

**USING AUTONOMOUS VIRTUAL AGENTS TO STUDY
THE PERCEPTION OF INTENTION**

by

PETER C. PANTELIS

**A dissertation submitted to the
Graduate School—New Brunswick
Rutgers, The State University of New Jersey
in partial fulfillment of the requirements**

for the degree of

Doctor of Philosophy

Graduate Program in Psychology

Written under the direction of

Dr. Jacob Feldman

and approved by

New Brunswick, New Jersey

October, 2013

ABSTRACT OF THE DISSERTATION

Using autonomous virtual agents to study the perception of intention

By PETER C. PANTELIS

Dissertation Director:

Dr. Jacob Feldman

Mental states (e.g., goals, beliefs, and intentions) may be attributed to agents on the basis of motion cues, and previous studies have successfully related low-level perceptual qualities of a stimulus agent's trajectory (i.e. speed, acceleration, or other manner of motion) to resulting subjective percepts. I argue for a powerful and novel experimental paradigm, in which I utilize a two-dimensional virtual environment populated by autonomous agents whose simulated vision, memory, and decision making capabilities can be manipulated. These agents—nicknamed “IMPs” (*Independent Mobile Personalities*)—navigate the environment, collecting “food” and interacting with one another. Their behavior is modulated by a small number of distinct goal states: attacking, exploring, fleeing, and gathering food. In a first study, subjects attempt to infer and report the IMPs' continually changing goal states on the basis of their motions and interactions. Although these programmed ground truth goal states are not directly observable, subjects estimate them accurately and systematically. I present a Bayesian model of the inference of goal states which accurately predicts subjects' responses, including their pattern of errors. In a second study, I use simulated evolution to create a pool of evolved IMPs which exhibit adaptive behavior. I operationally define IMPs sampled from this simulated evolution as being more rational compared

to non-evolved “control” IMPs, and find that subjects construe evolved IMPs as being both more intelligent and more human-like than non-evolved IMPs. In a final critical experiment, I demonstrate that subjects are better at discriminating the goal states of evolved IMPs than those of non-evolved IMPs. The two studies I present in this thesis provide empirical support for an account of adult “theory of mind” which asserts that 1) the inference of latent mental states can be understood as the inversion of a model of the generative processes producing the observable behavior of the agent, 2) this generative model reflects expectations of agent rationality, and 3) evolutionary fitness is a reasonable operational model of apparent agent rationality, to which subjects are sensitive. These experiments also demonstrate that using autonomous agents as stimuli opens up many basic research questions in the study of the interpretation of intentionality.

Acknowledgements

This research was supported by the Rutgers IGERT program in Perceptual Science, NSF DGE 0549115 (<http://perceptualscience.rutgers.edu/>). I also owe thanks to the Robert J. Glushko and Pamela Samuelson Foundation, as well as the Hellenic University Club of New York, for providing additional funding for my research and its presentation at scientific conferences. Rutgers graduate students Steven Cholewiak, Chia-Chien Wu, Kevin Sanik and Ari Weinstein contributed to the development of the IMPs environment in Matlab. Tim Gerstner and Gaurav Kharkwal additionally contributed to the further development of the IMPs environment in Java. I developed the computational model, presented in Chapter 2, with Josh Tenenbaum and Chris Baker at MIT. All of this dissertation research was performed under the patient mentorship of Jacob Feldman, to whom I give special thanks.

Portions of the research presented in Chapter 2 appeared in *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, and a close version of this chapter has been submitted (but not yet accepted) for journal publication (see References or Vita).

Dedication

This dissertation is dedicated to my supportive parents, to my sister Rachel, and to my fiancée (and perennial pilot subject) Michelle.

I would also like to express my gratitude to the many teachers, friends, collaborators, and helping hands I have encountered at Rutgers during my time here. You are too many to list by name, but I have been overwhelmed by how collectively kind you have been to me, and how generous with your time and energy.

Table of Contents

Abstract	ii
Acknowledgements	iv
Dedication	v
List of Tables	ix
List of Figures	xi
1. Introduction	1
2. Modeling the inference of goal states in intentional agents	3
2.1. Abstract	3
2.2. Introduction	3
2.3. Computational model	9
2.3.1. Approximate representation of states and actions	11
2.3.2. Learning phase	12
2.3.3. Inference phase	12
2.3.4. Free parameters of the Bayesian model	13
2.4. Experiment 1	14
2.4.1. Methods	14
Subjects	14
Stimuli	14
Procedure	15
2.4.2. Behavioral Results	16
2.4.3. Bayesian model performance	18
KL Divergence	18
Fitting and evaluating the model	19
Estimating the parameters of the DBN through sampling	26
2.4.4. Discussion	26
2.5. Experiment 2	28
2.5.1. Methods	28

Subjects	28
Stimuli and Procedure	28
2.5.2. Behavioral Results	29
2.5.3. Model performance	29
Model fit, early versus late in the experiment	30
2.5.4. Discussion	30
2.6. General Discussion and Conclusion	33
3. Interpreting the behavior of autonomous agents selected by simulated evolution	37
3.1. Abstract	37
3.2. Introduction	37
3.2.1. The relationship between agent, environment, and behavior.	38
3.2.2. The intentional stance	40
3.2.3. Artificial life, simulated evolution, and rationality	41
3.3. The IMPS environment	43
3.4. IMP Evolution	44
3.5. Experiment 1: Discriminating between evolved and non-evolved IMPs	47
3.5.1. Experiment 1: Evolution results	47
3.5.2. Methods	52
Subjects	52
Stimuli	52
Procedure	54
3.5.3. Experiment 1a: Basic discrimination	54
Results	55
3.5.4. Experiment 1b: Mismatched description	59
Results	61
3.5.5. Experiment 1c: “Turing test”	63
Results	64
3.5.6. Discussion	65

3.6. Experiment 2: The inference of mental states in evolved (vs. non-evolved) agents	66
3.6.1. Experiment 2: Evolution results	66
3.6.2. Methods	67
Subjects	67
Stimuli	67
Procedure	68
3.6.3. Results: Goal state inference	69
3.6.4. Results: Vision length estimation	72
3.6.5. Discussion	75
3.7. General Discussion and Conclusion	76
3.7.1. Future Directions	77
4. Conclusions	79
5. Appendices	81
5.1. Appendix A: IMP programming, Chapter 2	81
5.1.1. Environment	81
5.1.2. Perception	82
5.1.3. Memory	83
5.1.4. Goals and Actions	83
5.1.5. Path planning	84
5.2. Appendix B: IMP programming, Chapter 3	84
5.2.1. Decision making	87
5.2.2. Evolution	89
5.3. Appendix C: Experimental instructions, Chapter 3	90
5.3.1. Experiment 1a	90
5.3.2. Experiment 1b	91
5.3.3. Experiment 1c	91
Vita	98

List of Tables

2.1. Accuracy of subjects and model, with respect to ground truth goal state	16
2.2. 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.	18
2.3. 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.	20
2.4. Model performance with respect to modal subject responses in Exp. 1. This table shows the proportion of time that the modal subject response matched the maximum likelihood response of the model. Because subjects' modal 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%.	20
2.5. 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).	25
2.6. 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.	29
3.1. Confusion matrix, hit rates, false alarm rates, and d' for each of the IMPs' goal states, for non-evolved (randomly-parameterized) IMPs (pooled across subjects). Mean proportion of IMP time spent in each state is in parentheses. Overall response rates are included in the bottom row. . . .	70

3.2.	Confusion matrix, hit rates, false alarm rates, and d' for each of the IMPs' goal states, for IMPs sampled from the simulated evolution (pooled across subjects). Mean proportion of IMP time spent in each state is in parentheses. Overall response rates are included in the bottom row. . . .	70
5.1.	Transition probability matrices for IMPs used in Chapter 2, Experiment 1.	85
5.2.	Transition probability matrices for IMPs used in Chapter 2, Experiment 2.	86

List of Figures

- 2.1. (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. 8
- 2.2. Graphical model of the Dynamic Bayesian Network with which the IMP is modeled, S represents the environmental configuration, or scene, G represents the goal state of the IMP, and A represents the IMP *action*. S_{t+1} is conditional on both S_t and G_t , A_t depends on S_t and G_t , and the action taken by the IMP at time t , A_t , will influence the subsequent configuration of the environment at time $t + 1$, S_{t+1} 12
- 2.3. The discretization of the agent-centric space, used for the Bayesian model. 13
- 2.4. 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 blue, and the other IMPs were colored black. 15
- 2.5. 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. 17

2.6.	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.	21
2.7.	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.	22
2.8.	Experiment 1. Both the Bayesian model and the multinomial logistic regression output a normalized posterior distribution across the four possible goal states. At every quarter-second time slice, I calculate this distribution for either model and plot the AIC adjusted log relative likelihood of the subjects' data. Positive values favor the Bayesian model.	24
2.9.	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.	27
2.10.	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). <i>Red</i> represents the "attack" state, <i>blue</i> = "explore," <i>yellow</i> = "flee," and <i>green</i> = "gather." For the subjects' responses, <i>black</i> indicates the proportion of subjects who had not yet responded on the keyboard during a given trial.	31

2.11. Experiment 2. Both the Bayesian model and the multinomial logistic regression output a normalized posterior distribution across the four possible goal states. At every quarter-second time slice, I calculate this distribution for either model and plot the AIC adjusted log relative likelihood of the subjects' data. Positive values favor the Bayesian model.	32
2.12. Model performance as a function of how far back the trailing average (lag) parameter reaches.	33
3.1. The left panel shows the evolution of the IMP parameter which determines the transition probability of an IMP remaining in the "gather" state, given that it is currently in this gather state (these time series are smoothed with a 10-generation trailing average). The evolution was done three times. IMPs without any preference for transitioning to the gather state, versus one of the other three possible states, would have this parameter set to .25, represented in this figure with a black dotted line. The right panel shows a histogram of the frequency of all possible values for this parameter, pooled across all three runs of evolution—15000 IMPs in total.	48
3.2. A histogram of the frequency of all possible fields of vision for the IMPs, pooled across all three runs of evolution.	49
3.3. A histogram of the frequency of all possible vision ray lengths for the IMPs, pooled across all three runs of evolution. An asterisk (*) on the x-axis denotes the width of the environment, and a double asterisk (**) denotes the length of the environment's diagonal—the maximum value allowed for this parameter.	50
3.4. A 2D histogram of the joint frequencies of all combinations of parameter settings for memory resolution (x-axis) and vision ray length (y-axis). Darker bins contain a higher frequency of IMPs sampled from the evolution, across all generations of all three runs of evolution.	51

3.5.	A screenshot of a scene from Experiments 1a–1c. The subject’s task was to choose between the red and blue agent on the basis of its intelligence (Exps. 1a–1b) or based on which was more likely to be controlled by a human player (Exp. 1c). Subjects responded on a 6-level scale, from “strongly red” to “strongly blue.”	53
3.6.	ROC curves for all 12 subjects in Experiment 1a, each presented in a different color. Across various confidence thresholds (subjects responded on a 6-level scale), a “hit” is a scene for which the subject selected the evolved IMP, and a “false alarm” is a scene for which the subject selected the non-evolved IMP. The more “bowed” the ROC curve is above the line of no discrimination (represented with a solid black line), and the greater the area under the curve, the better the subject’s discrimination.	56
3.7.	Empirical cumulative distribution functions, along the parameter governing the probability of remaining in the <i>gather</i> state, for various IMP populations.	58
3.8.	Empirical cumulative distribution functions, along the field of vision parameter, for various IMP populations.	58
3.9.	Each evolved IMP was sampled from a random generation of a 1000-generation simulated evolution. Here, IMPs are placed into bins according to the generation of evolution from which they were sampled. Whereas subjects consistently rated evolved IMPs to be more intelligent (i.e. gave a rating greater than the chance level of 3.5, shown here with a dashed black line), IMPs sampled from later versus earlier generations of evolution do not appear to be perceived any differently with respect to intelligence.	60
3.10.	ROC curves for Experiments 1a–1c, with false alarm rate plotted on the x-axis and hit rate plotted on the y-axis.	62
3.11.	A histogram of the frequency of all possible fields of vision for the IMPs, over 1000 generations of evolution performed to generate stimuli IMPs for Experiment 2.	68

3.12. Empirical cumulative distribution functions for estimated and actual IMP vision ray lengths in Experiment 2.	73
---	----

1. Introduction

How do human beings make inferences about the minds of other agents in their world, on the basis of observing their behavior? This dissertation addresses this question, drawing inspiration from and applying the techniques of many disciplines connected to cognitive science: philosophy, developmental psychology, computational modeling, traditional psychophysics, game theory and decision theory, artificial life, and evolutionary algorithms. The synthesis of these elements, via an innovative experimental method, has allowed for a novel approach to the study of human “theory of mind.”

The experiments and theory presented in this thesis are organized into two chapters, each of which stands alone as a study addressing a more specific aspect of the larger question:

In Chapter 2, I introduce the experimental stimuli: Autonomous virtual agents with simulated cognitive and perceptual capabilities, which interact within a 2-dimensional virtual environment. The subject’s task is to observe these agents and infer their intentions on the basis of their motion behavior. I analyze their performance and present a Bayesian computational model of this inference process. This model relies on creating an accurate (but approximate) generative model of agent behavior, and inverting this model to perform inference.

