Cognition 130 (2014) 360–379

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Inferring the intentional states of autonomous virtual agents

Peter C. Pantelis [a,*], Chris L. Baker [b], Steven A. Cholewiak [a], Kevin Sanik [c], Ari Weinstein [c], Chia-Chien Wu [a], Joshua B. Tenenbaum [b], Jacob Feldman [a]

[a] Dept. of Psychology, Center for Cognitive Science, Rutgers University – New Brunswick, United States
[b] Dept. of Brain and Cognitive Sciences, Massachusetts Institute of Technology, United States
[c] Dept. of Computer Science, Rutgers University – New Brunswick, United States

A R T I C L E   I N F O

Article history:
Received 16 February 2013
Revised 16 September 2013
Accepted 13 November 2013

Keywords:
Theory of mind
Action understanding
Intention inference

A B S T R A C T

Inferring the mental states of other agents, including their goals and intentions, is a central problem in cognition. A critical aspect of this problem is that one cannot observe mental states directly, but must infer them from observable actions. To study the computational mechanisms underlying this inference, we created a two-dimensional virtual environment populated by autonomous agents with independent cognitive architectures. These agents navigate the environment, collecting "food" and interacting with one another. The agents' behavior is modulated by a small number of distinct goal states: *attacking*, *exploring*, *fleeing*, and *gathering food*. We studied subjects' ability to detect and classify the agents' continually changing goal states on the basis of their motions and interactions. Although the programmed ground truth goal state is not directly observable, subjects' responses showed both high validity (correlation with this ground truth) and high reliability (correlation with one another). We present a Bayesian model of the inference of goal states, and find that it accounts for subjects' responses better than alternative models. Although the model is fit to the actual programmed states of the agents, and not to subjects' responses, its output actually conforms better to subjects' responses than to the ground truth goal state of the agents.

1. Introduction

Comprehension of the goals and intentions of others is an essential aspect of cognition. Motion can be an especially important cue to intention, as vividly illustrated by a famous short film by Heider and Simmel (1944). The "cast" of this film consists only of two triangles and a circle, but the motions of these simple geometrical figures are almost universally interpreted in terms of dramatic narrative. Indeed, it is practically impossible to understand many naturally occurring motions without comprehending the intentions that contribute to them: a person running is interpreted as trying to get somewhere; a hand lifting a Coke can is automatically understood as a person intending to raise the can, not simply as two objects moving upwards together (Mann, Jepson, & Siskind, 1997). Much of the most behaviorally important motion in a natural environment is produced by other agents and reflects unseen mental processes. But the computational mechanisms underlying the inference of mental states, including goals and intentions, are still poorly understood.

Human subjects readily attribute mentality and goal-directedness to moving objects as a function of properties of their motion (Tremoulet & Feldman, 2000), and are particularly influenced by how that motion seems to relate to the motion of other agents and objects in the environment (Blythe, Todd, Miller, & The ABC Research Group, 1999; Barrett, Todd, Miller, & Blythe, 2005; Tremoulet & Feldman, 2006; Zacks, Kumar, Abrams, & Mehta, 2009; Gao, McCarthy, & Scholl, 2010; Pantelis & Feldman, 2012). The broad problem of attributing mentality to others has received a great deal of attention in the philosophical literature (often under the term *mindreading*), and has been most widely

* Corresponding author. Address: Indiana University – Bloomington, Dept. of Psychological and Brain Sciences, 1101 E. 10th Street, Bloomington, IN 47405, United States. Tel.: +1 812 856 7800.
E-mail address: pcpantel@indiana.edu (P.C. Pantelis).

0010-0277/$ - see front matter
http://dx.doi.org/10.1016/j.cognition.2013.11.011

studied in infants and children (Gelman, Durgin, & Kaufman, 1995; Gergely, Nádasdy, Csibra, & Bíró, 1995; Johnson, 2000; Kuhlmeier, Wynn, & Bloom, 2003). But the adult capacity to understand animate motion in terms of intelligent behavior has been less studied. Computational approaches to the problem of intention estimation have been scarce historically (for perhaps the earliest example, see Thibadeau, 1986), in part because of the difficulty in specifying the problem in computational terms. But new modeling approaches are emerging from various perspectives and disciplines in this rapidly-developing area of research (Feldman & Tremoulet, 2008; Baker, Saxe, & Tenenbaum, 2009; Crick & Scassellati, 2010; Kerr & Cohen, 2010; Pautler, Koenig, Quek, & Ortony, 2011; Burgos-Artizzu, Dollár, Lin, Anderson, & Perona, 2012).

Experimental stimuli in studies of the interpretation of intentionality from motion have, like the original Heider and Simmel movie, consisted almost exclusively of animations featuring motions crafted by the experimenters or their subjects to convey specific psychological impressions. Traditional psychophysics is then applied to relate attributes of the observed motion to the subjective impression produced (Blythe et al., 1999; McAleer & Pollick, 2008). While this method has yielded important insights, it suffers from certain critical limitations. Apart from the inefficiency of continual reliance on subjective intuition (e.g. via a subject pool) to generate new and varied stimuli scenes, handcrafted stimuli are opaque in that it is unclear exactly *why* the constituent motions convey the particular impressions they do, since they have been designed purely on the basis of the designers' intuitions—intuitions that are, in effect, the object of study. This makes it impossible to explore, for example, the relationship between observers' judgments of the agents' mental states and the true nature of the "mental" processes generating agent behavior. In this case, the independent and dependent variables are both direct reflections of subjective notions of what particular classes of behavior "should" look like.

Other studies have examined the perception of animate motion more systematically, either by varying the velocity and orientation of agents parametrically, or by manipulating parameters of simple programs generating agent behavior (Stewart, 1982; Dittrich & Lea, 1994; Williams, 2000; Tremoulet & Feldman, 2000, 2006; Gao, Newman, & Scholl, 2009; Gao & Scholl, 2011; Pantelis & Feldman, 2012). While this method avoids some of the aforementioned pitfalls of using handcrafted stimuli, our present study represents a substantial departure even from this approach. In the spirit of Dennett (1978)'s suggestion to "build the whole iguana," our goal was to create cognitively autonomous agents whose motions actually were, at least in a limited sense, driven by their own beliefs, intentions, and goals. To this end, we developed a 2D virtual environment populated with autonomous agents—virtual robots—who locomote about the environment under their own autonomous control, interacting with and competing with other agents in the environment. We refer to the agents as IMPs, for *I*ndependent *M*obile *P*ersonalities. Like agents in artificial life environments (e.g. Yaeger, 1994; Shao & Terzopoulos, 2007), IMPs have a complete, albeit severely restricted, cognitive architecture.

The IMPs can be understood to have one overall goal: to obtain "food" and bring it back to a home location. But at each time step, an IMP's behavior is modulated by its continually-updating "goal" state, which determines how it will respond to stimuli in the environment. An IMP can be in one of four discrete goal states: it can **explore** the environment, **gather** food, **attack** another agent, or **flee** from another agent (Fig. 2. These four states were loosely modeled on the "Four Fs" of animal ethology, action categories that are said to drive most animal behavior; see Pribram, 1960).

The agents obtain information about their environment via on-board perception, consisting of a simple visual module with a 1D retina (a perceptual ability reminiscent of that of the 2D characters in Abbott's (1884) novella *Flatland*). The agents progressively learn a map of their environment as they move about the environment. Lastly, the agents have a limited capacity to reason about how to accomplish their goals (for example, they can calculate the shortest path through the environment between their current location and a goal location). Thus the IMPs are complete, though crude, cognitive agents. Their observable actions are based entirely on what they "want", "know", and "think" about their environment.

The subjective appearance of IMP behaviors corresponding to their respective goal states are necessarily connected to the subjective intuitions of the programmers, but this connection is far more indirect than in the case of stimuli created via handcrafted animation. We can hardly predict how stimulus scenes will appear with any precision, given that the IMP subroutines connected with respective goal states execute within the complicated context of other modules in the IMP programming, and that these scenes are dynamically generated within multi-agent environments which are explicitly probabilistic. Manipulation of the parameters of particular IMP modules may have inherently unpredictable effects; for example, we have no strong and precise intuition about what it would "look like" if the resolution of an agent's vision were changed. Finally, whereas the semantics we attach to the various IMP goal states may be arbitrary and subjective (why "attack" and not "chase"?[1]), there is an objective ground truth to the existence of behavioral states contained in the IMP program, among which the IMP actually transitions, and each of which predisposes the IMP toward particular actions. There is therefore a ground truth basis for assessing subject's accuracy when they attempt to infer these underlying states.

In the studies below, we ask what human subjects can infer about the IMPs' goal states on the basis of observing them move about and interact within a sparse environment, and model how they might go about performing this inference. The appearance and behavioral repertoire of the IMPs are quite simple; they are rigid triangles which may only translate or rotate. This does not mean to imply that the perceptual features of these IMP stimuli exhaust the possibly important cues subjects may use to make

[1] For more on the semantics subjects attach to IMP behavior without being first supplied with our labels, see Pantelis et al. (2011).

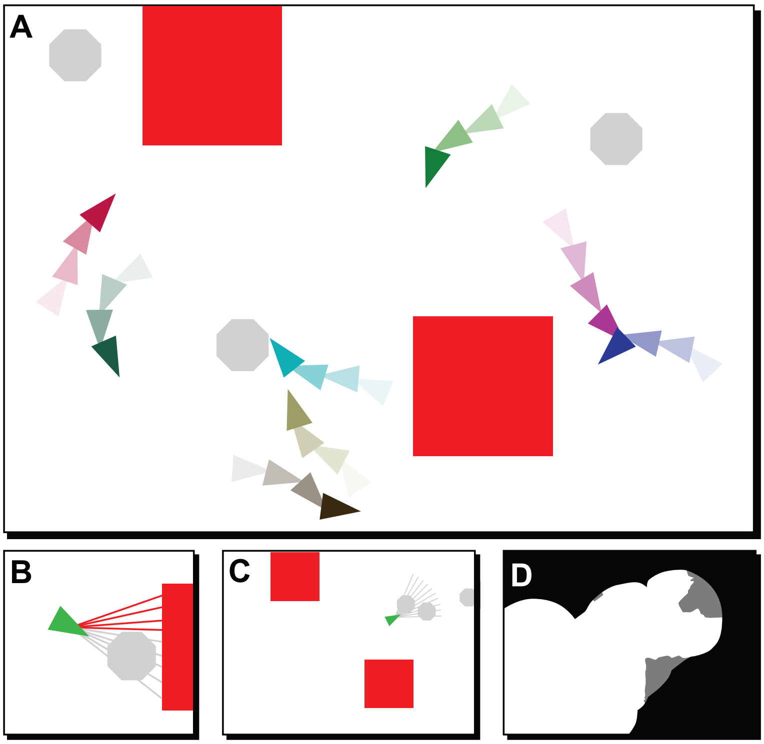


Fig. 1. (a) The virtual environment with its native IMPs (depicted as moving triangles). (b) The IMPs have autonomous vision by virtue of simulated 1d retinas, and by (c) exploring their environment they can (d) gradually develop a mental map of the objects and obstacles it contains.

