error bars show standard error (n=16).

(K). Our model can produce and interpret such behavior, but it does so without positing these explicit subgoals or the corresponding parse of the agent's motion into subsequences, each aimed to achieve a specific goal. Extending our model to incorporate a useful intermediate representation of goal sequences is an important direction for future work. Even without these complexities, however, we find it encouraging to see how well we can capture people's joint attributions of beliefs and desires as Bayesian inferences over a simple model of rational agents' planning and belief updating processes.

Acknowledgments This work was supported by the JSMF Causal Learning Collaborative, ONR grant N00014-09-0124 and ARO MURI contract W911NF-08-1-0242. We thank L. Kaelbling, T. Dietterich, N. Goodman, N. Kanwisher, L. Schulz, and H. Gweon for helpful discussions and comments.

References

Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition*, *113*, 329–349.

Bello, P., & Cassimatis, N. L. (2006). Developmental accounts of theory-of-mind acquisition: Achieving clarity via computational cognitive modeling. In *Proceedings of the Twenty-Eighth Annual Conference of the Cognitive Science Society* (pp. 1014–1019).

Bergen, L., Evans, O. R., & Tenenbaum, J. B. (2010). Learning structured preferences. In *Proceedings of the Thirty-Second Annual Conference of the Cognitive Science Society* (pp. 853–858).

Davis, L. S., & Benedikt, M. L. (1979). Computational models of space: Isovists and isovist fields. *Computer Graphics and Image Processing*, *11*, 49–72.

Dennett, D. C. (1987). *The intentional stance*. Cambridge, MA: MIT Press.

Feldman, J., & Tremoulet, P. D. (2008). *The attribution of mental architecture from motion: Towards a computational theory* (Tech. Rep. No. RuCCS TR-87). Department of Psychology, Rutgers University.

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational action. *Trends in Cognitive Sciences*, *7*(7), 287–292.

Gergely, G., Nádasdy, Z., Csibra, G., & Biró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*, 165–193.

Goodman, N. D., Baker, C. L., Bonawitz, E. B., Mansinghka, V. K., Gopnik, A., Wellman, H., et al. (2006). Intuitive theories of mind: A rational approach to false belief. In *Proceedings of the Twenty-Eighth Annual Conference of the Cognitive Science Society* (pp. 1382–1390).

Goodman, N. D., Baker, C. L., & Tenenbaum, J. B. (2009). Cause and intent: Social reasoning in causal learning. In *Proceedings of the Thirty-First Annual Conference of the Cognitive Science Society* (pp. 2759–2764).

Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories*. Cambridge, MA: MIT Press.

Hauskrecht, M. (2000). Value-function approximations for partially observable Markov decision processes. *Journal of Artificial Intelligence Research*, *13*, 33–94.

Lucas, C. G., Griffiths, T. L., Xu, F., & Fawcett, C. (2009). A rational model of preference learning and choice prediction by children. In *Advances in Neural Information Processing Systems 21* (pp. 985–992).

Morariu, V. I., Prasad, V. S. N., & Davis, L. S. (2007). Human activity understanding using visibility context. In *IEEE/RSJ IROS Workshop: From sensors to human spatial concepts (FS2HSC)*.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*(5719), 255–258.

Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: MIT Press.

Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences*, *10*(7), 309–318.

Ullman, T. D., Baker, C. L., Macindoe, O., Evans, O., Goodman, N. D., & Tenenbaum, J. B. (2010). Help or hinder: Bayesian models of social goal inference. In *Advances in Neural Information Processing Systems 22* (pp. 1874–1882).

Wellman, H. M. (1990). *The child's theory of mind*. Cambridge, MA: MIT Press.

Wimmer, H., & 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.

Yoshida, W., Dolan, R. J., & Friston, K. J. (2008). Game theory of mind. *PLoS Computational Biology*, *4*(12), 1–14.

Zettlemoyer, L. S., Milch, B., & Kaelbling, L. P. (2009). Multi-agent filtering with infinitely nested beliefs. In *Advances in Neural Information Processing Systems 21* (pp. 1905–1912).