In Chapter 3, I further demonstrate the experimental potential of these autonomous agents. By modifying their cognitive and perceptual capabilities through a simulated evolutionary process, I create a set of agents which have been selected for adaptive behavior and can therefore be considered more rational (compared to non-evolved agents) with respect to the demands of their environment. I then examine whether subjects are sensitive to this enhanced rationality, and ask whether subjects can better infer the mental states of more rational agents. These questions are motivated by theories from the philosophical and psychology literature, which posit that the assumption of the observed agent being rational is critical for the inference of its mental states.

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. I instead manipulate hidden internal states of agents, which influence agent behavior indirectly. This, I argue, provides a closer analogy to the inferential processes comprising a "theory of mind."

2. Modeling the inference of goal states in intentional agents

2.1 Abstract

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, I developed (with collaborators) 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*. I 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). I 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 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.

2.2 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 et al., 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 and 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 et al., 1999; Barrett et al., 2005; Tremoulet and Feldman, 2006; Zacks et al., 2009; Gao et al., 2010; Pantelis and 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 studied in infants and children (Gelman et al., 1995; Gergely et al., 1995; Johnson, 2000; Kuhlmeier et al., 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 are still scarce, 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 and Tremoulet, 2008; Baker et al., 2009; Crick and Scassellati, 2010; Pautler et al., 2011).

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 and Pollick, 2008). While this method has yielded important insights, it suffers from certain critical limitations. 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 agents' mental states, because the agents in such stimuli do not, of course, actually possess mental states; they are simply fictive extensions of the experimenters' intuitions.

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 and Lea, 1994; Williams, 2000; Tremoulet and Feldman, 2000, 2006; Gao et al., 2009; Gao and Scholl, 2011; Pantelis and Feldman, 2012). While this method avoids some of the aforementioned pitfalls of using handcrafted stimuli, the present study represents a substantial departure even from this approach. In the spirit of Dennett (1978)'s suggestion to “build the whole iguana,” the 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, I developed (with collaborators) 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. I refer to the agents as IMPs, for *Independent Mobile Personalities*. Like agents in artificial life environments (e.g. Jaeger, 1994; Shao and 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 (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. 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.

In the studies below, I ask what human subjects can infer about the IMPs' intentions on the basis of observing them move about the environment, and how they might go about performing this inference. 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, my main question is really about the capacity of 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). This paradigm can, similarly, shed light on the relationship between actual behavioral dispositions (like the tendency of an agent to transition into an "attack" state) and their psychological correlates (the subjective impression of intentions). In this sense, I see this 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 I 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 I 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 this paradigm, to correctly ascertain which of the four predefined goal states an agent is in at each point in time. I recognize the need for caution in referring to the four IMPs goal states as "mental states." The IMPs' cognitive capabilities are obviously very limited (see fuller description in Appendix A), and 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. I do not mean to imply that a simple decision matrix is in any real sense equivalent to a real human mental state. Nevertheless, I adopt this phrasing deliberately, because in the context of this study, the IMPs' states play the same role as 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.

In the two studies below, I 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, I 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) I was 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

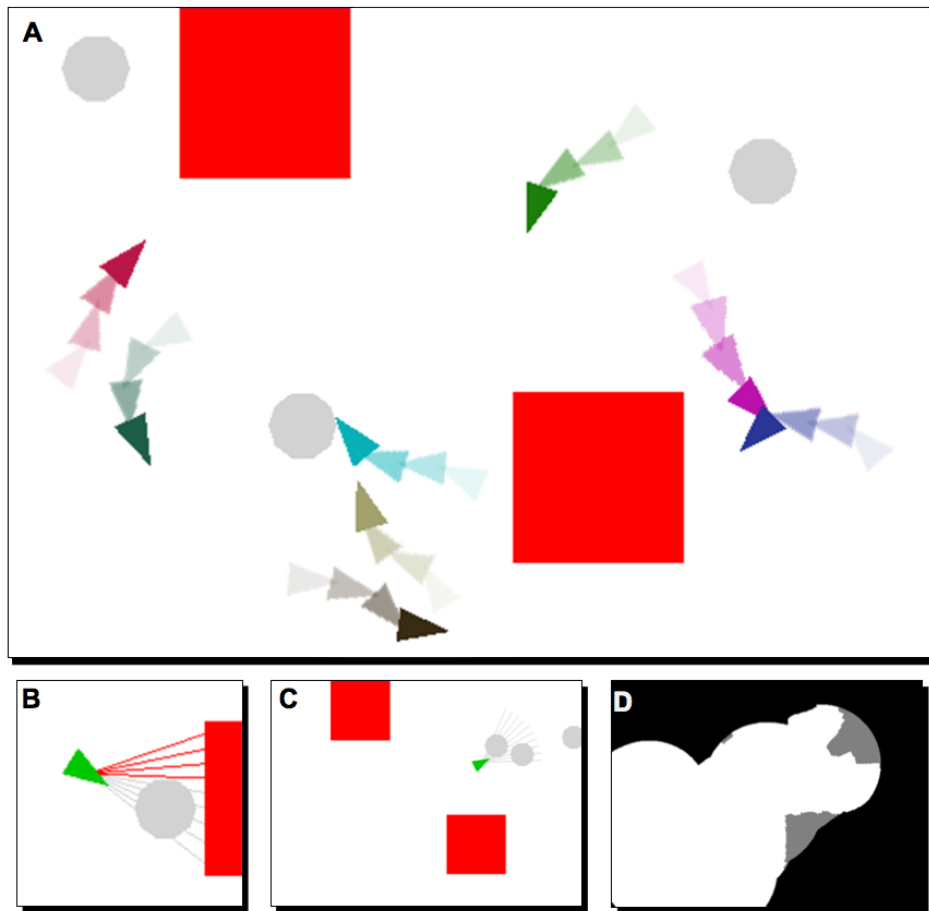


Figure 2.1: (a) The virtual environment with its native IMPs (depicted as moving triangles). (b) The IMPs have autonomous vision by virtue of simulated 1d retinas, and by (c) exploring their environment they can (d) gradually develop a mental map of the objects and obstacles it contains.

agreed with one another (reliability). I 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.

2.3 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 and 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 :

$$p(A|G, S). \quad (2.1)$$

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:

$$p(G|A, S) \propto p(A|G, S)p(G|S). \quad (2.2)$$

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 the IMPs’ behavior is not particularly well-defined, and even if it were, the rational planning computations required for inverse planning would be intractable

given the continuous state, multi-agent, partially observable MDPs required to express the IMPs' domain.

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, I 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 in real time—a truly desirable feature for both natural and artificial systems faced with real-world time constraints.

I formalize this idea by constructing a dynamic Bayesian network (DBN), shown

in Fig. 2.2, 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 parameters of the DBN are learned from prior experience; specifically, I use the actual generative model of the IMPs to produce repeated observations of IMPs' activities. These simulations serve as data for supervised learning.

The DBN employed by the model is only meant to be an approximation of the internal model of IMP behavior that may be harbored by subjects. Also, when compared to the true generative model governing IMP behavior, the 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, I provide technical details about 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.

2.3.1 Approximate representation of states and actions

I first simplify the set of possible IMP actions. Although the IMPs can take on various speeds and angular velocities, I classify all IMP actions as either *turn left*, *turn right*, or *move straight ahead*. Second, I simplify the set of possible environmental configurations. I coarsely discretize the agent-centric physical space into 9 sections (Figure 2.3). The nearest other IMP and nearest food to the target agent can each lie in one of these 9 sections; thus, the discretization scheme allows for $9 \times 9 = 81$ possible configurations of the environment (with respect only to the relative locations of the nearest other IMP and nearest food). It should be noted that subjects may very well rely on a qualitatively different discretization scheme, or may not rely on any such discretization at all.

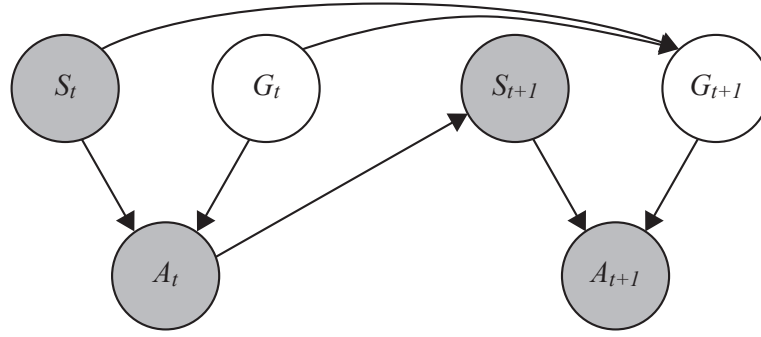


Figure 2.2: Graphical model of the Dynamic Bayesian Network with which the IMP is modeled, S represents the environmental configuration, or scene, G represents the goal state of the IMP, and A represents the IMP action. S_{t+1} is conditional on both S_t and G_t , A_t depends on S_t and G_t , and the action taken by the IMP at time t , A_t , will influence the subsequent configuration of the environment at time $t + 1$, S_{t+1} .

2.3.2 Learning phase

In order to empirically determine this conditional probability of an action at time t , $P(A_t|G_t, S_t)$, I 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 probability of transitioning among the four goal states, $P(G_{t+1}|G_t, S_t)$, represents another set of critical parameters of the DBN that 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 non-deterministic 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.

2.3.3 Inference phase

Once the parameters of the DBN have been learned, inference of the IMPs' goal states, given their observed actions, can be performed. I model the observer's inference by

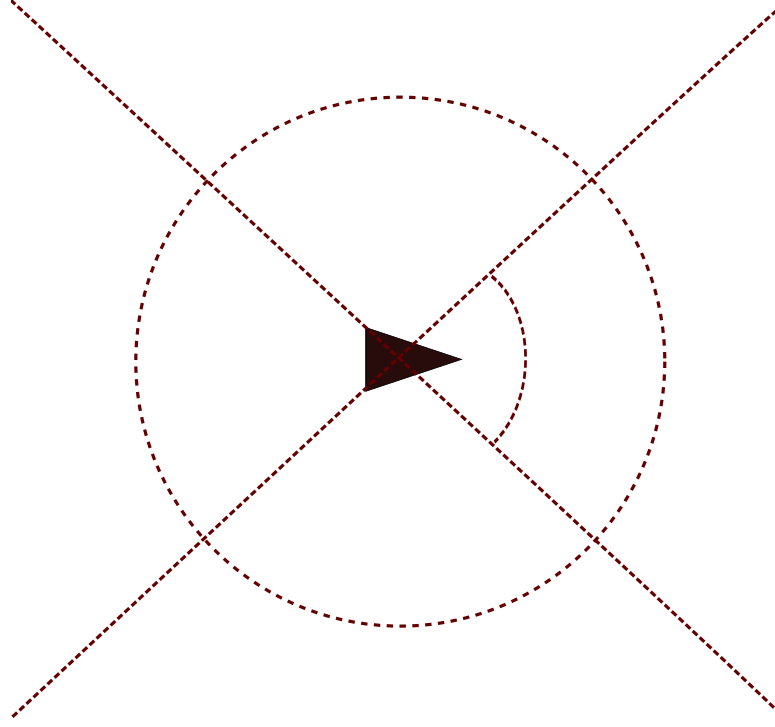


Figure 2.3: The discretization of the agent-centric space, used for the Bayesian model.

computing the marginal posterior probability of a goal at time t , given the state and action sequence up to that point:

$$P(G_t|S_{1:t}, A_{1:t}) \propto P(A_t|G_t, S_t) \sum_{G_{t-1}} P(G_t|G_{t-1}, S_{t-1}). \quad (2.3)$$

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 the probability of changing goals with the likelihood that each goal produced the observed actions.

2.3.4 Free parameters of the Bayesian model

The Bayesian model fits a great many parameters, in an attempt to approximate the generative model of the IMPs as a simplified DBN. However, the parameters of the DBN are all fit to simulated data; they serve to reduce the policy and transition probabilities

of an IMP to lookup tables. Only three free parameters are fit with respect to the subjects' actual responses: two for discretizing 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, the model's outputs are averaged over a trailing period 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. In later sections, I provide a quantitative comparison among these different methods of predicting subjects' responses, along with a non-model-based alternative (multinomial logistic regression) to my Bayesian approach.

2.4 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.

2.4.1 Methods

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 minutes.

Stimuli

Each subject viewed the same set of 20 scenes, generated in advance. Each pre-recorded scene was 60 seconds in duration, and was presented within a 400 x 400 pixel window, horizontally subtending approximately 13.5° of visual angle. Each scene was

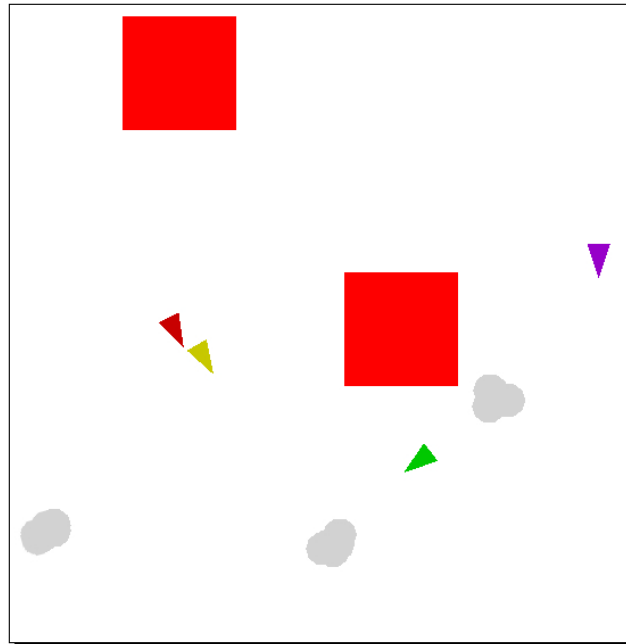


Figure 2.4: A frame from a sample scene viewed by subjects. The red IMP is in an “attack” state, the purple IMP is “exploring,” the yellow IMP is “fleeing,” and the green IMP is “gathering.” Note that colors were only shown during training scenes. During the remainder of the experiment, the target IMP was colored blue, and the other IMPs were colored black.