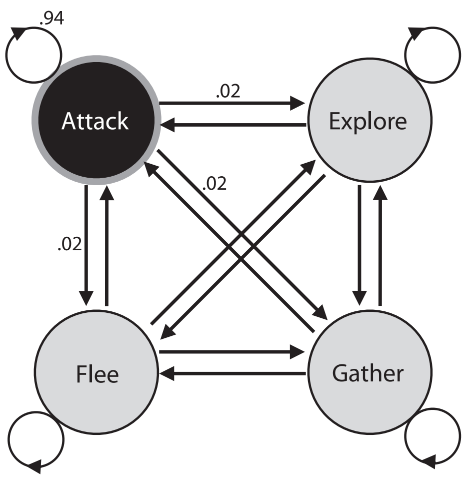


Fig. 2. The IMP's decision making module is programmed such that the IMP will transition stochastically among the four possible goal states (*attack*, *explore*, *flee* and *gather*) according to transition probability tables contained in its program (see Tables 7 and 8 in the Appendix for these transition tables). The IMP applies one of four distinct transition tables depending on its current circumstance (*nothing nearby*, *another IMP nearby*, *food nearby*, or *another IMP and food nearby*). This figure illustrates the decision making of an IMP from Experiment 1, with no other IMPs or food nearby, currently in the *attack* state. Given this situation, the IMP would remain in the *attack* state with probability .94, and transition to each of the other goal states with probability .02.

inferences about other agents (Gelman et al., 1995). Rather, we isolate these cues for experimental study.

Again, the IMPs' goal states are not directly observable, but are internal variables that determine how they respond to what they themselves perceive in the environment around them. Thus, our main question is really about the capacity of our human subjects to represent that which, in turn, represents: a mind. Traditional psychophysics concerns itself with the relationship between physical variables (e.g. luminance or sound amplitude) and their mental correlates (e.g. perceived lightness and loudness). Our paradigm can, similarly, shed light on the relationship between hidden processes generating observable actions (like the tendency of an agent to transition into an "attack" state) and their psychological correlates (the subjective impression of intentions). In this sense, we see our paradigm as a true "psychophysics of intention."

The idea of using autonomous virtual agents as psychophysical stimuli was previously explored in Pantelis and Feldman (2012). In that study, stimulus scenes were populated with simple reflex agents which differed in their behavioral tendencies (the way they reacted to other agents) but lacked perception, memory, or decision-making. The goal of that study was to use a parameterized space of behavioral tendencies as a way of generating agents with a wide range of "personalities," in order to map out subjects' subjective personality space (via multidimensional scaling). The current study has more ambitious aims, and the agents have a far more complex mental architecture. The IMPs environment is a setting for interactive intentional behavior, and affords many possibilities for the empirical study of intention perception.

For example, by modifying the IMPs' programming we can completely control the agents' hidden cognitive and perceptual capacities (e.g. their vision, memory, or behavioral dispositions) or the influence of these capacities on observable agent behavior. This allows us to study how modifying any of these capacities, or even deleting them entirely, might influence the way observers understand their mental properties.

In the current study we focus on one particularly central aspect of the computation of intention: human observers' ability to infer the "mental state" of agents on the basis of their actions—that is, in our paradigm, to correctly ascertain which of the four predefined goal states an agent is in at each point in time. We recognize the need for caution in referring to the four IMPs goal states as "mental states," as the IMPs' perceptual and cognitive capabilities are obviously very limited (see fuller description in A). The IMPs' various modules are relatively simplistic models of vision, memory, decision making, and path planning, and do not precisely correlate with the manner by which humans or other animals perform these tasks. The four IMP goal states are simply decision structures that determine how the agent's responses are conditioned on what it perceives and knows about its environment, and we do not mean to imply that these states and the simple decision matrices governing transitions among them are in any real sense equivalent to real human mental states and decision making, any more than we wish to imply that actual human vision works via the casting of rays from the eye.[2] Nevertheless, we adopt this phrasing deliberately, because in the context of our study, the IMPs' states play an analogous role to intentional mental states: they control the selection of action given the knowledge and perception accessible to the agent. They are "behavioral dispositions" in the very literal sense that they are internal characteristics that modulate the probability of behavior, and in this very concrete sense are loose analogs of the more complex intentional dispositions that govern human behavior.

These studies represent a departure from traditional psychophysics methods, in which physical attributes of a stimuli agent's trajectory are manipulated and the resulting subjective percept is studied. We instead manipulate hidden internal states of agents and parameters of the generative model with which the agents are programmed, which influences agent behavior indirectly. Inferences made about this hidden information define a computational process which, we argue, is in closer analogy to the actual inferential processes comprising a "theory of mind." This formalization of intention inference has also been explored with a virtual environment and classification algorithm developed by Kerr and Cohen (2010).

In the two studies below, we ask subjects to observe four IMPs interacting, one of which is designated as the target, and continually indicate using the computer keyboard what state they think the agent is in at each point in time. In effect, we asked in as direct a manner as possible whether the subjects could correctly divine the agent's internal state on the basis of its actions. Because this state is in fact simply a variable inside the agent's autonomous program (the "ground truth" goal state) we were then able to analyze how often, and under what circumstances, the subject's response was in fact correct (validity, in traditional statistical terminology), as well as how often subjects agreed with one another (reliability). We then introduce a computational model of the inference process, ask how often and under what circumstances it is able to estimate the true state, and evaluate how effectively it models subjects' responses. Most of these analyses (with the exception of reliability) are impossible using handcrafted displays, because the agents in such displays have no ground truth mental states.

[2] This was an early and incorrect theory of vision held by Plato, Euclid, and others. It is also how we model the IMPs' vision.

2. Computational model

If one wishes to attribute goals to an agent effectively, it is useful to have a good model of that agent. From an "ideal observer" (Knill & Richards, 1996; Geisler, 2004) or "rational analysis" (Anderson, 1989) perspective, the optimal solution to the goal attribution problem indeed relies crucially upon an accurate model of the agent's goal- and environment-dependent behavior. In the IMPs domain, such a model would express how an IMP's action A depends probabilistically on its goal G and the state of the environment S:

$$p(A|G,S). \tag{1}$$

Given the observed action, and this model of how the agent generates its behavior, the ideal observer works backwards to reason about the agent's underlying goal. The inference performed by the observer can be expressed as a problem of computing the posterior probability of the goal G by inverting the generative model using Bayes' rule:

$$p(G|A,S) \propto p(A|G,S)p(G|S). \tag{2}$$

This ideal observer approach to goal attribution has found past success when applied to scenarios involving simple, two-dimensional "grid-world" environments and restricted sets of possible goals and behaviors (Baker et al., 2009; Ullman et al., 2009). In these limited contexts, a wide range of natural and intuitive behaviors can be modeled as the result of rational "planning" in a Markov decision process (MDP), and the process of inverting this model using Bayesian inference can be called "inverse planning."

As is true with any model of a complicated reality, a useful model of the agent function will be compressive in nature. And because rationality is a powerful form of compression, the assumption of a rational agent can be (and has been) an exceedingly useful starting point when inferring beliefs, goals, and intentions. However, a rational model of our IMPs' behavior is not particularly well-defined, and even if it were, the computations required for optimal planning would be intractable given the continuous state, multi-agent, partially observable MDPs required to express the IMPs' domain (Kaelbling, Littman, & Cassandra, 1998).

On the other hand, the actual model generating IMP behavior exists (it was used to generate the stimuli) and makes sense (it accurately and precisely describes how IMP behavior depends on the underlying goal state and the state of the environment). If granted access to this complete generative model, one could therefore use Bayesian inference to optimally estimate the IMP's goal state, given the observed action. These considerations make the complete generative model of the IMPs' mental architecture an attractive basis for modeling the "true" Bayesian ideal observer.

Nevertheless, we do not adhere to this ideal observer approach. A complete generative model of the IMPs' mental architecture would necessarily include all perception, memory, and decision making processes undertaken by the IMP. It is doubtful that human observers actually harbor a full generative model of the observed agent—an analog of the computer program for generating the IMPs' behavior—in their heads, or that they observe sufficient data throughout the course of the experiment to induce this program. And even if subjects did have access to the true IMP generative model, the computations required for full Bayesian inference over this model are themselves intractable, due to the complexity of the space of the IMPs' potential mental states, actions, and physical configurations.

Still, it is possible to apply Bayesian reasoning over a model of agents' goal- and context-dependent behavior, using an *approximate* model of IMP behavior rather than the true underlying generative model. This model need not represent with perfect fidelity all aspects of the agent's actions, the scene, or the mapping from scenes and goals to actions. It need only capture the key features of the IMPs' situations, goals, actions, and the structure of the dynamic relationships between them that are necessary to support inferences that are accurate and precise enough for everyday social functioning. This approach also has the potential to yield tractable computations that can be performed efficiently and dynamically as the stream of perceptual input arrives.

We formalize this idea by constructing a dynamic Bayesian network (DBN), shown in Fig. 3, which represents observers' knowledge about the probabilistic, temporal dependencies between the IMPs' states (configuration of agents, food, and obstacles in the environment), goals and actions. To compress the IMPs' continuous, multidimensional state and action spaces, this DBN represents the IMPs' activities at a more abstract level by chunking similar states and actions into semantically coherent categories. The probabilistic relationships represented in the DBN are learned from prior experience; specifically, we use the actual generative model of our IMPs to produce repeated observations of IMPs' activities. These simulations serve as data for supervised learning.

The DBN employed by our model is only meant to be an approximation of the internal model of IMP behavior that may be harbored by subjects, and we must note that robustly modeling other sorts of agents—like humans—may very likely require a richer representation to be effective. Also, when compared to the true generative model governing IMP behavior, our approximate model is coarser in resolution with respect to states and actions, and therefore will appear noisier. This is not an uncommon outcome for Bayesian models: the gap between the true generative process (which can, in principle, be deterministic if all the relevant variables are known) and the modeled generative process becomes absorbed into the model's noise or stochasticity.

Next, we provide technical details about our representation, learning, and inference procedures, or in other words, how the DBN is first constructed and then used to reason about the IMPs' goal-directed behaviors.

2.1. Approximate representation of states and actions for the Bayesian model

We first simplify the set of possible IMP actions. Although the actual IMPs can take on various speeds and angular velocities (but never pause), the model classifies all IMP actions as either *turn left*, *turn right*, or *move straight ahead*.

Second, we simplify the set of possible environmental configurations. We coarsely discretize the agent-centric physical space into 9 sections (Fig. 4). The nearest other IMP and nearest food to the target agent can each lie in one of these 9 sections; thus, our discretization scheme allows for $9 \times 9 = 81$ possible configurations of the

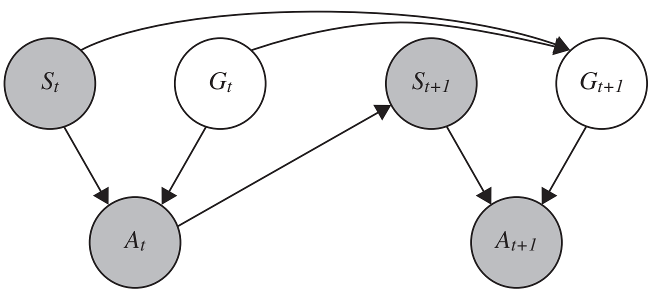


Fig. 3. Graphical model of the Dynamic Bayesian Network with which we *approximate* the actual generative model of the IMP. S represents the environmental configuration, or *s*cene, G represents the *g*oal state of the IMP, and A represents the IMP *a*ction. The probability of the IMP transitioning to G_{t+1} is conditional on both S_t and G_t; A_t depends on S_t and G_t. These probabilistic relationships are learned via extensive simulation of IMP behavior. The action taken by the IMP at time t (A_t) will also causally influence the subsequent configuration of the environment at time $t+1$, (S_{t+1}), which in turn influences A_{t+1}, but these relationships are not learned or represented explicitly by the computational model because S_{t+1} and A_{t+1} are both given (i.e. inputs known to the inferential algorithm) at time $t+1$.