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 the Appendix A, and example scenes can be viewed at <http://rucss.rutgers.edu/~jacob/demos/imps/>.

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. 2.4). 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.

2.4.2 Behavioral Results

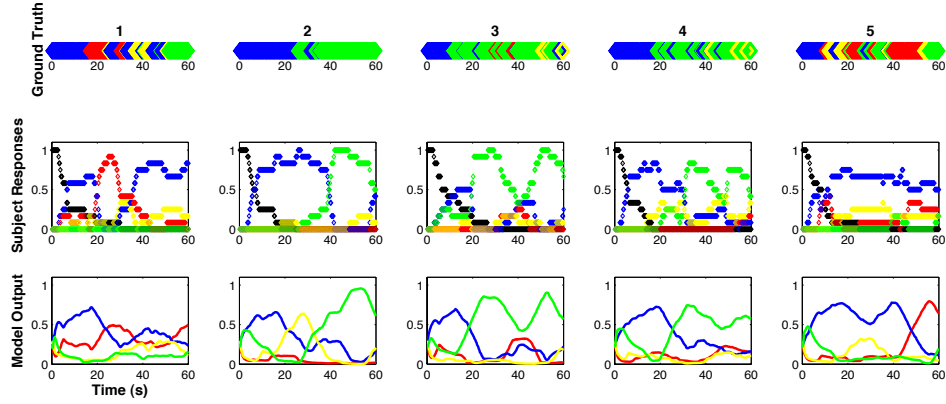
Figure 2.5 illustrates how subjects responded at they observed the 14 test scenes. The “ground truth,” programmed goal state of the target IMP is shown in the top horizontal bar for each scene. The proportion of subjects’ responses across the four response types is shown in the middle row.

I 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 2.1). Mean accuracy was 48%, approximately twice chance performance.

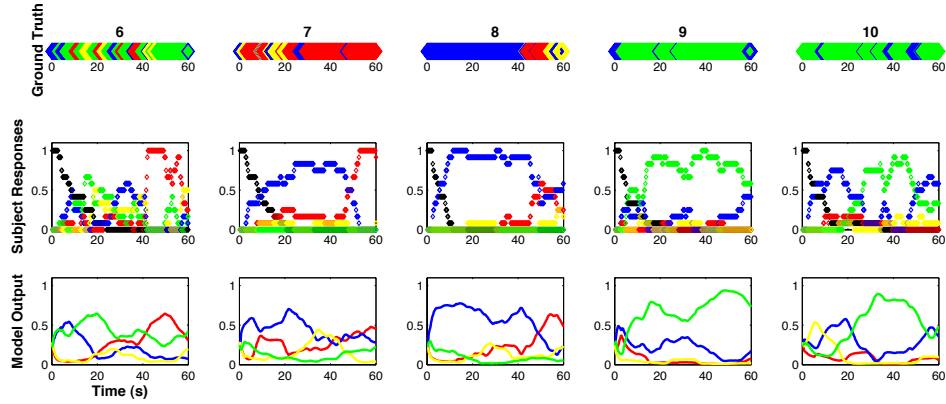
Table 2.1: Accuracy of subjects and model, with respect to ground truth goal state

	Scene														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	All
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

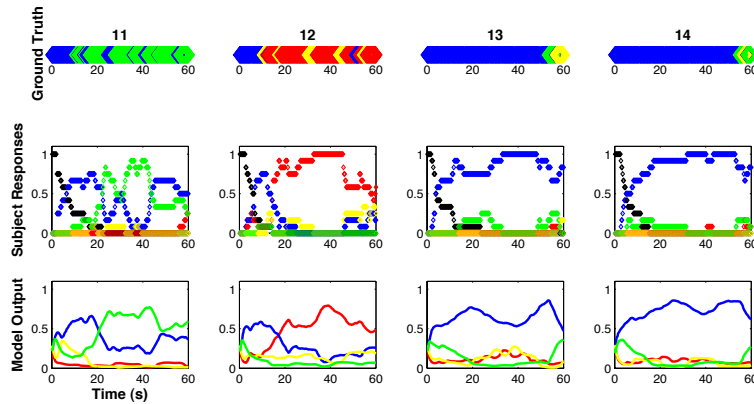
Another critical aspect of subject performance is intersubject agreement (reliability). Excluding portions of trials for which the most common response was “none yet given” (a response category represented by black in Fig. 2.5), an average of 8.7 out



(a) Scenes 1-5.



(b) Scenes 6-10.



(c) Scenes 11-14.

Figure 2.5: 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.

of 12 subjects (73%) agreed upon the modal 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 2.2). 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.

Table 2.2: 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

2.4.3 Bayesian model performance

KL Divergence

The Bayesian model outputs a posterior distribution across the four possible response types. I 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, I

use Kullback-Leibler (KL) divergence as my 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. KL divergence is also a ratio scale measure of performance; that is, a reduction of a factor of 2 in KL divergence corresponds to a representation that is twice as accurate.

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 2.2) 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 $KL = .630$.

Fitting and evaluating the model

As illustrated in Figure 2.5, the posterior distribution (output) of the Bayesian model across the four response types matches quite well with the distribution of subjects’ responses. I fit the model’s three free parameters using KL divergence as the performance metric. The best fitting parameters are not the global best fit: Because assessment of one particular configuration of parameters (while recruiting, in this case, ~ 1200 simulated scenes) may take multiple hours of computer time, the parameter space is coarsely discretized. 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. Additionally, rather than crossing radius \times angle, the model was assessed by keeping one of these parameters constant (at 100 pixels and $2\pi/3$ radians, respectively) and varying the other.

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) both helps to accommodate inertia 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.

Table 2.3: 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

Table 2.4: Model performance with respect to modal subject responses in Exp. 1. This table shows the proportion of time that the modal subject response matched the maximum likelihood response of the model. Because subjects' modal 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

I 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 the dynamical Bayesian model. 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.

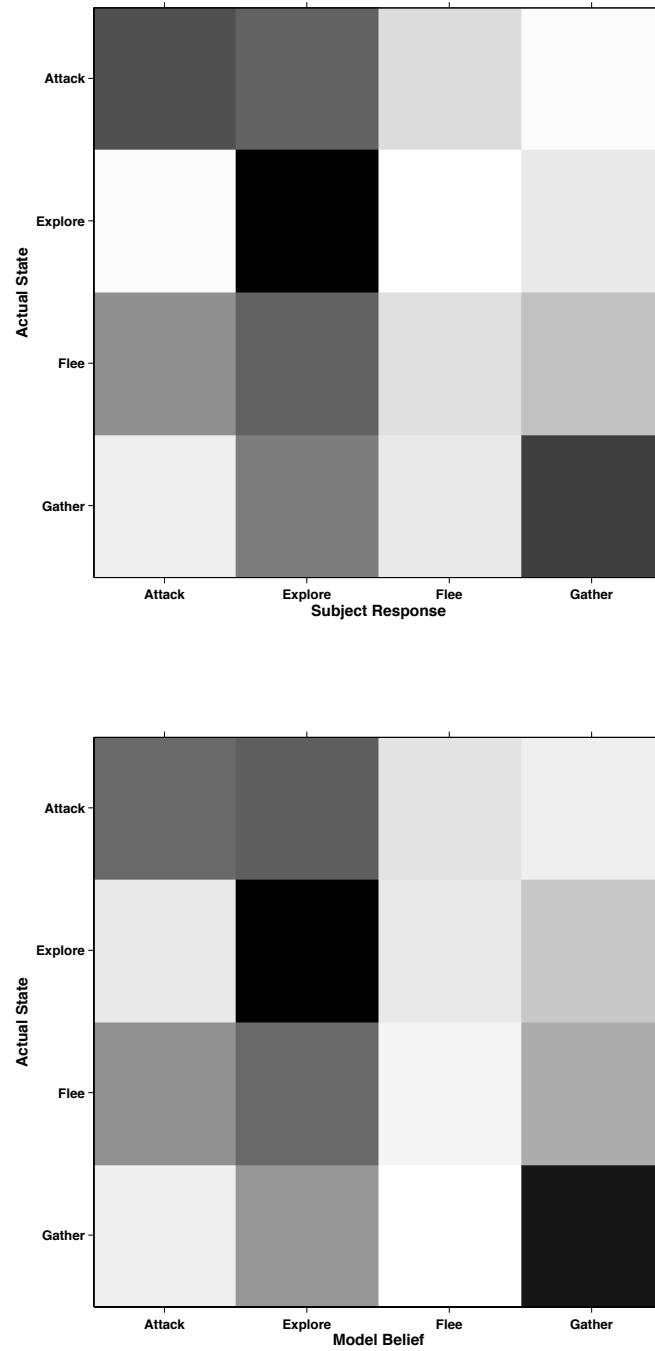


Figure 2.6: A visualization of subjects' and model's confusion matrices in Exp. 1. The probability of a response (by either the subjects [top] or model [bottom]), given the "ground truth" state of the IMP, is represented on a gradient from white (very low) to black (very high). The main diagonal represents "hits": correct detections of the actual IMP state. All other cells represent confusions.

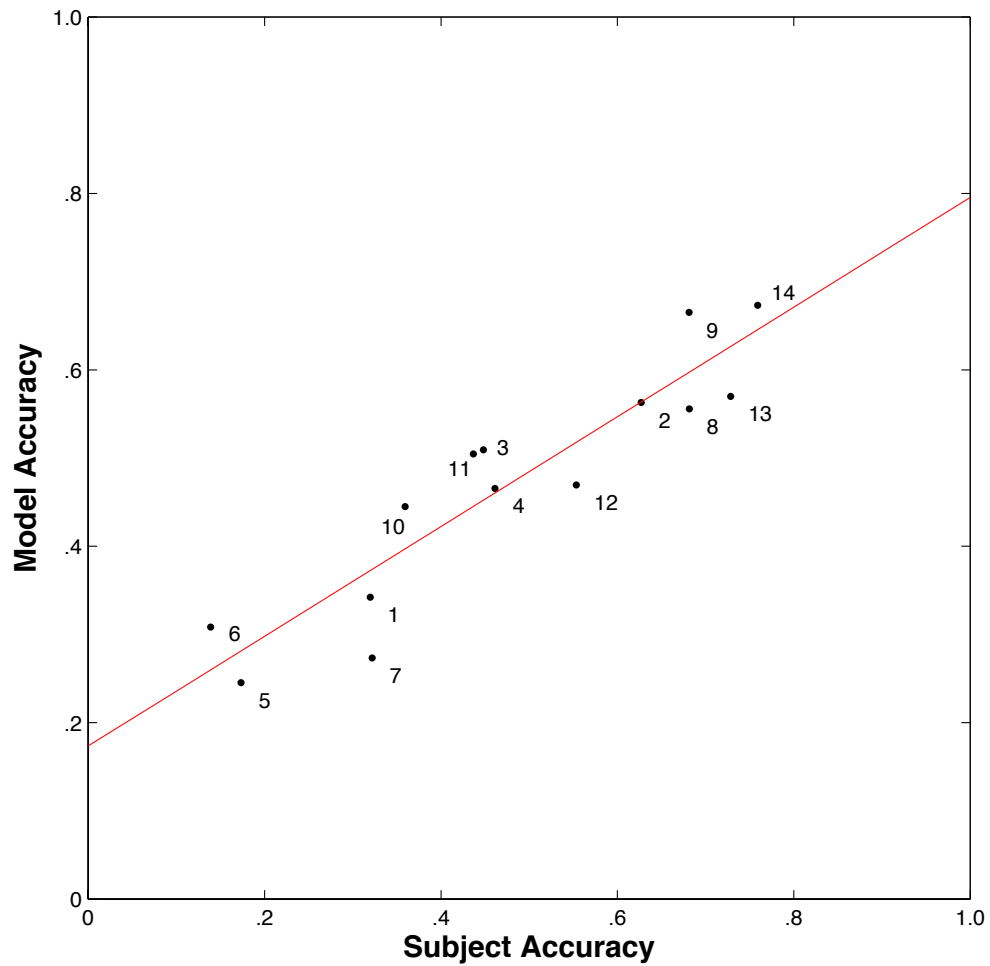


Figure 2.7: The correlation between subjects' accuracy and the model's accuracy, with respect to ground truth. Each point in the figure represents one of the 14 scenes from Exp. 1, and is labeled to correspond to the numbering in the other figures.

This is the same information with which I endow the Bayesian model, although in finer resolution—for input to the Bayesian 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, I 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 (I 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; I 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 the Bayesian model. To compensate for this difference in model parsimony, I calculated the Akaike information criterion (AIC) for both models. I then computed the difference in adjusted (natural) log likelihood (see Burnham and Anderson, 2002), which expresses the relative fit of the Bayesian and logistic models after adjusting for the number of fitted parameters. Figure 2.8 shows this difference over time for each of the 14 scenes (adjusted log likelihood values over zero favor the Bayesian model). As can be seen in the figure, the Bayesian model is nearly always preferred by a very large margin.

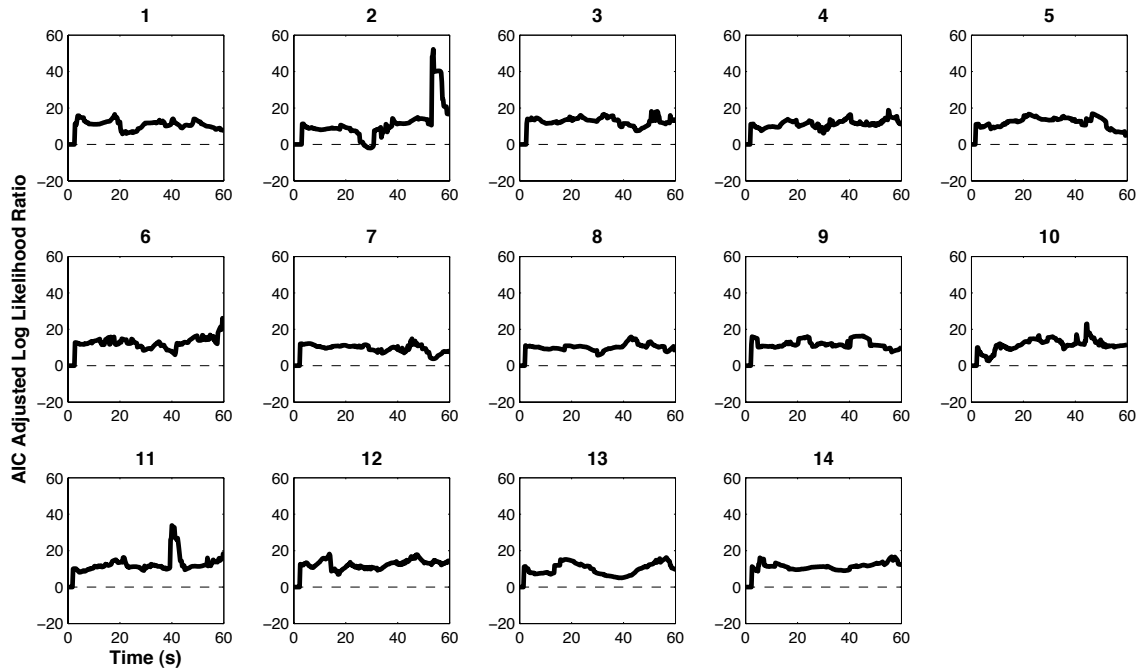


Figure 2.8: Experiment 1. Both the Bayesian model and the multinomial logistic regression output a normalized posterior distribution across the four possible goal states. At every quarter-second time slice, I calculate this distribution for either model and plot the AIC adjusted log relative likelihood of the subjects' data. Positive values favor the Bayesian model.

A benefit of the regression approach was that I could assess the diagnostic value of these input variables when taken in isolation (Blythe et al., 1999). Table 2.5 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 the Bayesian model.

The Bayesian model, despite being fit primarily to the IMP's programmed states (via simulated data), actually fits subjects' modal responses much better than it fits this ground truth (Tables 2.1 & 2.4). Consistent with this result, the model indeed makes similar errors to subjects, with respect to ground truth (Tables 2.2 and 2.3, Figure 2.6) and finds individual scenes to be similarly difficult (or easy) to classify accurately (Figure 2.7). 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.

Table 2.5: 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

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 Figure 2.9, I first collecting a large set (~ 2500) of simulations from which to sample. Then, I took 10 random samples from this larger set, of size 25, 100, 250, or 1000 (for the 2500 scene sample size, I used the entire set). Using KL divergence, I evaluated Bayesian models using IMP policies and goal state transition probabilities approximated from these samples of varying sizes. As shown in Figure 2.9, the model’s performance improves with greater sample size, with diminishing returns once one draws from samples of 1000 simulations or more.

2.4.4 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? I assessed a dynamic Bayesian model as a candidate solution to the problem of goal inference subjects faced. The Bayesian model must first learn from a large amount of (simulated) training data, and then can approximate subject performance very well.

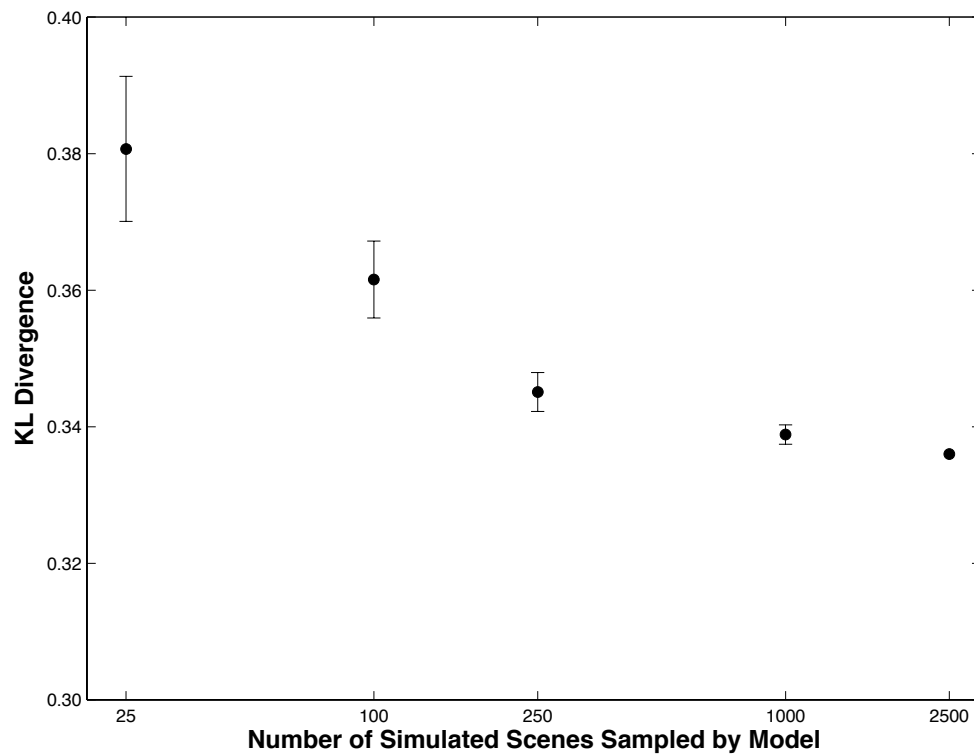


Figure 2.9: Model performance as a function of how many simulations are run to approximate the IMP policy and transition matrix. As there are four IMPs populating each scene, each simulated scene actually samples four IMP trajectories.

2.5 Experiment 2

In Experiment 2, I manipulated the transition matrices of the IMPs such that their probability of transitioning to the *attack* or *flee* states would be greater (see Tables 2.6 and 5.2). In doing so, I intended to create IMPs whose behavioral dispositions would be different from those utilized in Experiment 1, thus allowing us test the generalizability of the Bayesian computational model to a new set of subjects viewing a new set of IMPs. Additionally, the order of scene presentation was randomized, allowing for valid comparisons between the patterns of subject performance earlier versus later in the experiment.

2.5.1 Methods

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 minutes.

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.

Table 2.6: 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

2.5.2 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. 2.10), an average of 7.9 out of 11 subjects (72%) agreed upon the modal response at any given time—an intersubject reliability also very similar to that of Exp. 1.

2.5.3 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 ~ 1000 simulated scenes, the average KL divergence of the Bayesian model’s output distribution and the subjects’ response distribution was .382. Figure 2.10 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 the best fit for Exp. 2. Figure 2.12 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 ($KL = .424$). Figure 2.11 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.

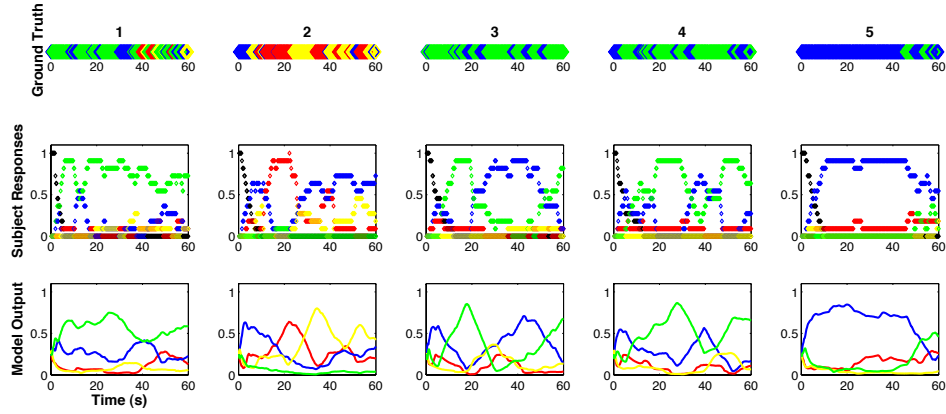
To provide some additional context for the performance of these two models, a null model performed better than it did in the previous experiment ($KL = .783$), as did a model that only learns and applies the base rates of subject responses in Exp. 2 ($KL = .569$, see Table 2.6 for these base rates).

Model fit, early versus late in the experiment

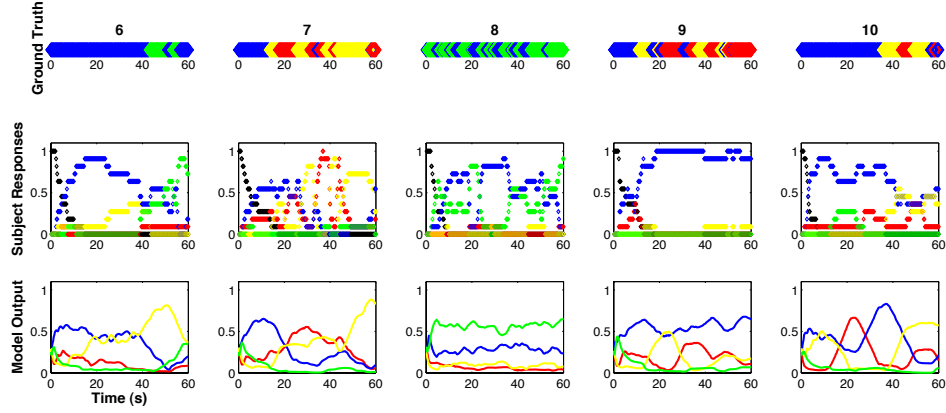
One might hypothesize that as subjects become more attuned to the nature of the agents over the course of the experimental session, their performance will come to more closely conform to the Bayesian model's output, because this model relies on accurately approximating the underlying generative model governing agent behavior. However, the model did not fit subjects' responses better for later trials versus earlier trials. I 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 the Bayesian 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 the model), there is no basis for learning. Whatever subjects learned about the nature of the IMPs apparently was confined to the initial training period.

2.5.4 Discussion

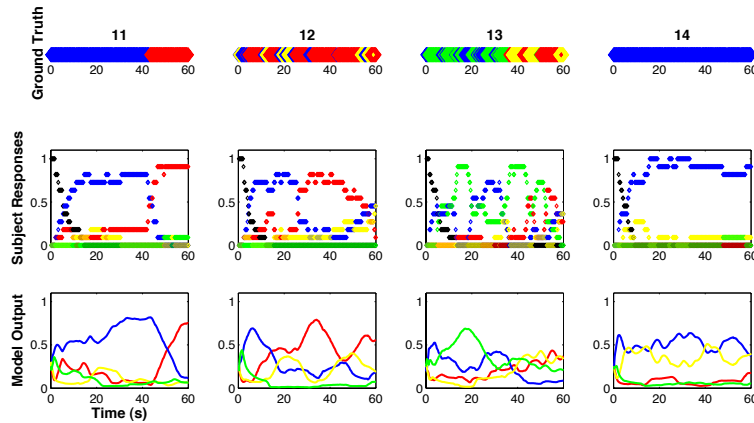
The pattern of behavioral results observed in Exp. 1 largely replicated with a new group of subjects viewing IMPs whose program was slightly altered. Both subject accuracy (with respect to ground truth) and reliability (with respect to one another)



(a) Scenes 1-5.



(b) Scenes 6-10.



(c) Scenes 11-14.

Figure 2.10: 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.