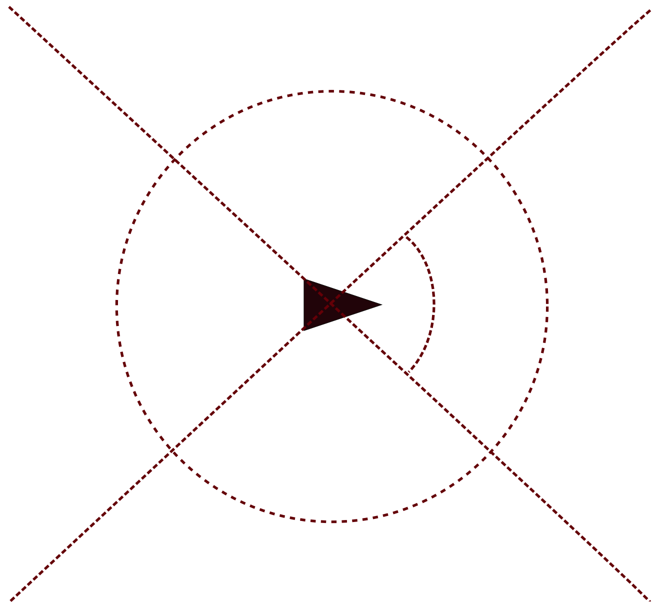


Fig. 4. The discretization of the agent-centric space, used for the Bayesian model.

environment (with respect only to the relative locations of the nearest other IMP and nearest food).

It should be noted that our subjects may very well rely on a qualitatively different discretization scheme, or may not rely on any such discretization at all. We also do not mean to imply that this representational scheme is somehow globally optimal (though subsequent model fitting demonstrates that this representational scheme is adequate for the task of modeling a large portion of subject responses). As is often the case, the particulars of the model are both the product of *a priori* considerations and subsequent tinkering. For example, we presumed that a coordinate frame centered at the IMP would be consistently more relevant for this class of inference than one centered at the origin of the 2D simulation environment. Further, the small section carved into the agent-centric space directly in front of the agent was included as a possible location of food or another IMP (Fig. 4) due to *a priori* intuitions that subjects would find this section to be particularly salient perceptually, and knowledge that an agent "carrying" food would typically be positioned with this food at a consistent distance from the tip of its "nose." After implementing these decisions, we could have carved the remaining agent-centric space into sections of any arbitrary number, size, shape, and location. The coarsest possible discretization of the continuous environment (and the objects it contains) would have been achieved by allowing it to take on only one of two values—for instance, "something near the IMP" or "nothing near the IMP". We added some minimal expressive power to the representation of the environment by allowing a distinction of whether that "something" is an IMP or food, and dividing the agent-centric space, very simply, into far and near sections (with a circle) and sections to the left, right, front, and back of the IMP (with two lines intersecting at the centroid of the IMP).

2.2. Learning phase

In order to empirically determine this conditional probability of an action at time $t, P(A_t|G_t, S_t)$, we sample the generative model by running a large number of simulated IMPs environments. When an IMP is in a particular goal state, in the context of a particular environmental configuration, the selected action is tabulated. Eventually, this learning process yields a table that approximates the IMP's "policy" (conditional probability of an action) given any goal state/environment combination.

The conditional probabilities of an IMP transitioning among the four goal states, $P(G_{t+1}|G_t, S_t)$, also must be learned through this sampling process. The stochasticity in the actual generative model governing agent behavior reflects the programmed probabilities of an IMP transitioning among the various goal states. But the DBN is nondeterministic because the person observing the behavior (the subject) has uncertainty about the beliefs of the agent—i.e. what the agent perceives or remembers about its environment at any given time point.

2.3. Inference phase

Once the parameters of the DBN have been learned, inference of the IMPs' goal states, given their observed actions, can be performed. We model the observer's inference by computing the marginal posterior probability of a goal at time t, given the state and action sequence up to that point:

$$P(G_t|S_{1:t}, A_{1:t}) \propto P(A_t|G_t, S_t)\sum_{G_{t-1}} P(G_t|G_{t-1}, S_{t-1}). \quad (3)$$

At $t = 1$, each goal state is believed to be equally likely. At each subsequent time step, this computation yields a probability distribution across the four possible goal states which integrates the probability of changing goals with the likelihood that each goal produced the observed actions.

2.4. Free parameters of the Bayesian model

Only three free parameters are fit with respect to experimental subjects' responses: two for discretizing the representation of the agent-centric environment, and one response "lag" parameter (rather than using the model output for the point in time aligning precisely with the subject's response, a trailing average of the model's outputs is taken which goes back an amount of time determined by this parameter). For comparison, a very simple statistical model, only learning the base rates of subjects' responses across the four response types (attack, explore, flee, and gather) and attempting to predict subjects' responses on this basis, would fit the very same number of free parameters.

2.5. Alternative models

Our Bayesian model explicitly represents the IMP states and the possible transitions among them. Because this model parses the dynamic scene, integrating perceptual cues associated with respective states at a given time with prior estimates of the relative plausibility of transitions among states, it is able to integrate information across time in a more powerful manner than would be enabled by a simpler feature- or cue-based account (such as multinomial logistic regression).

To provide a quantitative comparison between our dynamic, model-based account and a cue-based alternative, we construct a large family of logistic regression models and compare their performance to that of our Bayesian model, in both experiments. We expect both modeling approaches to fit subjects' classification judgments reasonably well, but if the Bayesian model is superior it should be able to fit these data with greater model parsimony (fewer free parameters) and, relatedly, should generalize better across settings.

3. Experiment 1

The first experiment tested subjects ability to successfully categorize the IMPs' behaviors and detect transitions among the IMPs' goal states. The four possible underlying states were explained transparently to the subjects during an initial training phase.

3.1. Methods

3.1.1. Subjects

Twelve undergraduate students in introductory psychology classes at Rutgers University participated in the experiment, and received course credit for their participation. Two additional subjects' data were excluded due to failure to follow experimental instructions (the subject did not respond during entire experimental trials, or pressed inappropriate keys). Each experimental session lasted approximately 30 min.

3.1.2. Stimuli

Each subject viewed the same set of 20 scenes, generated in advance. Each pre-recorded scene was 60 s in duration, and was presented within a 400 $\times$ 400 pixel window, horizontally subtending approximately 13.5° of visual angle. Each scene was populated with 4 identically parameterized IMPs at randomized starting positions, 15 gray food objects (divided evenly into 3 clusters, with each cluster initially placed at a random starting position), and two square red obstacles (placed at the same locations in each scene). A fuller description of the virtual environment and the programming of the IMPs can be found in A, and example scenes can be viewed at http://www.indiana.edu/~brainlab/pantelis.

3.1.3. Procedure

Five initial training scenes were shown. Subjects were instructed to simply observe the action and try to determine what was happening within the scenes. During training, each IMP's true goal state was reflected in its color (see Fig. 5). After subjects watched these 5 scenes, they were asked what they thought the IMPs were doing, and what the colors might mean. It was then explained to them that the colors actually corresponded to the underlying mental or behavioral state of the IMP, and that an IMP could be in one of four of these states at a given time: "attacking" another agent, "exploring" its environment, "fleeing" from another agent, or attempting to "gather" food.

Each subject then viewed 15 additional scenes, the first of which was treated as practice and excluded from analysis. In these scenes, IMPs did not change color; that is, the subjects' task was to infer the underlying state of an IMP solely from its behavior and context. The target IMP was colored black, and the other 3 were colored blue. Subjects were instructed to pay attention to the black agent in each scene, and indicate on the keyboard which state they thought this target agent was in at any given time. Four keys represented the 4 respective possible states; subjects were instructed to press a key as soon after a scene began as possible, and thereafter to press a key only when they thought the target IMP had transitioned into a new state. Subjects each viewed the same 20 total scenes, and in the same order.

3.2. Behavioral results

Fig. 6 illustrates how subjects responded at they observed the 14 test scenes. The "ground truth," programmed goal state of the target IMP is shown in the top horizontal bar for each scene. The proportion of subjects' responses across the four response types is shown in the middle row.

We first examined subjects' performance by measuring the proportion of time that their classifications matched the ground truth state of the target IMP (validity; see Table 1). Mean accuracy was 48%, approximately twice chance performance.

Another critical aspect of subject performance is intersubject agreement (reliability). Excluding portions of trials for which the most common response was "none yet given" (a response category represented by black in

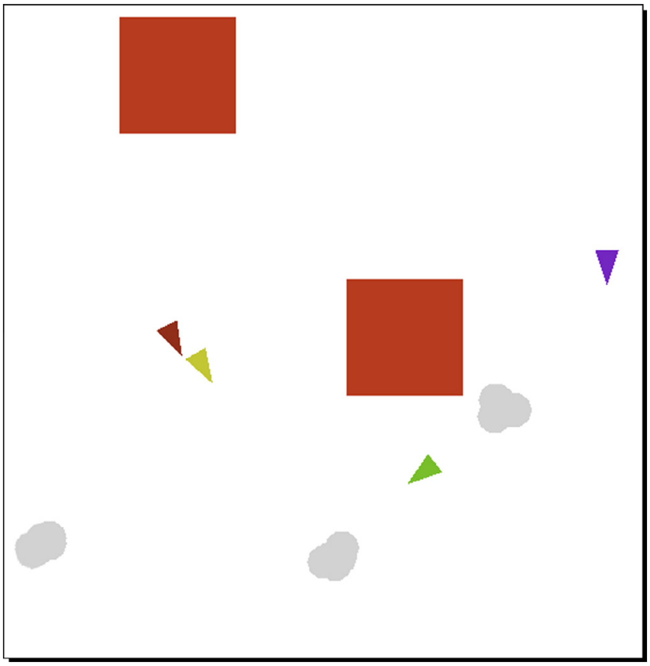


Fig. 5. A frame from a sample scene viewed by subjects. The red IMP is in an "attack" state, the purple IMP is "exploring," the yellow IMP is "fleeing," and the green IMP is "gathering." Note that colors were only shown during training scenes. During the remainder of the experiment, the target IMP was colored black, and the other IMPs were colored blue. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 6), an average of 8.7 out of 12 subjects (73%) agreed upon the mode response at any given time.

A comparison of estimated goal states to actual ones shows a number of interesting patterns, as illustrated by the inter-state confusion matrix (Table 2). The analysis reveals one dominant source of subject "errors." Subjects generally did not initiate responding immediately at the start of each trial; 13% of overall trial time was prior to the initial response. As IMPs were most likely to be in the explore state at the beginnings of trials, these errors of omission account for a large proportion of subjects' misclassifications for this action type. Otherwise, subjects' detection of the explore state was 79%. Accuracy was lower for the other states. For example, when an IMP was in the flee state, subjects were actually most likely to respond *attack* or *explore*, and the hit rate for *flee* was only 10%.

Subjects' response rates across the four types were well-calibrated to the actual time the IMPs spent in each state: Subjects' responded *explore* most frequently, followed by *gather*, *attack*, and *flee*. Overall, subjects tended to slightly overestimate *explore* at the expense of *gather*, which was slightly underestimated.