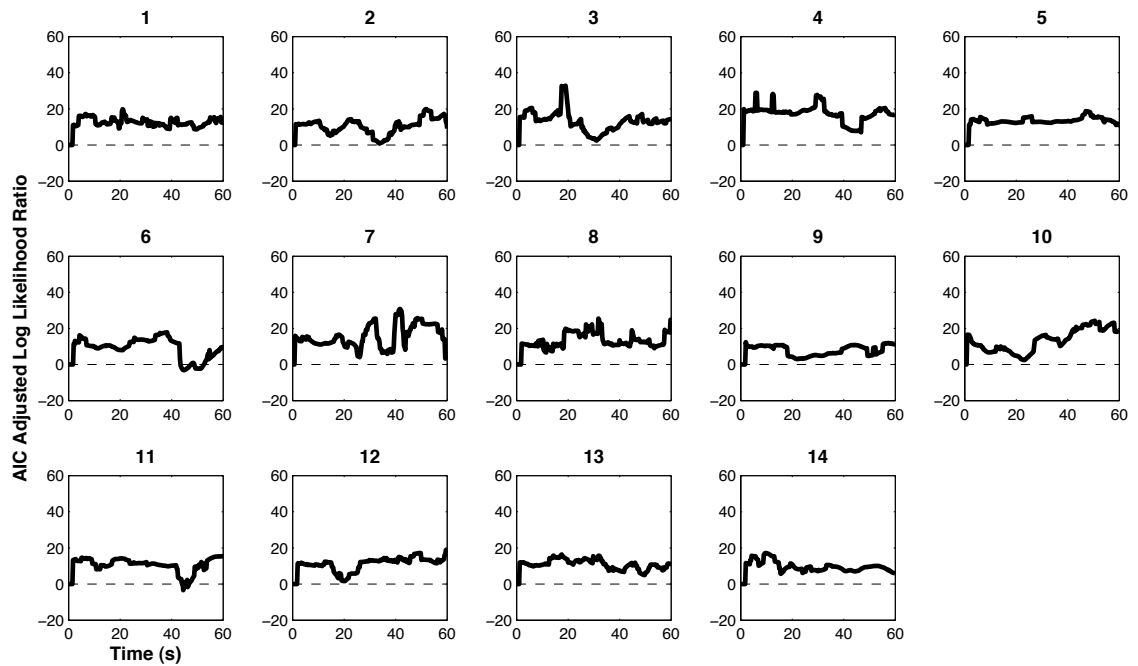


Figure 2.11: Experiment 2. Both the Bayesian model and the multinomial logistic regression output a normalized posterior distribution across the four possible goal states. At every quarter-second time slice, I calculate this distribution for either model and plot the AIC adjusted log relative likelihood of the subjects' data. Positive values favor the Bayesian model.

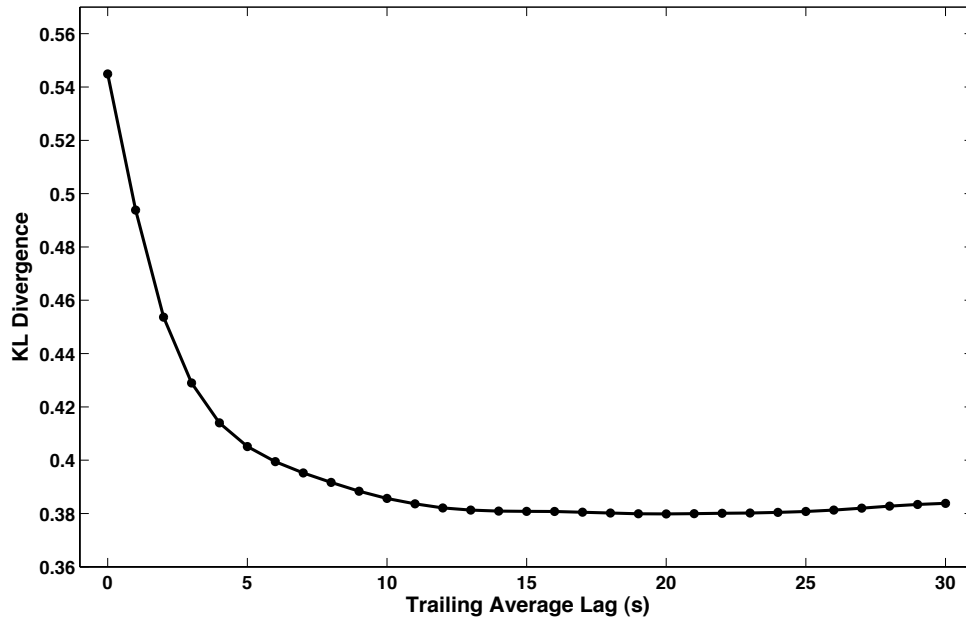


Figure 2.12: Model performance as a function of how far back the trailing average (lag) parameter reaches.

were consistent with figures obtained in Exp. 1. But, more critically, I demonstrated that my 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 the computational model’s proficiency as a model for the inference process used by subjects.

2.6 General Discussion and Conclusion

The 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. I found that under-segmentation of the state trajectory was far more common than over-segmentation (cf. Zacks and Tversky, 2001). That is, subjects often missed brief excursions into other states, but rarely indicated a transition between states when one had not occurred. Of course, most such brief excursions entail virtually no observable change in behavior. This finding merely 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 naturally captured by the dynamic Bayesian model.

Subjects' intuitions were fairly consistent across experiments. The optimal discretization of the environment surrounding the IMP (with respect to the 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 the 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 subjects.

The development of a useful and robust model of another agent's behavior is indeed central to my treatment of the problem of mental state estimation. In the case of my 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, Baker et al., 2009; Ullman et al., 2009). Computational approaches that rely on inverting either the true generative model of the agent, an approximation of this generative model (as in this paper), or

an idealized generative model of a rational agent can each be extremely compressive, reducing the representation of the agent “function” to a compact and generalizable set of rules.

I briefly mention how my computational approach to mental state inference relates to the extensive literature on “mindreading” (Nichols and Stich, 1998). This literature has centered on a debate between *simulation theory*, which posits that determination of others’ intentions is based on tacit simulation of one’s own behavior (e.g. Heal, 1996; Goldman, 2006), and *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 and Nichols, 2003). The computational model I present might be said to have elements of both approaches. On the one hand, it estimates mental states via a model (or theory) of agents that explicitly represents mental states and relates them to actions. On the other hand, I in fact derive this model from extensive simulations of agent behavior, allowing the probability of actions conditioned on the goal state and the scene— $P(A|G, S)$ —to be tabulated. However, one should note that the model I use to simulate and understand other agents need not, and does not, rely on the observer’s *own* decision processes, as in simulationist accounts.

The debate between theory theory and simulation theory has been hampered, I 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. I 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 I have proposed, will help focus future debate over underlying principles.

These 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. My results confirm that subjects can indeed detect mental states

systematically (though of course not perfectly) and make it possible to more directly investigate the computational mechanisms underlying this essential cognitive function. In future work, I 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. Even more ambitious future versions of our environment may be used 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.

3. Interpreting the behavior of autonomous agents selected by simulated evolution

3.1 Abstract

The ability to make inferences about the minds of other agents in the world is highly adaptive to the human (and many other) species, as evidenced by how natural and commonplace these inferences are to our cognitive repertoire. In the psychophysical study of how goals, intentions, beliefs, or other mental states are attributed to agents on the basis of their motion, several previous studies have related low-level perceptual qualities of the agent's trajectory (such as speed) to the resulting subjective percept. In this study, I aim to relate a far more abstract, but powerful, variable to resulting inferences about the mental states of agents: the rationality of the agent. To this end, I created (with collaborators) a 2-dimensional virtual environment, populated by autonomous agents with simulated perception, memory, and decision making. Critically, the modular manner by which we program the agents allows us to modify them toward more adaptive strategies through simulated evolution. We then operationally define agents sampled from this simulated evolution as being more rational compared to non-evolved (randomly-programmed) agents. Subjects were highly sensitive to the evolved status of the agents: They construed evolved agents as being more intelligent, and also more human-like, than non-evolved agents. Furthermore, subjects were more accurate in making judgments about the "mental" states of evolved agents. These results provide critical evidence in support of philosophical and psychological models of theory of mind that posit rationality as a central basis for the attribution of mental states.

3.2 Introduction

There is a special class of objects in the world which is especially salient to the human cognitive and perceptual systems: objects which are alive and are construed to have

minds. The behavior of these objects reflects underlying mental processes, and therefore is interpreted in a qualitatively different manner from that of inanimate objects without intentions (Dennett, 1987; Leslie, 1987, 1994). What makes a particular object seem animate, and what does an object's behavior reveal about its mental state? These questions have historically received a great deal of attention in developmental psychology (e.g. Gelman et al., 1995; Gergely et al., 1995; Johnson, 2000; Williams, 2000) and philosophy (e.g. Heal, 1996; Nichols and Stich, 1998; Stich and Nichols, 2003; Goldman, 2006), and are receiving increasing attention in psychophysics with adult subjects (Blythe et al., 1999; Tremoulet and Feldman, 2000; Barrett et al., 2005; Tremoulet and Feldman, 2006; McAleer and Pollick, 2008; Zacks et al., 2009; Gao et al., 2010; Pantelis and Feldman, 2012) and in computational modeling (Feldman and Tremoulet, 2008; Baker et al., 2009; Crick and Scassellati, 2010; Pautler et al., 2011; Pantelis et al., 2013).

3.2.1 The relationship between agent, environment, and behavior.

The behavior of an agent reflects its external context, or circumstances, but the causal influence of the external physical world on the agent's behavior is indirect, mediated by unseen, internal processes. Different varieties of internal mental or dispositional states—e.g., goals, intentions, beliefs, and emotions—have often been studied as distinct from one another, but all satisfy the following condition: they are internal states of the agent which conditionalize the agent's behavior. Depending on the state of the agent, the observer can expect different behavior. This is why the consideration of mental states plays a critical role in satisfactorily modeling the agent (i.e. possessing a “theory of mind”), and this is why these hidden mental states can be inferred on the basis of behavior by inverting such a model.

This treatment of the problem of inferring the mental state of an observed agent assumes that the agent exhibits behavior that is connected causally with its goals, or intentions, or beliefs, etc. This conception of agency is at least as old as Thomas Aquinas. If this assumption is not at least partially true, then the agent's behavior will be unpredictable on the basis of its mental state, and, inversely, one will be wholly unable to infer the mental state of the agent on the basis of the behavior. In other

words, the behavior becomes non-diagnostic of the hidden state. The agent must be “minimally rational” (Cherniak, 1981).

If one does not model the agent as being in some way rational—doing things *for a reason*—then it is meaningless to consider this object to be an agent, and its behavior ceases to be different from that generated by any other object class, guided by rules constrained by alternative considerations like Newtonian mechanics (Stewart, 1982; Williams, 2000). Positing that an object has a mind—a qualitatively different model for the generative processes producing the object’s behavior—ceases to add explanatory power if this mind (and its beliefs, goals, intentions that lead to decisions) does not have any observable consequence vis-a-vis connecting means to ends.

But beyond this minimum standard, what does it mean to be rational? The normatively rational agent acts in a manner that maximizes expected utility across possible outcomes. Under decision theory (which generalizes to game theory if there are multiple “players”), this agent takes into account all of the possible factors which are relevant to the problem in order to arrive at the optimal decision, which implies a body of knowledge and reasoning framework that “if not absolutely complete, is at least especially clear and voluminous” (Simon, 1955, p. 99). This burden can be difficult to meet in practice, especially when the set of relevant environmental variables is poorly-defined, continuous, dynamic, or not fully observable to the agent. When other agents are added to the environment, and the rational strategy will depend on the strategies adopted by these other agents (i.e. in a game theoretic setting), this further compounds the problem.

An alternative perspective on agent choice relaxes the demands on the agent in terms of the amount of information, time, and computational power it is expected to have at its disposal when making a decision (Simon, 1955; Kahneman, 2003). This approach—“bounded rationality”—sacrifices the analytic precision and optimality of stricter mathematical models of rational choice, in favor of more approximate—but generally effective, and computationally tractable—methods of reasoning under uncertainty. Dennett (1987) simplifies the definition of agent rationality even further, and firmly embeds the concept of rationality into the environmental niche of the agent:

What is (approximately) rational for the agent is what is adaptive with respect to its (empirical) evolutionary success, even if “the demands of nature and the demands of a logic course are not the same”, and our cognitive and perceptual faculties may be “nothing more than a bag of tricks” (p. 51).

3.2.2 The intentional stance

Modeling the agent as a rational actor, by one of the above standards, can be quite useful in trying to understand its behavior. Dennett (1987) defines this strategy of taking the “intentional stance” in this manner:

[F]irst you decide to treat the object whose behavior is to be predicted as a rational agent; then you figure out what beliefs that agent ought to have, given its place in the world and its purpose. Then you figure out what desires it ought to have, on the same considerations, and finally you predict that this rational agent will act to further its goals in the light of its beliefs. (p. 17)

The Baker et al. (2006, 2007) approach (also applied in Ullman et al., 2009; Tauber and Steyvers, 2011) is the *inversion* of Dennett’s intentional stance, formalized computationally. The inferential process can be conceived as “inverse planning”—i.e. positing a rational actor and then inferring the goals that would be the most sensible given the observed behavior:

At issue in both theories is a mapping that is fundamental to a theory of mind: a mapping from intentions to predicted actions, and from observed actions to intentions. This mapping is a many-to-many relation (Searle, 1984; Malle et al., 2001), which is what makes theory of mind so difficult to define and tackle computationally.

The intentional stance serves to simplify a complicated physiological and psychological reality (the vastly complex underlying process generating another agent’s behavior) in a manner that is adaptive in its explanatory and predictive power. Like all models, the intentional stance is an approximation. It may be applied to inanimate

objects in the world as well as truly intentional ones¹, and also to groups and markets, which may (or may not) behave rationally, but are not sentient nor volitional, except in the abstract. The intentional stance may even be applied to the self as one attempts to rationalize one's behavior after the fact.