3.3. Bayesian model performance

3.3.1. KL divergence

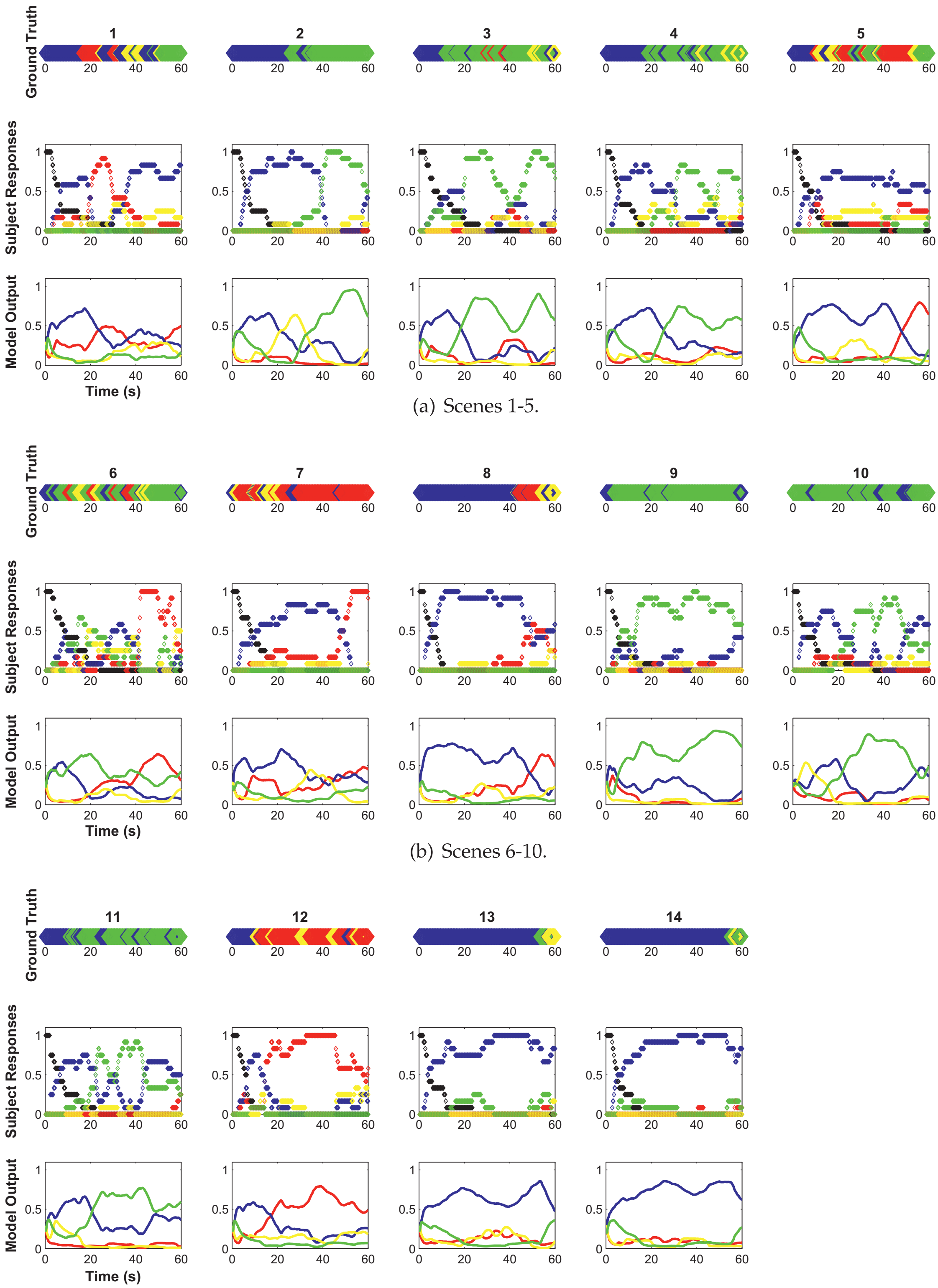
The Bayesian model outputs a posterior distribution across the four possible response types. We consider a model to be a good fit if this distribution matches well with the distribution of the (12) subjects' responses across these four types. For this reason, we use Kullback–Leibler (KL) divergence as our model performance metric.

KL divergence is a non-symmetric measure relating two probability distributions. If M is the model's output distribution, and S is the subjects' response distribution, then the KL divergence is the number of extra bits required to encode S using M instead of S. Thus, a lower KL divergence represents a better fit, with a minimum possible KL of 0 indicating that the two distributions are exact matches, and a maximum possible KL being arbitrarily large, depending on the smoothing parameter (ϵ) inserted into the model distribution in lieu of zero values.

As a baseline, a "null" model—believing the agent to always be equally likely to be in any of the four goal states (attack, explore, flee, gather)—would fit subjects' responses (on average) at KL = .863. A slightly less naïve model, which knows the distribution of subject responses (see Table 2) and believes the probability of agent being in the four respective goal states to always be in proportion to these empirically determined response rates, fits subjects' responses at KL = .630.

3.3.2. Fitting and evaluating the model

As illustrated in Fig. 6, the posterior distribution (output) of our Bayesian model across the four response types matches quite well with the distribution of subjects' responses. We fit the model's three free parameters using KL divergence as our performance metric. Because assessment of one particular configuration of parameters (while recruiting, in this case, $\sim$1200 simulated scenes) is computationally inefficient, and a precisely optimal setting of these parameters—fitting the data marginally better—is unnecessary for demonstrating the efficacy of the model with respect to general claims we make about it, we



(a) Scenes 1-5.

(b) Scenes 6-10.

(c) Scenes 11-14.

Fig. 6. Experiment 1. This figure shows, over time, the underlying "ground truth" state of the agent (top row for each scene), the distribution of subject responses (middle row), and the output distribution of the Bayesian model (bottom row). *Red* represents the "attack" state, *blue* = "explore," *yellow* = "flee," and *green* = "gather." For the subjects' responses, *black* indicates the proportion of subjects who had not yet responded on the keyboard during a given trial. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Accuracy of subjects and model, with respect to ground truth goal state.

	Scene														All
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
Subjects	.32	.63	.45	.46	.17	.14	.32	.68	.68	.36	.44	.55	.73	.76	.48
Model	.34	.56	.51	.47	.25	.31	.27	.56	.67	.45	.50	.47	.57	.67	.47

Table 2
Confusion matrix for subjects' responses in Exp. 1 (averaged across subjects). Mean proportion of IMP time spent in each state is in parentheses, and mean proportion of time subjects spent in each response category is at the bottom of each column.

Actual state	Subject response				
	None	Attack	Explore	Flee	Gather
Attack (.16)	.05	.42	.38	.11	.04
Explore (.39)	.25	.04	.60	.03	.08
Flee (.08)	.07	.28	.38	.10	.17
Gather (.37)	.07	.07	.32	.08	.46
	.13	.13	.44	.07	.22

coarsely discretize the parameter space. Radius was allowed to take on values of 50, 70, 100, 130, 160, or 190 pixels. Angle was allowed to take on values of $\pi/6, \pi/3, \pi/2, 2\pi/3$, or $5\pi/6$ radians.[3]

We found that the model fit best when discretizing the space with a 100 pixel radius and an angle of $5\pi/6$ radians. The model also fit subjects' responses best if, rather than using the model output at the particular point in time aligning with the subject's response, a trailing average of the model's outputs—going back up to 12 s—was used. Using a trailing average (rather than taking the model's prediction at a given quarter second timeslice[4]) helps to accommodate inertia and reaction time lag in subjects' responses, and tends to naturally smooth out artificial discontinuities resulting from the harsh discretization of the agent-centric space. Using this model, with its best fitting free parameters (100 pixels, $5\pi/6$ radians, 12 s), resulted in an average KL divergence (with respect to subjects' distribution of responses) of .334.[5]

We additionally fit a multinomial logistic regression model, predicting subjects' responses from several underlying variables at any given time. This is a discriminative, statistical approach that lacks the temporal component of our dynamical Bayesian model, and is reminiscent of the various cue-based categorization approaches tested in Blythe et al. (1999).[6]

The following variables enter into the regression model as continuous input variables: relative angle of nearest other agent, distance to nearest other agent, relative angle of nearest food, distance to nearest food, agent turning velocity, and agent speed. This is the same information with which we endow the our Bayesian model, although in finer resolution—for input to our model, all of these variables are highly discretized. For example, whereas the Bayesian model only knows whether the target IMP is turning left, right, or moving straight ahead, the regression model has access to its precise turning velocity.

In order to prevent overfitting this model to the data set, we applied a cross-validation procedure. Because there were 6 potential input variables to be included in the regression, there were 63 possible combinations of variables, and therefore 63 candidate regression models to test. For the cross-validation procedure, for each candidate model, the data set was first split randomly into a training set and a testing set (we used 25%/75%, 50%/50%, and 75%/25% splits). The candidate model was then fit to the training set and assessed with the testing set, using percent of subject responses correctly predicted as the performance measure. This procedure was repeated 10 times for each model, and the results were averaged to provide an assessment of the generalizability of the candidate model.

Individual models performed similarly across the three training/testing split conditions. Several regression models performed about equally well; we selected the model that generalized to test sets best, on average, across all three. This model employed four input variables: relative angle of nearest other agent, distance to nearest other agent, distance to nearest food, and agent turning velocity.

This regression model is thus not a straw man, but a fair and robust treatment of this approach. However, the regression model includes 15 free parameters, compared to the 3 free parameters employed by our model. To compensate for this difference in model parsimony, we calculated the Akaike information criterion (AIC) for both models. We then computed the difference in adjusted log likelihood (see Burnham & Anderson, 2002), which expresses the relative fit of the Bayesian and logistic models after adjusting for the number of fitted parameters. Fig. 9 shows this difference over time for each of the 14 scenes (adjusted log likelihood values over zero favor the Bayesian

[3] π would be the maximum setting for this parameter, and would divide the agent-centric space into two semicircular sections of equal size instead of four sections.

[4] The model was assessed at each quarter second because each second of animation contained at least four frames before linear interpolation (see A), and the analysis was performed on this raw frames.

[5] Because parameter fitting selected an extreme value of $5\pi/6$ radians, we tested several values further in this direction, between $5\pi/6$ and π. Increasing this angle parameter slightly beyond this setting resulted in very marginal improvement of fit (KL = .333). Increasing this parameter still further made the fit marginally worse, with KL = .335 at the maximum parameter setting of π. We therefore conclude that $5\pi/6$ is an approximately optimal parameter setting.

[6] The nature of the modeled categorization task, in this case, also differs from that of Blythe et al. (1999) in that subjects categorized entire agent trajectories in that study, whereas in our study subjects infer the internal state of an agent which may change several times over the course of a scene.

model). As can be seen in the figure, the Bayesian model is nearly always preferred.

A benefit of the regression approach was that we could assess the diagnostic value of these input variables when taken in isolation (Barrett et al., 2005). Table 5 shows how well the various parameters predicted subjects' responses; this is a class of simple models, each employing 6 free parameters fit to subjects' data. Using KL divergence again as the performance metric, a regression model that only knows the distance to the nearest food to the agent fits subjects' responses best. However, such a model's effectiveness lies only on its ability to discriminate between the two most frequent response types: *explore* and *gather*, and this model will never select *attack* or *flee* as the most likely IMP state. Similarly, a model that instead only uses the distance to the nearest other agent performs well by distinguishing between *explore* and *attack*, but will never select *flee* or *gather*. The regression approach needs to integrate information across several input variables, employing a great number more free parameters, to better capture features shared by the patterns of subjects' responses and our Bayesian model.

The Bayesian model, despite being fit primarily to the IMP's programmed states (via simulated data), actually fits subjects' mode responses much better than it fits this ground truth (Tables 1 and 4). Consistent with this result, the model indeed makes similar errors to subjects, with respect to ground truth (Tables 2 and 3, Fig. 7) and finds individual scenes to be similarly difficult (or easy) to classify accurately (Fig. 8). That is, when the Bayesian model did not predict ground truth accurately, it tended to make errors that were similar to those made by human subjects.

3.3.3. Estimating the parameters of the DBN through sampling

The policy and transition probabilities of the IMPs are approximated through sampling of the actual generative model—i.e. by running the simulation a large number of times. The larger the sample, the closer the model will actually approximate the policy of the IMPs. This accumulation of data can be considered the learning mechanism of the model.

For the analysis summarized in Fig. 10, we first collecting a large set (~ 2500) of simulations from which to sample. Then, we took 10 random samples from this larger set, of size 25, 100, 250, or 1000 (for the 2500 scene sample size, we used our entire set). Using KL divergence, we evaluated Bayesian models using IMP policies and goal state transition probabilities approximated from these samples of varying sizes. As shown in Fig. 10, the model's performance improves with greater sample size, with diminishing returns once one draws from samples of 1000 simulations or more.