If human theory of mind relies in the application and inversion of the intentional stance (i.e. a rational model of the agent), then 1) people ought to be sensitive to the level of perceived rationality in observed agents, and 2) the more rational the behavior of the agent, the better people ought to be at inferring their mental states. Testing these predictions experimentally requires:

- a means for creating intentional agents to be used as stimuli
- a definition of agent rationality that is well-defined and non-circular
- a way to manipulate this level of rationality in stimulus agents, and
- a “ground truth” for assessing the accuracy of human subjects as they estimate the mental states of the agents.

Below I present an experimental paradigm that satisfies all of these criteria, presenting a 2D simulation environment to subjects in which autonomous agents (nicknamed “IMPs” for “Independent Mobile Personalities”) compete in a simple game. These agents have modular perceptual and cognitive capabilities which determine their behavior and can be manipulated parametrically. Importantly, for the purposes of this study, the parameters governing the perceptual and cognitive architecture of the IMPs can also be *evolved*. I use an IMP's fitness within the environment as my operational definition of rationality, and let a simulated evolution select for this fitness.

3.2.3 Artificial life, simulated evolution, and rationality

Artificial life can be loosely defined as the modeling and simulation of biological processes or behaviors, with the goal of imitating life with increasing fidelity. The

¹Though, to the materialist, modeling an object—animate or inanimate—as “truly” having a mind is never a veridical account, which would invalidate this distinction.

study of artificial life cuts across many disciplines, perspectives, and eras (Chaitin, 1970; Braitenberg, 1984; Yaeger, 1994; Carnahan et al., 1997; Shao and Terzopoulos, 2007), and need not involve any evolutionary or genetic algorithms. But when evolutionary programming is applied to the subfield of artificial life, it is perhaps applied the most literally—The “organisms,” “generations,” and “environments” to be modeled and simulated are hardly metaphorical at all.

Evolutionary approaches differ in detail, but require the same collection of elements: 1) A representation of the agent that can be altered in some way, 2) A method for generating a population of agents, 3) a means of assessing the “fitness” of each agent, and 4) A strategy for creating subsequent generations of agents based on this assessment (i.e. “selection and genetic operators,” Mitchell and Forrest, 1993).

The relative adaptiveness of various strategies can be assessed as proportional to their respective frequencies in the evolved population (Bicchieri, 2009). Depending on the particular evolutionary algorithm and domain, this iterative process may ultimately converge to one optimal or approximately optimal strategy. In many other interesting cases, the resulting population of agents may have multiple modes representing co-evolved stable equilibria, for example in the simulated co-evolution of pursuit and evasion, or predators and prey (Reynolds, 1994; Cliff and Miller, 1996; Nolfi and Floreano, 1998). These treatments successfully model the concept of evolutionarily stable strategies first put forth by Smith and Price (1973), which merged the concept of the Nash equilibrium from game theory with the study of evolutionary dynamics. However, evolution—even the simulated variety—is a complex, noise-sensitive process. The outcomes are not always clean nor stable, and the dynamics can be chaotic (Nowak and Sigmund, 2004).

The evolutionary algorithm I employ meets a basic minimum threshold of effectiveness for the purposes of the psychophysics experiments presented in this study: IMP strategies which are more adaptive (i.e. more rational within the constraints of this domain) must appear with greater frequency in the evolved population. I sample IMPs from various generations of simulated evolution and consider these sampled IMPs to be operationally “more rational” than randomly parameterized (non-evolved)

IMPs.

3.3 The IMPS environment

IMPs are virtual robots rendered on the computer screen as moving triangles. Pantelis and Feldman (2012) also used moving triangular automata, modeled as reflex agents, as psychophysical stimuli, and Pantelis et al. (2011) built on this approach, employing more complex agents with architectures closely resembling those used in the present study. As in the animated film created by Heider and Simmel (1944), the agents populating the virtual environment are rendered as basic geometric shapes (in the case of the IMPs, isosceles triangles), an approach that is designed to isolate the motion behavior of these stimuli as the critical aspect of the scene to be connected to how subjects perceive them.

The ability of the IMPs to survive and thrive during each simulated generation of evolution involves finding and collecting “food” in the environment and bringing it back to a pre-specified “cache” location while avoiding obstacles and the harassment of other IMPs.

An IMP’s behavior is determined at all times by three factors: 1) what behavioral state it is in, 2) its knowledge of whether food and other imps are nearby (modulated by both perception and memory), and 3) its method for navigating this environment. An IMP can be in one of four states: “explore,” “attack,” “hide,” or “gather,” and each IMP follows the same simple program given that it is in one of these four states. What distinguishes the agent behaviors, and allows them to be evolved, are their differential capabilities building and utilizing mental maps of the environment and their probabilities for transitioning between their possible behavioral states. The strategy for a respective IMP (its location within the “IMP strategy space”) can therefore be represented by a long set of discrete and continuous parameters, and the search for the optimal strategy, given the constraints on the IMP architecture and the rules of the “game”, can be conceived as a search of this high-dimensional strategy space.

One might have a loose intuition about what IMP parameter settings might be better than others, but attempting to analytically derive the optimal (normatively rational) agent strategy for an IMP playing this dynamic game within a continuous, constantly-changing environment—itsself populated with a virtual infinity of other possible IMPs—would be a futile effort. I therefore attempt, through simulated evolution, to generate IMPs which, though probably not optimal in their parameterization, behave more rationally than other IMPs which have not been selected through this process.

3.4 IMP Evolution

Each IMP's program is governed by 22 parameters,² distributed over four cognitive and perceptual modules:

- **Vision.** The IMP's vision module is modeled as a 1-dimensional retina, which receives a 1D array of information from a finite set of vision rays cast forward from the center of the IMP's body. If a ray intersects with an object in the environment, the vision module returns its color. Three parameters govern vision: the number of rays cast (an integer from 1 to 50), the angular field of vision (a continuous value from 0° to 195°), and the length of the rays (a continuous value from 0 to 56, with 56 being the approximate length of the environment's diagonal). The rays are projected at equal angular intervals across the field of vision. As a result, the larger the field of vision, the wider the space between rays (given a constant number of rays).
- **Memory.** Using input returned from its vision module, the IMP continually updates a mental map of its environment. The environment is subdivided into an $n \times n$ grid, where n is a parameter which is allowed to take on an integer value from 60 to 175. The IMP may believe each of the square subdivisions in this grid

²Technically, only 18 of these parameters are free to vary. 16 of the parameters represent the probabilities in a 4×4 goal state transition table (see "Decision making"), and because the rows in this transition table are constrained to sum to 1, this eliminates 4 degrees of freedom in the IMP program.

to contain another IMP, food, an obstacle (walls are considered to be obstacles), or to be unknown.

- **Decision making.** The IMP is a finite state automaton, and can be in one of four states at any given time: *attack*, *explore*, *flee*, or *gather*. When an IMP makes a “decision,” it decides which of the four states it will be in at the next iteration (i.e. frame) of the simulation. It also decides on a target location towards which to move, which is a function of the state in which the IMP resides and its current memory representation of the environment.

One parameter determines how frequently the IMP makes a decision to both update its state and its target location. This parameter can take on an integer value from 1 (i.e. the IMP updates its state/target each at each and every frame) to 90 (i.e. the IMP updates its state/target every 90 frames).

The IMP transitions to a new state (or remains in the same state) stochastically, according to a 4×4 transition table. Thus, this 4×4 transition table represents 16 of the IMP’s parameters (though only 12 of them are free to vary; see footnote).

- **Path planning.** To assist in determining the shortest unimpeded path to a target location in the environment, the IMP builds a path planning grid. This grid subdivides the environment into square sections, each of which the IMP may believe is either safe to be traversed, or possibly contains an obstacle. This is an $n \times n$ grid, where n is a parameter which is allowed to take on an integer value from 1 to 50. If the value is set to be very low, then the IMP path planning algorithm will typically return that there are no unimpeded paths to the target location, and the program will default to taking a direct path to the target location, without regard to whether an obstacle may be standing in the way.

An evolution initializes with a population of 5 IMPs, each with a randomly generated set of parameters. During each generation, these 5 IMPs interact within the environment for 3600 frames (displayed during psychophysics experiments at a rate of 60 frames per second). An IMP begins with 50 “health” points, and automatically

loses .01 points each frame. If the IMP's health ever reaches zero, the IMP "dies" and becomes a piece of food, which may then be treated like any other piece of food that exists in the environment.

If the IMP neither collects any food nor collides with any other IMP, it will have a "health" of 14 at the end of a generation, which is a fairly common outcome. But each time an IMP successfully brings a piece of "food" back to its initial location, it receives a benefit of 50 points. If an IMP is in the "attack" state, and collides with another IMP that is not in the attack state, the other IMP loses .4 points and the attacking IMP sustains no damage. If the other IMP is also in the attack state, then both IMPs lose .2 points. If neither is in the attack state, then both IMPs lose .1 points. These costs and benefits define the utility of various outcomes in this environment, and the health of the agent at the end of a generation is used to assess its "fitness."

At the end of a generation, the algorithm assesses the health of the surviving IMPs. Dead IMPs (those whose health reached zero) are discarded. Five new IMPs are created for the next generation. Each new IMP is spawned from a previously surviving IMP, with probability in proportion to the health of that parent IMP. Then, the parameters of the new agent are randomly tweaked, or "mutated," such that every new IMP bears some close resemblance to an IMP from the previous generation, but is slightly altered.

For Experiments 1a–1c, three runs of evolution were simulated (each of 1000 generations), each with a randomized starting population. In these evolutions, all of the parameters were free to mutate after every generation. For Experiment 2, one run of evolution was simulated (of 1000 generations), with the memory and pathing resolutions held constant (at 150 and 25, respectively), only evolving the vision and decision making parameters. Further details about the precise nature of the evolutionary algorithm are included in Appendix B.

3.5 Experiment 1: Discriminating between evolved and non-evolved IMPs

Evolved IMPs exhibit behaviors selected for their adaptiveness, and are therefore more rational than non-evolved (randomly-parameterized) IMPs with respect to their simulated environment. In a series of experiments (1a–1c), I ask whether subjects can discriminate between evolved and non-evolved IMPs on the basis of their perceived intelligence (Exps. 1a–1b) or on the basis of which seem more convincingly human (Exp. 1c).

In the following section, I describe how this pool of evolved, more rational IMPs was generated.

3.5.1 Experiment 1: Evolution results

Across three 1000-generation evolutions (populated with 15000 total IMPs), 90.4% of IMPs survived to the next generation, with 31.4% of IMPs successfully collecting at least one piece of food.

The evolutionary algorithm I have used may be best considered as a local search that spends more time in more adaptive areas of the agent space, rather than a process which asymptotically converges to one or more stable equilibria. Consider the results of the three evolution runs, with respect to the IMP parameter determining the transition probability of an IMP remaining in the “gather” state, given that it is currently in this gather state (Fig. 3.1). In none of the three runs does this parameter converge to any particular value; rather, the respective populations oscillate across a wide range over the course of an 1000-generation simulation.

Nevertheless, the IMPs exhibit a preference to remain in the gather state, i.e. the evolving populations rarely dip below .25 on this parameter (.25 is the setting at which the IMP does not prefer remaining in this state versus transitioning to one of the other three). And if one examines a histogram of settings of this parameter for the 15000 total IMPs created and tested with this evolutionary algorithm, a setting of approximately .7-.8 for this parameter appears to have been the most successful. It makes intuitive sense that a relatively high setting for this parameter would be adaptive, because the

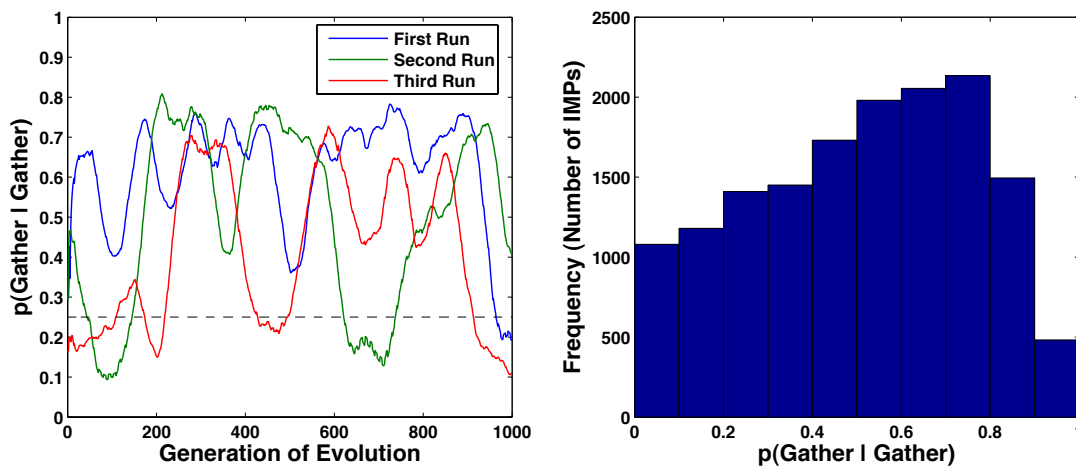


Figure 3.1: The left panel shows the evolution of the IMP parameter which determines the transition probability of an IMP remaining in the “gather” state, given that it is currently in this gather state (these time series are smoothed with a 10-generation trailing average). The evolution was done three times. IMPs without any preference for transitioning to the gather state, versus one of the other three possible states, would have this parameter set to .25, represented in this figure with a black dotted line. The right panel shows a histogram of the frequency of all possible values for this parameter, pooled across all three runs of evolution—15000 IMPs in total.

successful gathering of food—which entails dragging this food from one location to another, a task which can be sidetracked by brief excursions into non-gather states—is an adaptive behavior, given the cost structure of the environment.