Table 5
Perceptual variables as predictors of subjects' responses, when entered into a logistic regression. Average KL divergence is shown, with respect to the distribution of subjects' responses (lower is better).

Variable	KL divergence
Distance to nearest food	.436
Distance to nearest agent	.479
Speed of agent	.598
Relative angle of nearest agent	.622
Relative angle of nearest food	.628
Turning velocity of agent	.628

Table 3
Confusion matrix for model's responses in Exp. 1. Mean proportion of IMP time spent in each state is in parentheses, and mean proportion of time model spent in each response category is at the bottom of each column.

Actual state	Model belief			
	Attack	Explore	Flee	Gather
Attack (.16)	.36	.38	.14	.12
Explore (.39)	.13	.55	.13	.19
Flee (.08)	.29	.36	.11	.24
Gather (.37)	.12	.28	.09	.51
	.18	.41	.11	.30

3.4. Discussion

In Experiment 1, subjects were asked to continually categorize the behavior of a target IMP as reflecting one of four possible underlying goal states. Because these goal states existed in the program of the IMPs, there was a "ground truth" basis for assessing subjects' accuracy, and subjects' responses indeed showed moderate agreement with this ground truth. More impressive was the very high level of agreement *among* subjects, which suggested that subjects approached this categorization task in a similar fashion.

By what method might subjects perform this task? We assessed a dynamic Bayesian model as a candidate solution to the problem of goal inference our subjects faced. The Bayesian model must first learn from a large amount of (simulated) training data, and then can approximate subject performance very well. And the more closely the model approximates the actual generative model of the IMP, the better the model fits subjects' responses. This result is consistent with our claim that subjects successfully invert an accurate model of IMP behavior to perform inference about their goal states.

Table 4
Model performance with respect to mode subject responses in Exp. 1. This table shows the proportion of time that the mode subject response matched the maximum likelihood response of the model. Because subjects' mode response was "no response", on average, for the first 8 s of each scene (they failed to respond for the first 13.4% of trials), chance performance for a model is 21.7% and maximum overall performance is 86.6%.

Scene														Overall
1	2	3	4	5	6	7	8	9	10	11	12	13	14	
.68	.58	.73	.80	62	.44	.62	.83	.89	.66	.57	.84	.93	.92	.72

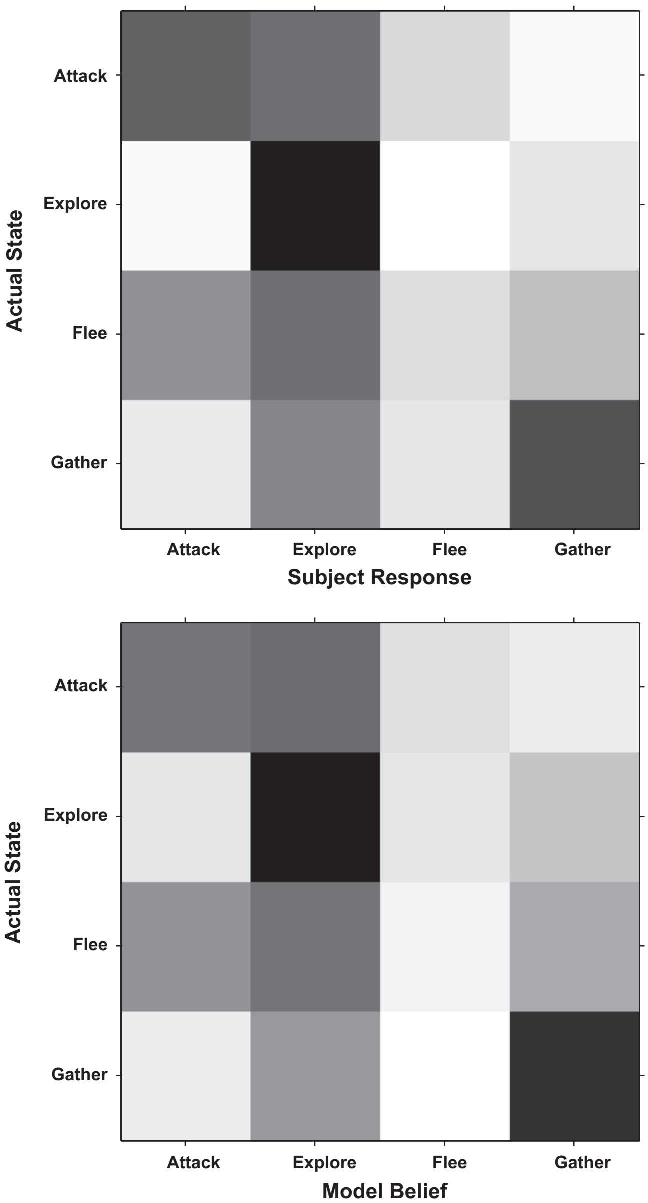


Fig. 7. A visualization of subjects' and model's confusion matrices in Exp. 1. The probability of a response (by either the subjects [top] or model [bottom]), given the "ground truth" state of the IMP, is represented on a gradient from white (very low) to black (very high). The main diagonal represents "hits": correct detections of the actual IMP state. All other cells represent confusions.

4. Experiment 2

In Experiment 2, we manipulated the transition matrices of the IMPs such that their probability of transitioning to the *attack* or *flee* states would be greater (see Tables 6 and 8). In doing so, we intended to create IMPs whose behavioral dispositions would be different from those utilized in Experiment 1, thus allowing us test the generalizability of our computational model to a new set of subjects viewing a new set of IMPs. Additionally, we randomized the order of scene presentation so that we could make valid comparisons between the patterns of subject performance earlier versus later in the experiment.

The ability of the experimenter to systematically manipulate a stimulus agent's behavioral program, as we do in this experiment, is a main advantage of our advocated approach of creating psychophysics experiments using simulated autonomous agents. In this case, the agent manipulation is somewhat subtle: we adjust the parameters of the IMPs' decision making module. But one could manipulate any of the other modules (vision, memory, and path planning) in any number of ways, or remove them entirely, or add additional modules, or swap them with qualitatively different—and perhaps more intelligent and realistic—modules. Which manipulations one wishes to make will, of course, depend on the scientific question

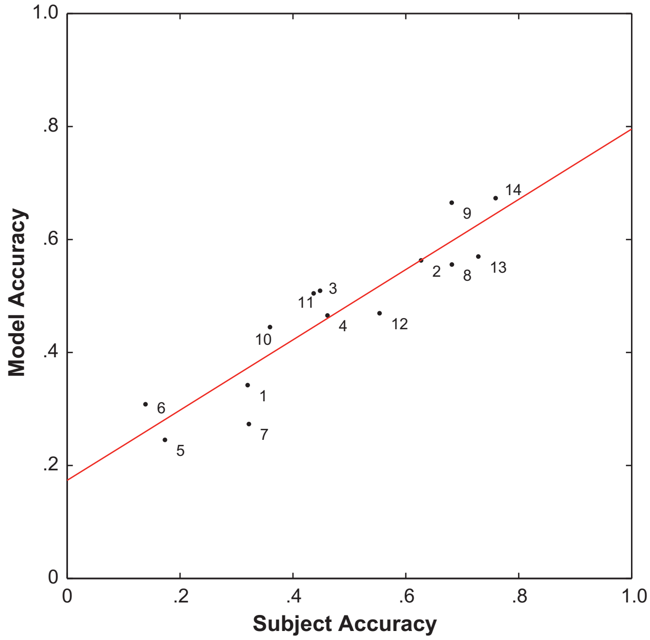


Fig. 8. The correlation between subjects' accuracy and the model's accuracy, with respect to ground truth. Each point in the figure represents one of the 14 scenes from Exp. 1, and is labeled to correspond to the numbering in the other figures.

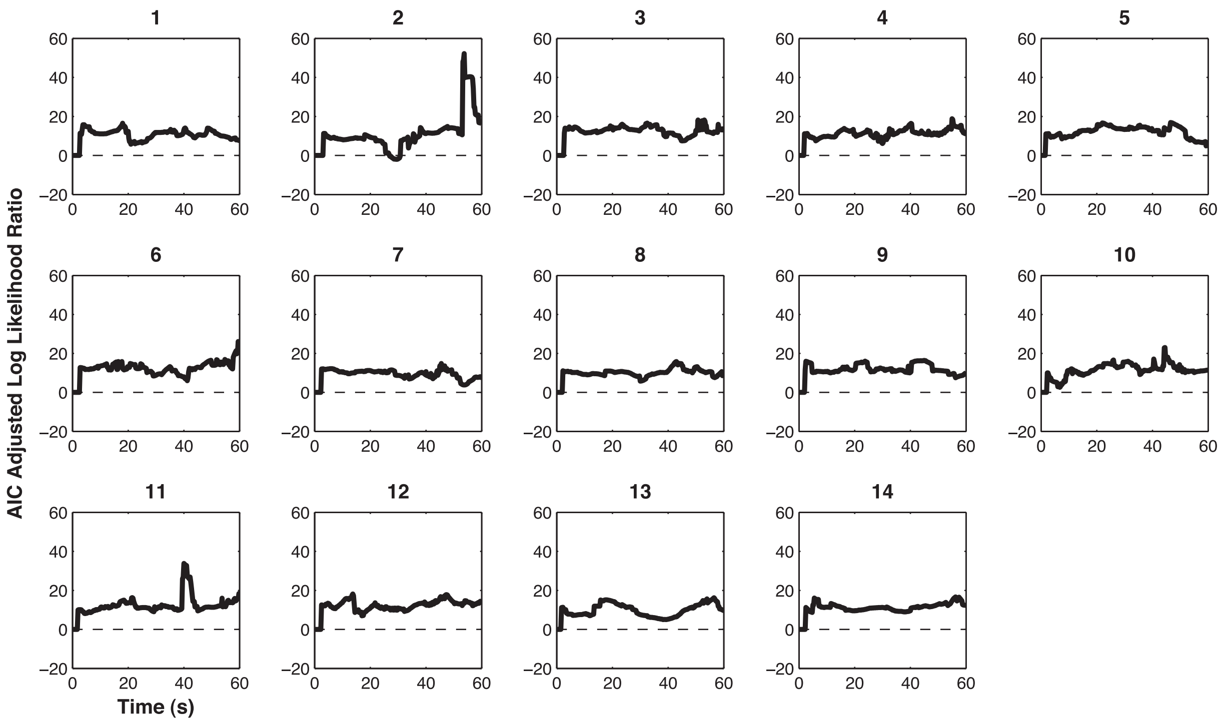


Fig. 9. Experiment 1. Both our Bayesian model and the multinomial logistic regression output a normalized posterior distribution across the four possible goal states. At every quarter-second time slice, we calculate this distribution for either model and plot the AIC adjusted log relative likelihood of the subjects' data. Positive values favor the Bayesian model.

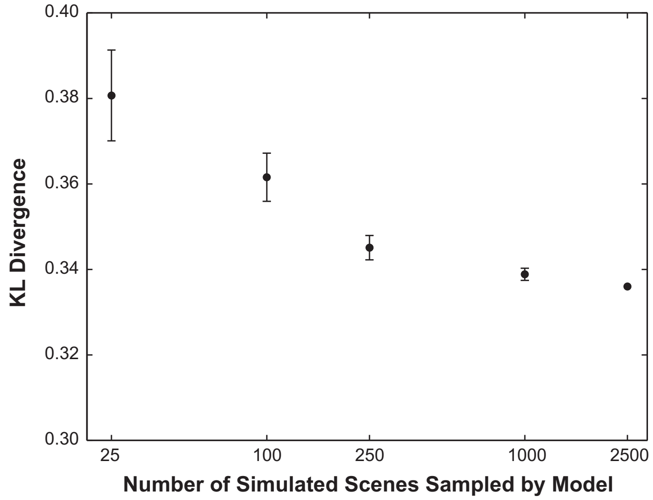


Fig. 10. Model performance as a function of how many simulations are run to approximate the IMP policy and transition matrix. As there are four IMPs populating each scene, each simulated scene actually samples four IMP trajectories.

Table 6
Confusion matrix for subjects' responses in Exp. 2 (averaged across subjects). Mean proportion of IMP time spent in each state is in parentheses, and mean proportion of time subjects spent in each response category is at the bottom of each column.

Actual state	Subject response				
	None	Attack	Explore	Flee	Gather
Attack (.18)	.01	.39	.41	.12	.07
Explore (.45)	.13	.08	.59	.06	.14
Flee (.13)	.02	.19	.43	.25	.10
Gather (.25)	.08	.06	.38	.03	.46
	.08	.14	.49	.09	.20

at hand, and may be as subtle, drastic, or nuanced as the space of possible IMPs will allow.

4.1. Methods

4.1.1. Subjects

Eleven undergraduate students in introductory psychology classes at Rutgers University participated in the experiment, and received course credit for their participation. One additional subject's data were excluded due to failure to follow experimental instructions. Each experimental session lasted approximately 30 min.

4.1.2. Stimuli and procedure

The stimuli and procedure were identical to that of Experiment 1, with the following exceptions:

(1) The scenes were generated using IMPs with modified goal state transition matrices.
(2) The five training scenes were presented to each subject in random order. The first test scene, which was also regarded as practice and thrown out for each subject, was the same for each subject. The following 14 test scenes, included in analysis, were presented to each subject in random order.

4.2. Behavioral results

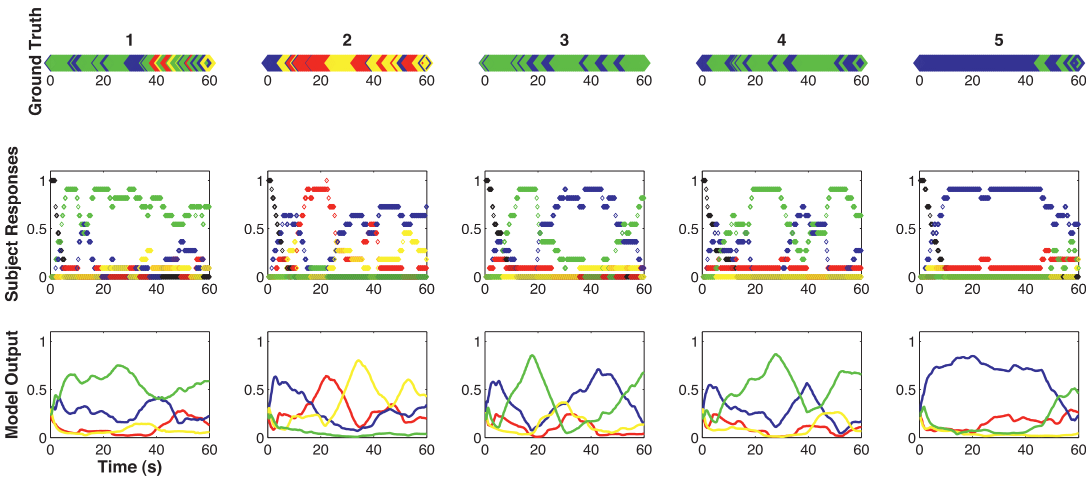
Subjects' overall accuracy with respect to ground truth was 48%, which matched performance in Exp. 1. Excluding portions of trials for which the most common response was "none yet given" (represented by black in Fig. 11), an average of 7.9 out of 11 subjects (72%) agreed upon the mode response at any given time—an intersubject reliability also very similar to that of Exp. 1.

4.3. Model performance

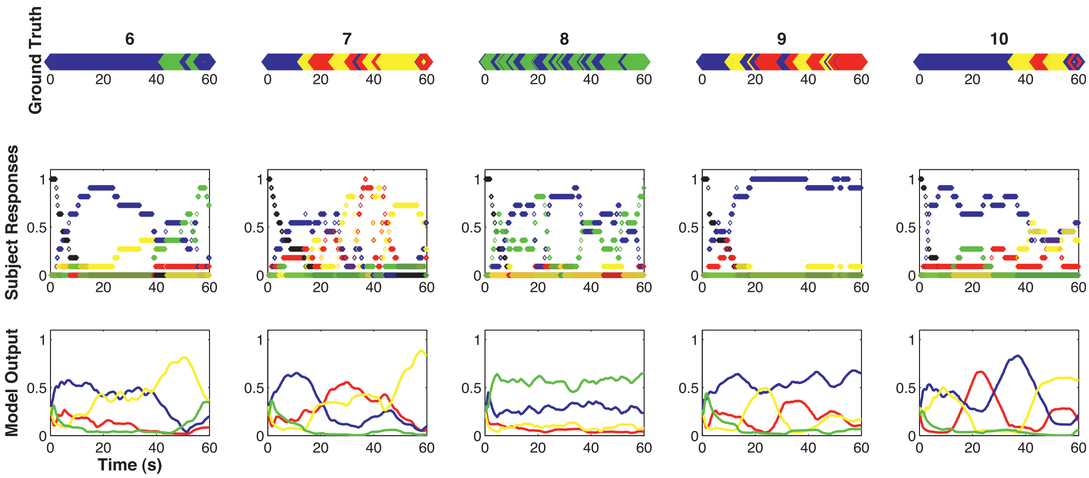
Reusing the free parameters originally fit to Exp. 1's data (100 pixel radius, $5\pi/6$ radian angle, 12 s trailing average) but approximating the policy and transition probabilities of the IMPs with a new set of $\sim$ 1000 simulated scenes, the average KL divergence of the Bayesian model's output distribution and the subjects' response distribution was .382. Fig. 11 illustrates the "ground truth" mental state of the IMP, the distribution of subjects' responses, and the model's output (using this parameterization) for each of the 14 scenes.

As in Exp. 1, this discretization of the agent-centric space results in our best fit for Exp. 2. Fig. 13 illustrates how the model's performance changes as the length of the trailing average (lag) parameter increases. Holding the other two parameters constant (at 100 pixels and $5\pi/6$ radians), model performance asymptotes at around 14 s. As in Exp. 1, setting this parameter to 12 s is at or near optimal.

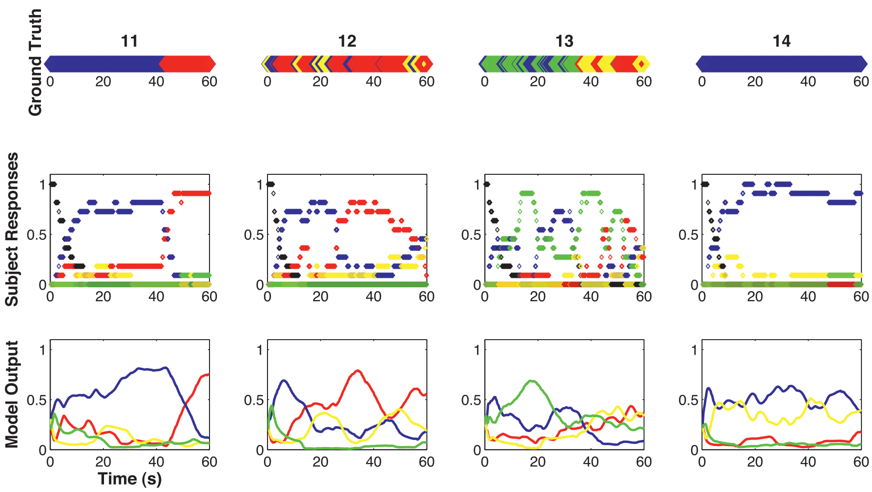
The best cross-validated multinomial logistic regression, using its 15 free parameters originally fit to Exp. 1, did not generalize as well to the new subjects' data (KL = .424). Fig. 12 shows that the adjusted log likelihood (compensating for the number of parameters in both models via AIC) strongly favors the Bayesian model over the logistic one, as in Exp. 1.



(a) Scenes 1-5.



(b) Scenes 6-10.



(c) Scenes 11-14.

Fig. 11. Experiment 2. This figure shows, over time, the underlying "ground truth" state of the agent (top row for each scene), the distribution of subject responses (middle row), and the output distribution of the Bayesian model (bottom row). *Red* represents the "attack" state, *blue* = "explore," *yellow* = "flee," and *green* = "gather." For the subjects' responses, *black* indicates the proportion of subjects who had not yet responded on the keyboard during a given trial. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

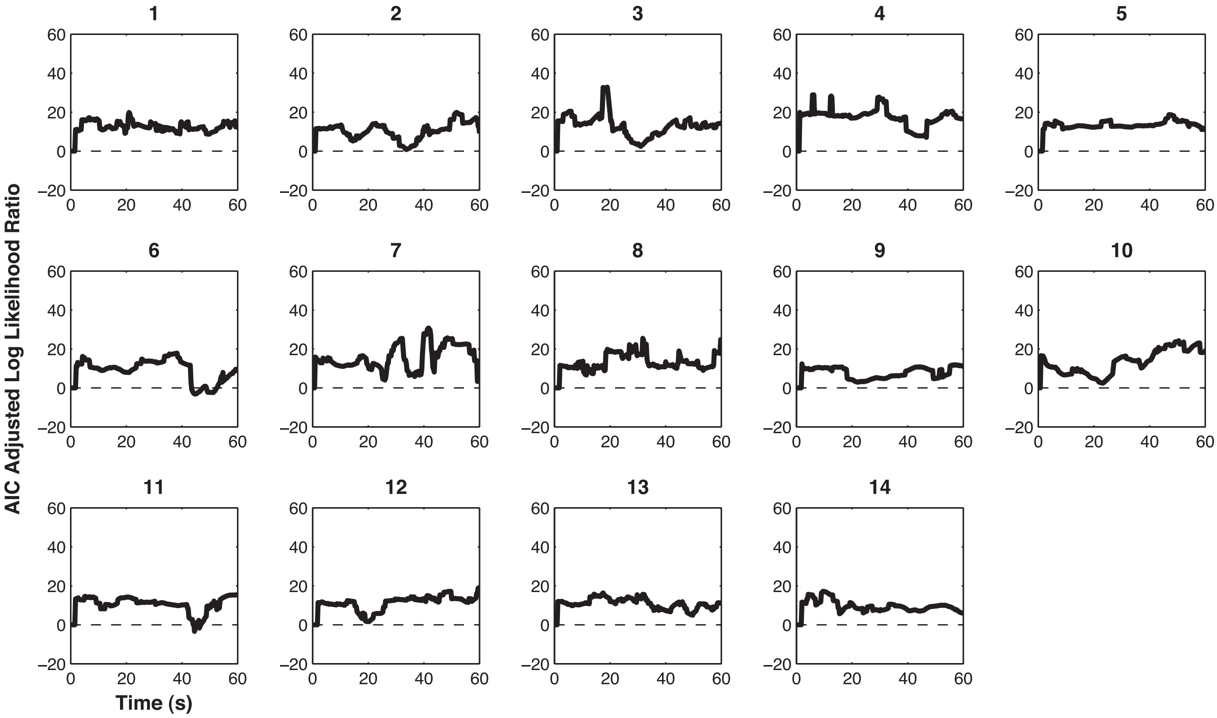


Fig. 12. Experiment 2. Both our Bayesian model and the multinomial logistic regression output a normalized posterior distribution across the four possible goal states. At every quarter-second time slice, we calculate this distribution for either model and plot the AIC adjusted log relative likelihood of the subjects' data. Positive values favor the Bayesian model.