Another parameter which showed one clear mode across the evolutionary population was for the IMP field of vision (Fig. 3.2). Wider fields of vision were more adaptive—and the wider, the better. However, for another vision parameter, governing the length of vision rays “cast” from the front of the IMP (Fig. 3.3), the distribution has two modes, likely indicative of two predominant visual strategies. Extremely short vision rays were not adaptive, but IMPs that cast rays of approximate 8 times their own body length (one-fifth the width of the environment) fared well, as did IMPs that cast much longer rays approaching the maximum.

The reason why longer vision rays do not necessarily dominate has to do with tradeoffs in the way the IMPs’ visual module works. Longer vision indeed allows for the IMP to detect objects in the environment at great distances. However, because the

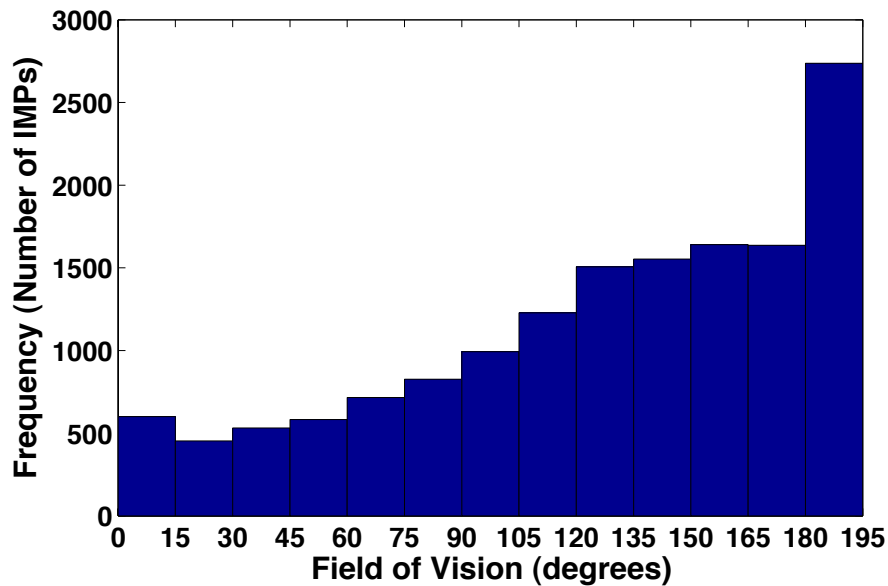


Figure 3.2: A histogram of the frequency of all possible fields of vision for the IMPs, pooled across all three runs of evolution.

IMP does not know how far away these objects are from the 1D information it collects, longer vision may indeed hinder the construction of an accurate mental map of the environment. Because the “walls” of the environment are considered to be obstacles, an IMP with vision rays of maximum length may also look around and believe itself to be surrounded by impassable objects, even if these objects are actually quite distant. Also, when it attempts to gather food, it may detect food at various distances and opt to pursue the more distant food, unaware that an easier meal was far closer. By contrast, an IMP with shorter vision rays may not be able to detect objects at great distances, but it can localize these objects better as it builds its mental map, and is less likely to waste time pursuing distant food of which it is unaware.

The optimal setting for a particular IMP parameter, however, not only depends on the utility structure of the environment, but also on how the other parameters are set. In other words, a particular “gene” on the IMP “chromosome” evolves in contexts both external (the environment) and internal (the other “genes”) to the IMP. I have already speculated that an IMP having relatively shorter vision ray length might be

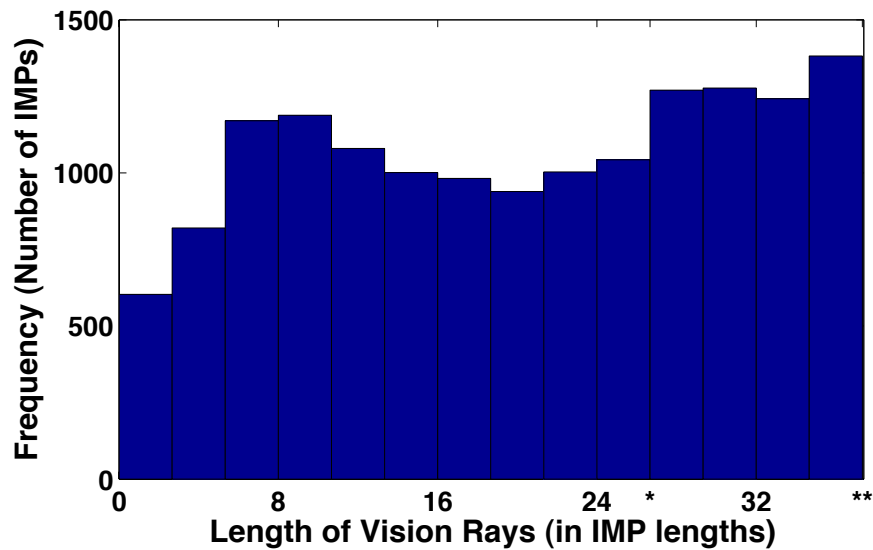


Figure 3.3: A histogram of the frequency of all possible vision ray lengths for the IMPs, pooled across all three runs of evolution. An asterisk (*) on the x-axis denotes the width of the environment, and a double asterisk (**) denotes the length of the environment's diagonal—the maximum value allowed for this parameter.

able to create a more precise mental map of its environment, leading to more adaptive behavior. Examining a joint histogram of vision ray length and memory resolution (Fig. 3.4) supports this claim. IMPs which have both relatively short vision ray length and near-maximal memory resolution appear in the evolved population with great frequency.

Analysis of the evolved IMPs could be continued in many different directions, but I will abridge it here, having established that IMPs sampled from this simulated evolution would not be sampled uniformly from the IMP strategy space. Rather, evolution has selected these IMPs, pushing them toward areas of this space which should be more adaptive—and more rational given these IMPs' environmental niche. In the following psychophysics experiments, I examine whether subjects are sensitive to the enhanced rationality of this pool of IMPs.

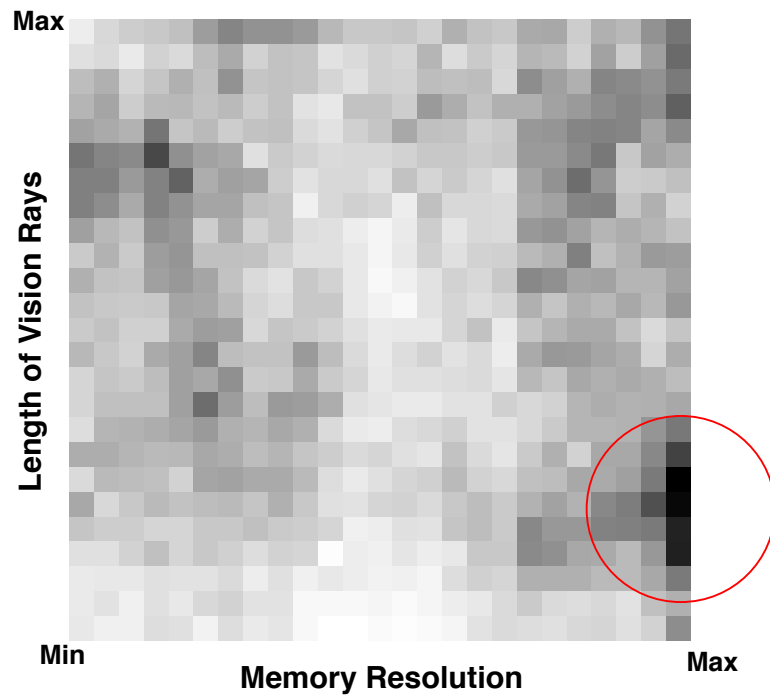


Figure 3.4: A 2D histogram of the joint frequencies of all combinations of parameter settings for memory resolution (x-axis) and vision ray length (y-axis). Darker bins contain a higher frequency of IMPs sampled from the evolution, across all generations of all three runs of evolution.

3.5.2 Methods

Subjects

Thirty-six undergraduate students (12 in each condition) in introductory psychology classes at Rutgers University participated in the experiment, and received course credit for their participation. Each experimental session lasted approximately 30 minutes.

Stimuli

Twelve sets of scenes, each set consisting of 25 scenes, were generated in advance. In each experiment (1a–1c), one of these 12 sets of scenes was presented to each of 12 subjects. Therefore, each of the 12 sets of scenes was shown to 3 different subjects, one in each experiment. The only difference in the experimental procedure, among these three experiments, was in the task instructions (reported below in the experiments' respective sections).

Each prerecorded scene was 60 seconds in duration, and was presented within a 950×950 pixel window, horizontally subtending approximately 31° of visual angle. Each scene was populated with 5 IMPs, 5 clusters of food (depicted as gray octagons), and 20 obstacles (4 large obstacles and 16 smaller obstacles, depicted as red squares). A screenshot of this 2D environment is shown in Figure 3.5. This environmental configuration (60 seconds, 5 IMPs, 5 food clusters, 20 obstacles) was identical to the evolutionary environment used to generate the evolved IMPs. All of the IMPs and objects were initialized at random locations in each scene, with obstacles being allowed to overlap with each other or the “walls” of the environment.

When the individual scenes were generated and inspected, scenes were thrown out if one of the two target IMPs “died” over the course of the scene. Scenes were also thrown out if—in a very rare circumstance—any of the IMPs in the scene ceased moving altogether without having died, presumably as a result of some degenerate set of incompatible parameter settings in its program.

Details about the IMPs environment, not covered here or in above sections, can be found in Appendix B.

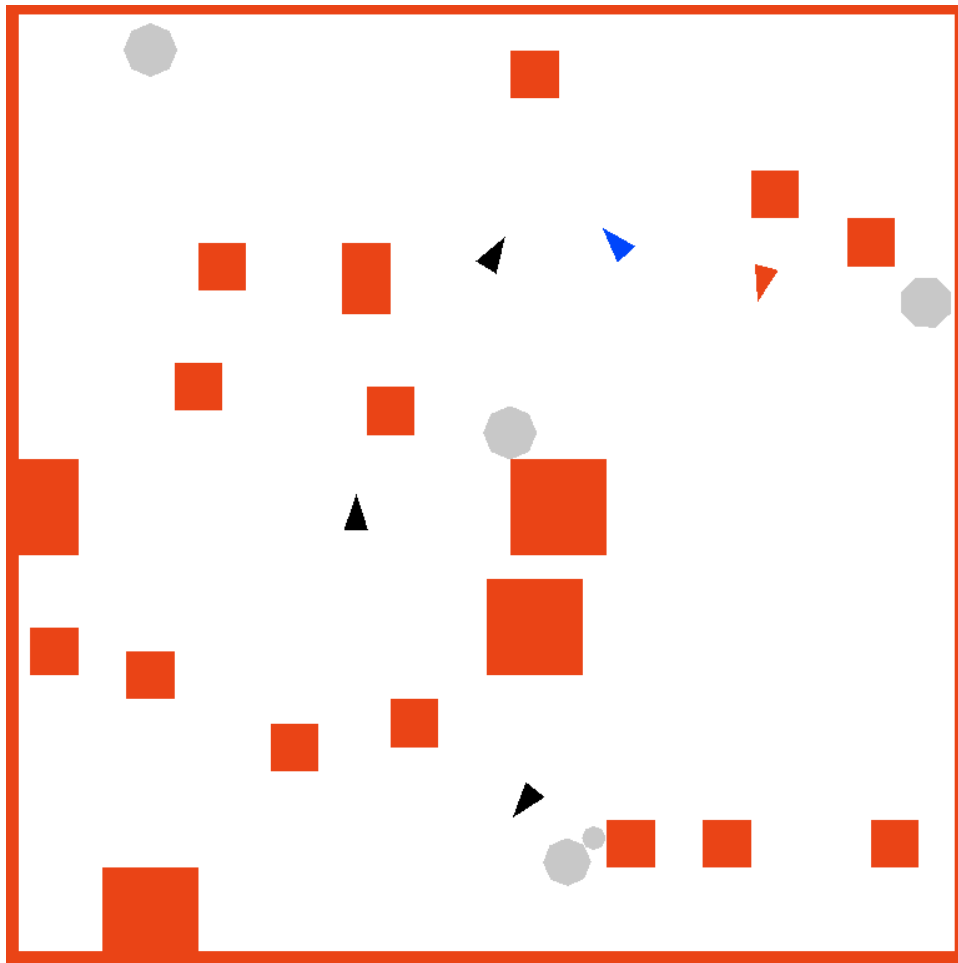


Figure 3.5: A screenshot of a scene from Experiments 1a–1c. The subject’s task was to choose between the red and blue agent on the basis of its intelligence (Exps. 1a–1b) or based on which was more likely to be controlled by a human player (Exp. 1c). Subjects responded on a 6-level scale, from “strongly red” to “strongly blue.”

Procedure

In each scene, three of the IMPs were colored black, while target IMPs were colored red and blue, respectively. After first reading a description of the environment—with specific emphasis on what was good or bad for the IMPs—subjects watched a sequence of 25 scenes. In the first scene, considered practice for the subject and not analyzed, all IMPs were randomly parameterized. In each of the subsequent 24 scenes, the black IMPs were randomly parameterized. Of the red and blue IMPs, one was randomly parameterized, and the other was an evolved IMP sampled from a random generation of simulated evolution. Whether the evolved IMP was colored red or blue was counterbalanced across trials.