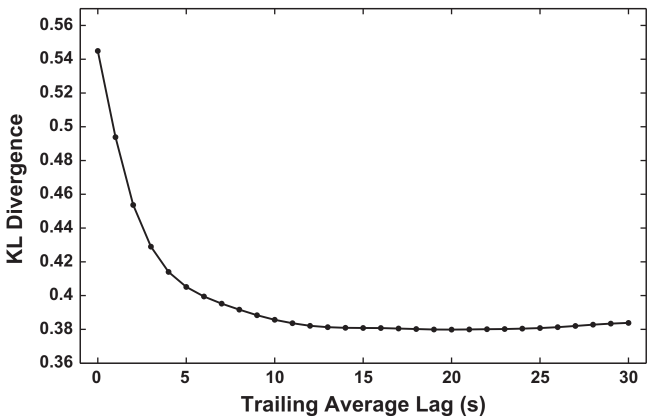


Fig. 13. Model performance as a function of how far back the trailing average (lag) parameter reaches.

To provide some additional context for the performance of these two models, a null model performed better than it did in the previous experiment (KL = .783), as did a model that only learns and applies the base rates of subject responses in Exp. 2 (KL = .569, see Table 6 for these base rates).

4.3.1. Model fit, early versus late in the experiment

One might hypothesize that as subjects become more attuned to the nature of the agents over the course of the experimental session, their performance will come to more closely conform to our model, because this model relies on accurately approximating the underlying generative model governing agent behavior. However, the model did not fit subjects' responses better for later trials versus earlier trials. We performed a repeated measures ANOVA with trial number (1–14) as the independent variable and model accuracy with respect to subject response over the course of the trial (i.e. scene) as the dependent variable. There was no main effect of trial number on conformity to our model ($F[13, 130] = 1.16,\ p = 0.32$). This is perhaps not surprising; because the subject received no feedback over the course of these 14 trials as to whether or not his or her responses were correct (with respect to ground truth or to our model), there is no basis for learning. Whatever subjects learned about the nature of the IMPs apparently was confined to the initial training period.

4.4. Discussion

The pattern of behavioral results we observed in Exp. 1 largely replicated with a new group of subjects viewing IMPs whose program was slightly altered. Both subject accuracy (with respect to ground truth) and reliability (with respect to one another) were consistent with figures obtained in Exp. 1. But, more critically, we demonstrated that our modeling approach—and indeed the specific free parameters fit to the previous data—generalized robustly to this new set of data. Lastly, the data from Exp. 2 revealed that subjects' learning over the course of the experiment did not influence our computational model's proficiency as a model for the inference process used by subjects.

5. General discussion and conclusion

Our data show that subjects are proficient at estimating the IMPs' ground truth goal states, both in terms of reliability (intersubjective agreement) and in terms of validity (accuracy in estimating the true IMP state). Although this internal goal state is only implicit in the IMPs' behavior, subjects can divine it; they can "tell what the agents are thinking," and tend to concur with one another.

Naturally, subjects' performance does not align perfectly with the underlying programming of the IMPs. We found that under-segmentation of the state trajectory was far more common than over-segmentation (cf. Zacks & Tversky, 2001). That is, subjects often missed brief excursions into other states, but rarely indicated a transition between states when one had not occurred. This finding reinforces the idea that detecting a change in intentional state is a concrete computational process that requires sufficient data or evidence in order to yield useful, robust results. That subjects' responses tend to have inertia—their data indicate a tendency to consider not just the IMP's momentary behavior, but observations made during preceding timepoints—is a feature captured by our dynamic Bayesian model's representation of an *a priori* distribution over goal states—$P(G_t|G_{t-1}, S_{t-1})$—which is partially conditional on the cumulative estimate of the IMP's state through the previous time step. This cumulative estimate—G_{t-1}—is itself a distribution over the four possible states which implicitly reflects observations and inferences made over all preceding timepoints.

Subjects' intuitions were fairly consistent across experiments. The optimal discretization of the environment surrounding the IMP (with respect to our model's fit) was the same for each independent sample of subjects. And the expectations and intuitions influencing how a subject would respond in the experiment remained stable across test trials; any learning that occurred did so during the initial practice trials. In other words, the internal model subjects held for our IMPs and their behavior did not appear to change much over the course of the experimental session. Perhaps this should not be surprising, given that subjects did not receive feedback validating or invalidating their responses. Nonetheless, this argues against the role of extensive learning in our subjects.

The development of a useful and robust model of another agent's behavior is indeed central to our treatment of the problem of mental state estimation. However, one important question remains unanswered: How does one access a good generative model of agent behavior? In the case of our computational treatment, an approximate model of the IMPs is "learned" and tabulated via observations made during simulations. But building a model of the observed agent need not require any learning at all: In some cases, one can derive a model of the agent from a simplifying assumption (e.g., the agent is rational) and the prescribed behavior given this assumption (e.g., the agent will behave rationally with respect to its goals, beliefs, preferences, and possible actions), as in previous Bayesian treatments of action understanding (Baker et al., 2009; Ullman et al., 2009). Dennett (1987) suggests that agents who have been subjected to natural selection—such as humans—have evolved to behave approximately rationally most of the time, and this is what allows an observer of these evolved agents to use an abstract, normative standard of rationality to model them.

Yet, querying one's own decision making apparatus does seem to be an attractive approach compared to considering another's complex mental machinery in the abstract. This argument perhaps favors the simulationist account, which posits that determination of other's intentions is based on tacit simulation of one's own behavior (e.g. Heal, 1996; Goldman, 2006, contra *theory theory* or *model theory*, which argues instead that reasoning about another agent's mind relies on a rich representation of goals, beliefs, or intentions and how they relate to behavior, e.g. Stich & Nichols, 2003). However, the system must also be quite flexible. Even if one uses intuitions about one's own decision making process as a starting point, one must be able to tweak this model in light of circumstantial knowledge about the agent's situation and the nature of the agent itself. And the more one is allowed to tweak the self-simulation—the more a question of "What would I do in this situation?" becomes "What would I do in this situation…if I were not me?"—the blurrier distinctions between simulation theory and theory theory become.

The debate between theory theory and simulation theory has been hampered, we would argue, by a dearth of concrete computational models of intention estimation in the literature, which has left somewhat unclear exactly what each position entails or predicts. We hope that the concrete framing of the intention estimation problem provided by the IMPs virtual environment paradigm, along with the computational model for intention estimation that we have proposed, will help focus future debate over underlying principles.

Our methods pave the way towards a true "psychophysics of intention," in which the subjects' perception of *psychological* characteristics of motion in the environment can be studied in the same way that perception of *physical* properties has been studied for decades. They also enable new experimental directions in the study of intentionality, only a few of which have been exploited in this study. For example, if agent behavior can be generated by an underlying program in real time, this allows for immersive experimental paradigms (as in Gao et al., 2010; Pantelis & Feldman, 2012) in which subjects' interaction with agents within the virtual environment—in addition to their explicit judgments—may shed light on underlying cognitive mechanisms. But perhaps most importantly, we argue that using autonomous agents like IMPs as experimental stimuli, and tasking subjects with inferring aspects of their generative program, brings the psychophysics of theory of mind into closer analogy with the modeled process. The inference of mental states is indeed an instance of a more general class of problems faced by the human brain, in which the goal is to estimate the parameters of underlying generative processes of the world.

In future work, we hope to expand the range of behaviors and degree of intelligence exhibited by the IMPs, which, after all, are still extremely limited compared to human agents. Eventually, our hope is to use an improved version of our environment to study comprehension of more cognitively complex phenomena—that is, to move beyond the ''Four F's'' and closer to the range of behavior exhibited by real human agents.

Acknowledgements

An earlier report of some of the data discussed in this paper was presented in the proceedings of the 2011 meeting of the Cognitive Science Society. This research was developed as part of the Rutgers IGERT program in Perceptual Science, NSF DGE 0549115 (http://www.perceptual-science.rutgers.edu/). Special thanks to the Computational Cognitive Science Group at MIT.

Appendix A. IMP programming

The experiment was programmed in MATLAB using the Psychtoolbox libraries (Brainard, 1997; Pelli, 1997; Kleiner, Brainard, & Pelli, 2007). Scenes were rendered offline at a rate of approximately 0.33 frames per second. By recording these frames offline, and saving the timing and essential visual information (location and orientation of all objects) for each frame, we could display scenes to subjects at a later time as animations, bypassing the intensive computations necessary for generating the displays in real time.

To achieve a suitable smoothness and speed during replay to subjects, we linearly interpolated the locations and orientations of objects in the scene between frames, and each scene was sped up 15×. Thus, an originally rendered simulation consisting of ~ 300 frames over 900 s was transformed into a displayed scene of ~ 1500 frames over 60 s (a change from approximately 0.33–25 frames per second). During the experiments, subjects' keyboard responses were recorded after the presentation of each frame: ~25 Hz.

A.1. Environment

Obstacles in the environment are stationary and cannot be moved or traversed by the IMPs. Thus, IMPs must go around them to gain access a blocked location. In addition, like a tall wall, obstacles in an IMP's line of sight occlude its view of other IMPs or food. Because the obstacles cannot move, they create a stable environment that the IMPs can use for path planning.

Food objects and other IMPs are distinct in shape and color, allowing IMPs to identify them against the background. Food is located in clusters, which allows agents to reasonably expect more food to be available at the same location upon later return. When food is ''consumed'' by the agent—which can only be done when the IMP delivers food to its personal predetermined ''cache'' location—it shrinks and disappears, over the course of 1 s. Food remains perceptible to other IMPs once an IMP has grabbed it, and may indeed be stolen from its grasp.

A.2. Perception

The IMP agents are endowed with two perceptual modules: touch and vision.

The IMPs' ''touch'' module is programmed as a simple contact identifier. When an IMP comes into contact with an object, it is made known to its program whether this contacted object is food or another agent.

Modeled as a 1-dimensional retina, an IMP's ''vision'' module allows it to identify color in its field of vision as it navigates the 2-dimensional environment (see Fig. 1b). The algorithm for simulating vision casts a series of rays radially at equal angular intervals from the center of the IMP's heading (i.e. the IMP's ''eye''). Because of the intrinsic geometry of this ray casting, visual resolution is reduced for more distant stimuli.

Three parameters constrain IMP vision: the number of rays cast by the IMP, the distance to which these rays are cast, and the angular field of vision. These parameters were fixed for all IMPs in Experiments 1 and 2: an IMP casts 20 rays, each extending 100 pixels, at equal angular intervals across a 135° field of vision.

Because no depth information is directly available to the vision module, the IMPs need to observe an object from multiple angles to estimate its location with precision. This situation is also known as the inverse projection problem, and useful estimates are made possible in this case by the assumption that food objects are of constant color and size. When rays $[r_m \ldots r_n]$ cast by the IMP detect color corresponding to food, the IMP ascertains that food is located somewhere within the triangular region bounded by the edges of rays r_m and r_n and it uses the centroid of this triangular region as an initial estimate of the food object's location. By successively viewing the food object, the IMP builds up a running average of these centroid-based estimates, increasing the accuracy of its estimated location. The IMP arrives at these successive views without any planning, however, and if other agents or obstacles occlude part of the food-colored region, this can limit the accuracy and precision of the IMP's visual estimates.

A.3. Memory

Using its vision module, an IMP is able to develop a ''mental map'' of the environment. The mental map allows the agents to keep a record of its estimates of the locations of obstacles, food, and other agents, and to plan paths that either intersect or avoid these objects.

The initial map is a *tabula rasa*, with all regions' contents unknown, but it is quickly enriched by experience. The mental map is dynamically updated based upon the input from the vision module; areas where an object may exist are filled in, and areas where nothing is observed are recorded as empty (see Fig. 1d). This dynamic updating is primarily used for food objects and other agents because obstacles are represented as static elements in a binary map. When an object (or nothing) is observed by the vision module, the memory records with a high degree of confidence that this object type is there, or that this area is empty. This confidence about the contents (or lack thereof) of a previously visited area then decreases over time until this region of the map is again considered unknown.

A.4. Goals and actions

An IMP will be in one of four goal states at any given time: *attack, explore, flee, or gather*. The IMP program converts these goal states into actions by deciding upon a particular target location in the environment, and then instructing the IMP to move toward this location according to its path planning algorithm (described below).

If the IMP is in the *attack* state, it finds the nearest location in its mental map where another IMP may be located, and sets this as its target location.

If the IMP is in the *explore* state, it finds the nearest location in the mental map that is unknown—that is, as yet unseen by the IMP—and sets this as its target location.

If the IMP is in the *flee* state, it finds the nearest location in its mental map where another IMP may be located, orients itself in the opposite direction, and moves in this direction.

If the IMP is in the *gather* state, and is touching food, it "grabs" this food and then sets its target location to a predetermined random $[x, y]$ location in the environment designated as the IMP's "cache." If the IMP is not touching food, it finds the nearest location in its mental map where a food object may be located (regardless of the size of the "food" region in its mental map located there), and sets this as its target location.

If the IMP is in the *attack*, *flee*, or *gather* state and has a mental map that contains no information about the locations of food or other IMPs, it will default to the *explore* state. Because the IMP's mental map initializes as a *tabula rasa*, the IMP always begins the simulation in the *explore* state.

The IMP transitions stochastically among these four states, conditional on whether there is food or another IMP located nearby (according to its mental map). In Experiments 1 and 2, food was considered "nearby" if it was fewer than 250 pixels away, and another IMP was considered "nearby" if it was fewer than 100 pixels away. Thus, there are 4 possible situations on which the IMP conditionalizes its behavior—nothing nearby, food nearby, IMP nearby, or food and IMP nearby. Each of these situations corresponds to a 4×4 transition table contained in its program.

The conditional state transition tables for the IMPs used in Experiments 1 and 2 are shown in Tables 7 and 8, respectively.

A.5. Path planning

The action state of the IMP determines a particular target location toward which the IMP must move. The final module of the IMP's cognitive architecture allows it to find the shortest path from its location to this target location, given its present knowledge of the environment (represented by its mental map).

Table 7
Transition probability matrices for the IMPs used in Experiment 1.

		Transition to:			
		Attack	Explore	Flee	Gather
No nearby objects (default)					
	Attack	.94	.02	.02	.02
Transition from:	Explore	.05	.80	.05	.10
	Flee	.02	.02	.94	.02
	Gather	.04	.04	.04	.88
Food nearby					
	Attack	.90	.04	.02	.04
Transition from:	Explore	.02	.84	.02	.12
	Flee	.02	.02	.84	.12
	Gather	.02	.02	.02	.94
Another IMP nearby					
	Attack	.97	.01	.01	.01
Transition from:	Explore	.08	.83	.08	.01
	Flee	.10	.04	.85	.01
	Gather	.03	.03	.03	.91
Food and other IMP nearby					
	Attack	.97	.01	.01	.01
Transition from:	Explore	.20	.40	.20	.20
	Flee	.06	.02	.90	.02
	Gather	.04	.04	.04	.88

Table 8
Transition probability matrices for the IMPs used in Experiment 2.

		Transition to:			
		Attack	Explore	Flee	Gather
No nearby objects (default)					
	Attack	.94	.02	.02	.02
Transition from:	Explore	.05	.80	.05	.10
	Flee	.02	.02	.94	.02
	Gather	.05	.10	.05	.80
Food nearby					
	Attack	.92	.03	.02	.03
Transition from:	Explore	.04	.80	.04	.12
	Flee	.01	.01	.92	.06
	Gather	.04	.04	.04	.88
Another IMP nearby					
	Attack	.93	.02	.03	.02
Transition from:	Explore	.10	.79	.10	.01
	Flee	.03	.02	.93	.02
	Gather	.06	.03	.06	.85
Food and other IMP nearby					
	Attack	.95	.01	.03	.01
Transition from:	Explore	.25	.30	.25	.20
	Flee	.03	.01	.95	.01
	Gather	.13	.04	.13	.70

The IMPs used an iterative implementation of the Floyd–Warshall path finding algorithm to find solutions for the all-pairs shortest path problem. If an area of the IMP's mental map is believed to unoccupied by an object (food, agent, or obstacle)—excluding unknown regions of the map—then the IMP assumes it may traverse this area while planning its path. This algorithm always acts on the current locations of objects in the mental map, and does not take into account whether the extrapolated trajectories of other IMPs in the environment may ultimately block a path at some predictable time point.

References

Abbott, E. F. (1884). *Flatland: A romance of many dimensions*. London: Seely & Co..

Anderson, J. R. (1989). A rational analysis of human memory. In H. L. Roedinger & F. Craik (Eds.), *Varieties of memory and consciousness* (pp. 195–210). Hillsdale, NJ: Lawrence Earlbaum Associates.

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

Barrett, H. C., Todd, P. M., Miller, G. F., & Blythe, P. W. (2005). Accurate judgments of intention from motion cues alone: A cross-cultural study. *Evolution and Human Behavior, 26*, 313–331.

Blythe, P., Todd, P. M., Miller, G. F., & The ABC Research Group (1999). How motion reveals intention: Categorizing social interactions. In G. Gigerenzer, P. M. Todd, & G. F. Miller (Eds.), *Simple Heuristics That Make Us Smart* (pp. 257–286). New York, NY: Oxford University Press.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 443–446.

Burgos-Artizzu, X. P., Dollár, P., Lin, D., Anderson, D. J. & Perona, P. (2012). Social behavior recognition in continuous video. In *2012 IEEE Conference on Computer Vision and Pattern Recognition* (pp, 1322–1329).

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multi-model inference: A practical information-theoretic approach*. New York: Springer.

Crick, C., & Scassellati, B. (2010). Controlling a robot with intention derived from motion. *Topics in Cognitive Science, 2*, 114–126.

Dennett, D. C. (1978). Why not the whole iguana? *Behavioral and Brain Sciences, 1*, 103–104.

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

Dittrich, W. H., & Lea, S. E. G. (1994). Visual perception of intentional motion. *Perception, 23*, 253–268.

Feldman, J. & Tremoulet, P. D. (2008). The attribution of mental architecture from motion: Towards a computational theory (Tech. Rep. No. TR-87). Piscataway, NJ: Rutgers University Center for Cognitive Science.

Gao, T., McCarthy, G., & Scholl, B. J. (2010). The wolfpack effect: Perception of animacy irresistably influences interactive behavior. *Psychological Science, 21*(12), 1845–1853.

Gao, T., Newman, G. E., & Scholl, B. J. (2009). The psychophysics of chasing: A case study in the perception of animacy. *Cognitive Psychology, 59*, 154–179.

Gao, T., & Scholl, B. J. (2011). Chasing vs. stalking: Interrupting the perception of animacy. *Journal of Experimental Psychology: Human Perception & Performance, 37*(3), 669–684.

Geisler, W. S. (2004). Ideal observer analysis. In L. M. Chalupa & J. S. Werner (Eds.), *The visual neurosciences*. Cambridge, MA: MIT Press.

Gelman, R., Durgin, F., & Kaufman, L. (1995). Distinguishing between animates and inanimates: Not by motion alone. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 151–184). New York, NY: Oxford University Press.

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

Goldman, A. (2006). *Simulating minds: The philosophy, psychology and neuroscience of mindreading*. New York: Oxford University Press.

Heal, J. (1996). Simulation, theory, and content. In P. Carruthers & P. K. Smith (Eds.), *Theories of theories of mind*. New York: Cambridge University Press.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology, 57*, 242–259.

Johnson, S. C. (2000). The recognition of mentalisic agents in infancy. *Trends in Cognitive Development, 16*, 637–656.

Kaelbling, L. P., Littman, M. L., & Cassandra, A. R. (1998). Planning and acting in partially observable stochastic domains. *Artificial Intelligence, 101*, 99–134.

Kerr, W. & Cohen, P. (2010). Recognizing behaviors and the internal state of the participants. In *IEEE 9th International Conference on Development and Learning* (pp. 33–38).

Kleiner, M., Brainard, D. & Pelli, D. (2007). What's new in Psychtoolbox-3? In *Perception 36 ECVP Abstract Supplement*.

Knill, D. C., & Richards, W. (1996). *Perception as Bayesian inference*. Cambridge, UK: Cambridge University Press.

Kuhlmeier, V., Wynn, K., & Bloom, P. (2003). Attribution of dispositional states by 12-month-olds. *Psychological Science, 14*(5), 402–408.

Mann, R., Jepson, A., & Siskind, J. M. (1997). The computational perception of scene dynamics. *Computer Vision and Image Understanding, 65*(2), 113–128.

McAleer, P., & Pollick, F. E. (2008). Understanding intention from minimal displays of human activity. *Behavior Research Methods, 40*(3), 830–839.

Pantelis, P. C., Cholewiak, S. A., Ringstad, P., Sanik, K., Weinstein, A., Wu, C., et al. (2011). Perception of intentions and mental states in autonomous virtual agents. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd annual conference of the cognitive science society* (pp. 1990–1995). Austin, TX: Cognitive Science Society.

Pantelis, P. C., & Feldman, J. (2012). Exploring the mental space of autonomous intentional agents. *Attention, Perception, & Psychophysics, 74*(1), 239–249.

Pautler, D., Koenig, B., Quek, B., & Ortony, A. (2011). Using modified incremental chart parsing to ascribe intentions to animated geometric figures. *Behavior Research Methods, 43*(3), 643–665.

Pelli, D. G. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437–442.

Pribram, K. H. (1960). A review of theory in physiological psychology. *Annual Review of Psychology, 11*, 1–40.

Shao, W., & Terzopoulos, D. (2007). Autonomous pedestrians. *Graphical Models, 69*, 246–274.

Stewart, J. A. (1982). Perception of animacy. Unpublished doctoral dissertation, University of Pennsylvania.

Stich, S., & Nichols, S. (2003). Folk psychology. In *The blackwell guide to philosophy of mind* (pp. 235–255). Oxford: Basil Blackwell.

Thibadeau, R. (1986). Artificial perception of actions. *Cognitive Science, 10*, 117–149.

Tremoulet, P. D., & Feldman, J. (2000). Perception of animacy from the motion of a single object. *Perception, 29*, 943–951.

Tremoulet, P. D., & Feldman, J. (2006). The influence of spatial context and the role of intentionality in the interpretation of animacy from motion. *Perception & Psychophysics, 68*(6), 1047–1058.

Ullman, T. D., Baker, C. L., Macindoe, O., Evans, O., Goodman, N. D. & Tenenbaum, J.B. (2009). Help or hinder: Bayesian models of social goal inference. In Advances in neural information processing systems (vol. 22, pp. 1874–1882).

Williams, E. W. (2000). Causal reasoning by children and adults about the trajectory, context, and animacy of a moving object. Unpublished doctoral dissertation, University of California, Los Angeles.

Yaeger, L. S. (1994). Computational genetics, physiology, metabolism, neural systems, learning, vision, and behavior or PolyWorld: Life in a new context. In C. Langton (Ed.), *Proceedings of the artificial life III conference* (pp. 263–298). Reading, MA: Addison-Wesley.

Zacks, J. M., Kumar, S., Abrams, R. A., & Mehta, R. (2009). Using movement and intentions to understand human activity. *Cognition, 112*, 201–216.

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin, 127*, 3–21.