To recall, three runs of evolution had been simulated to provide a pool of evolved IMPs for these experiments. In each experiment, 4 subjects (one-third) viewed evolved IMPs sampled from each of the respective simulations.

At the end of each scene, the subject was instructed to indicate, on the keyboard, which IMP (red or blue) had behaved more intelligently (in Exps. 1a-1b), or was more likely to have been controlled by a human player (Exp. 1c). This subject was allowed to respond on a 6-level scale, from “strongly red” to “strongly blue.” The experiment moved on to the next trial after the subject made a response.

3.5.3 Experiment 1a: Basic discrimination

Because the “random” agents are drawn uniformly from the parameter space, and the evolved agents are sampled from the same space, but tend to be concentrated in more adaptive regions, there is indeed overlap between the two IMP distributions that sets an upper limit on discrimination. In this sense, the task sets up as a signal detection problem, with the “signal” distribution being the evolved IMPs and the “noise” distribution being the randomly parameterized IMPs. I ask whether subjects can discriminate between these signal and noise distributions on the basis of their perceived intelligence.

The experimental instructions included a plain English explanation of what

was good or bad for IMPs as they interacting within the scenes (see Appendix C). This explanation was consistent with the actual programmed cost structure of the environment, determining IMP “health.”

Subjects could respond on a 6-level scale, from “strongly red” to “strongly blue”, depending on how confident they were that one of the two IMPs had behaved more intelligently than the other. Color was randomized and counterbalanced with respect to whether the red or blue IMP was the target (evolved) or lure (randomized) on a given trial. If I assign numbers to the ratings, from 1 (strongly believe the random IMP was more intelligent) to 6 (strongly believe the evolved IMP was more intelligent), an average rating of 3.5 represents chance performance, or no discrimination.

Results

On average, subjects provided a rating of 4.12, significantly greater than this chance 3.5 level ($t[11] = 7.60, p < .001$). Therefore, they rated evolved agents to be significantly more intelligent than non-evolved agents. Because subjects provided a range of discrete ratings (i.e. this was not simply a 2AFC task), and because this experiment lends itself to a signal detection framework (as explained above), I additionally constructed ROC curves for subjects’ responses, as a further visualization of their discrimination ability on this task. A caveat applies: Because targets and lures were presented simultaneously rather than on separate trials, a rating of 6 for the target IMP is assumed to be automatically indicate a rating of 1 for the lure IMP (and vice versa), which artificially constrains the ROCs to be symmetrical. Therefore, these are not bona fide ROC curves, but are nevertheless illustrative of general trends in subject responses. As shown in Figure 3.6, every subject showed discrimination above chance, judging the evolved IMPs to be more intelligent.

On what basis did subjects make their discrimination? During debriefing, nearly all subjects indicated that an IMP’s proficiency in collecting food was a main criterion for their decision making. An IMP’s ability to navigate its environment without either getting stuck at obstacles, or thrashing back and forth without making progress in any direction, was a secondary criterion expressed by many subjects.

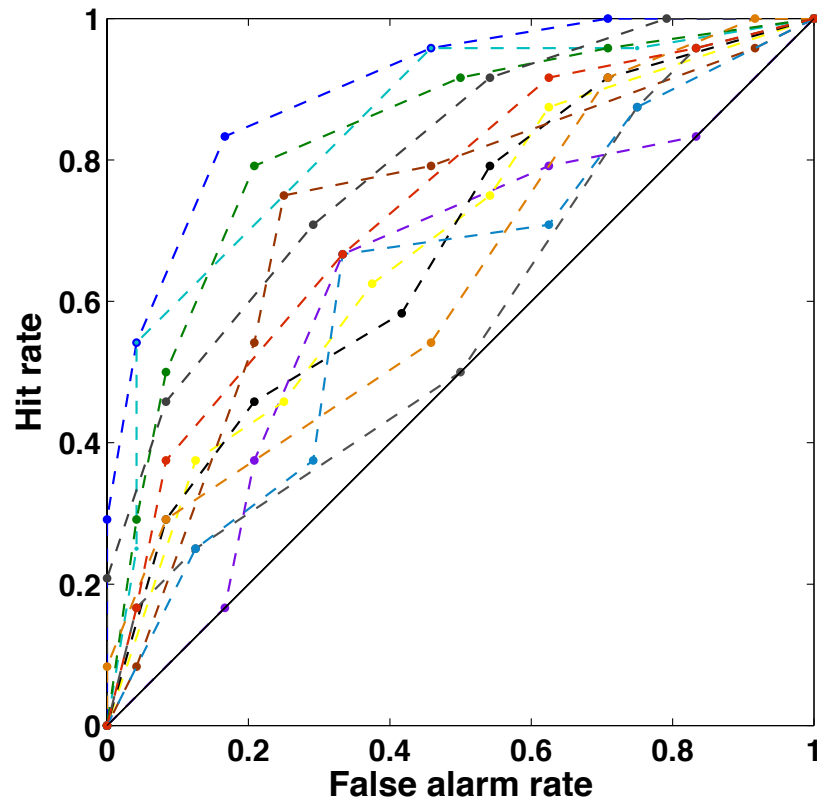


Figure 3.6: ROC curves for all 12 subjects in Experiment 1a, each presented in a different color. Across various confidence thresholds (subjects responded on a 6-level scale), a “hit” is a scene for which the subject selected the evolved IMP, and a “false alarm” is a scene for which the subject selected the non-evolved IMP. The more “bowed” the ROC curve is above the line of no discrimination (represented with a solid black line), and the greater the area under the curve, the better the subject’s discrimination.

This suggests that the evolution procedure I used selected for productive behaviors (collecting food) and selected against counterproductive behaviors (navigating the environment poorly), and that subjects were able to detect this greater efficiency in the evolved IMPs.

One of the primary IMP traits enabling an IMP to successfully gather food is a greater tendency to remain in the *gather* state, reflected in one of its transition probability parameters: $p(\text{Gather}|\text{Gather})$. As displayed in the empirical cumulative distributions along this parameter (Fig. 3.7, red lines), IMPs selected by subjects as more intelligent indeed tended to be higher on this parameter.

By Kolmogorov-Smirnov (K-S) test, the selected and rejected IMP samples are drawn from significantly different distributions, with respect to this $p(\text{Gather}|\text{Gather})$ parameter ($p < .001$). Selected IMPs also exhibited a much stronger tendency to transition from *flee* to *gather* ($p[\text{Gather}|\text{Flee}]; p < .001$).

Ability to navigate the environment is less straightforwardly related to any particular IMP parameter; different combinations of parameters governing memory resolution, path planning, and vision interact to produce more or less effective navigation. Observed navigation ability is also a noisy reflection of the IMP programming; IMPs find themselves in more or less challenging circumstances depending on where they are randomly placed among obstacles in the environment. For these reasons, the selected and rejected IMP distributions were quite similar with respect to these navigation-related parameters, analyzed in isolation. They differed along the field of vision parameter ($p < .05$, Fig. 3.8). However, this difference was small, and given the strong tendency for evolved IMPs to be high on this parameter, this observed difference in Experiment 1a may merely reflect this parameter's high correlation with a more robust, primary independent variable: the evolved status of the IMP.

I treat this evolved status as a categorical variable here, but each IMP was in fact sampled from a different generation of its respective evolution. Therefore, one might expect that IMPs sampled from later generations of evolution would appear more intelligent than those sampled from earlier stages. On the other hand, in the above analysis of the evolutionary dynamics, I observed that IMP parameters did

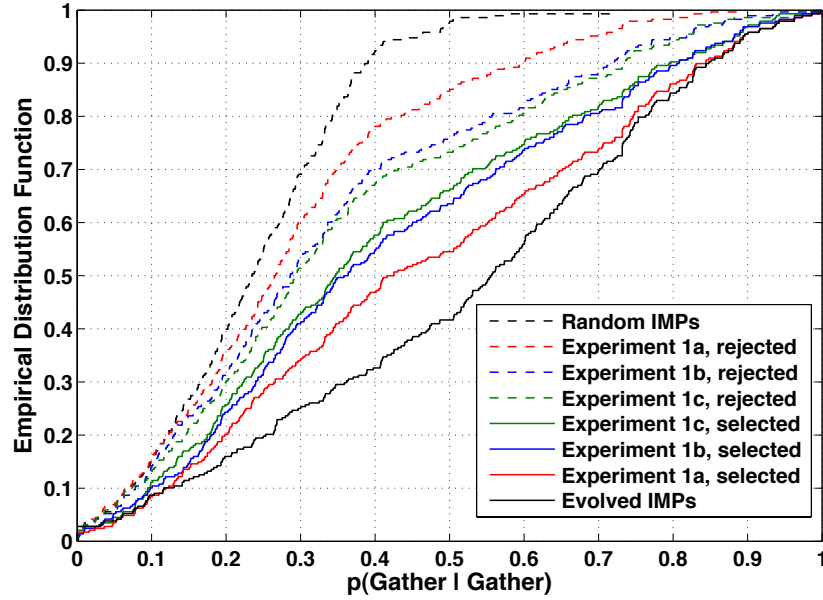


Figure 3.7: Empirical cumulative distribution functions, along the parameter governing the probability of remaining in the *gather* state, for various IMP populations.

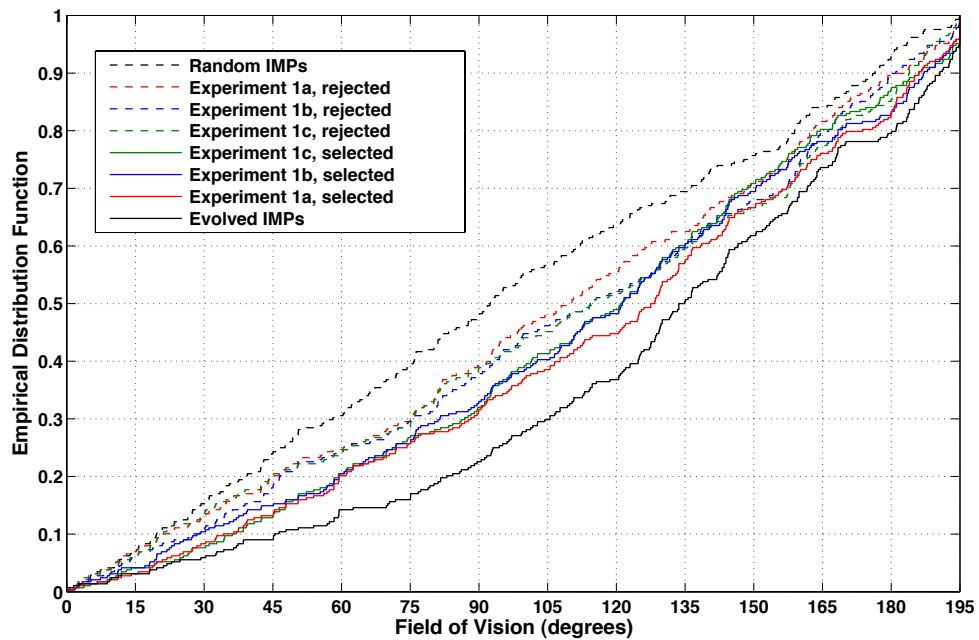


Figure 3.8: Empirical cumulative distribution functions, along the field of vision parameter, for various IMP populations.

not tend to converge over time, but rather oscillated frequently over the course of a 1000-generation evolution, only exhibiting general tendencies in the aggregate.

Over all evolved IMPs presented to subjects, there was no correlation between generation of evolution and intelligence rating ($r[286] = -.050, p = .399$). As shown in Figure 3.9, perceived intelligence peaked for IMPs sampled from between generations 300 and 400, and was lowest for IMPs sampled between generations 900 and 1000. A reinspection of Figure 3.1 may explain this result. As the three runs of evolution oscillated between values on one critical IMP parameter ($p[\text{Gather}|\text{Gather}]$), each of the three evolutions coincidentally oscillated toward a high value between generations 300 and 400, and all three randomly swung low between generation 900 and 1000. All of the analyses undertaken thus far appear to indicate that the simulated evolution indeed produced IMPs that were, on average, perceived to be more intelligent than randomly-generated IMPs, but that this evolution itself wandered noisily around the IMP space across generations.

3.5.4 Experiment 1b: Mismatched description

In Experiment 1a, subjects perceived evolved IMPs to be more intelligent than non-evolved IMPs. Subjects were provided with a description of the associated costs and benefits of various things that could happen to the IMP, and this description was consistent with the actual utility structure of the environment in which the IMPs evolved, and in which the IMPs appeared to subjects. In Experiment 1b, subjects were instead given a misleading description of what was good or bad for the IMPs (see Appendix C). Subjects were instructed that collecting food was slightly beneficial to the IMP, but that pursuing food was “usually more trouble than it is worth.” The payoff matrix for IMPs colliding with other IMPs was presented in a manner distorted from the true programmed payoff matrix.

In this experiment (and in Experiment 1c), each of 12 subjects saw a set of 24 scenes which were identical to those seen by one corresponding subject in Experiment 1a, and in the same order. Therefore, critically, all of the low-level perceptual cues intrinsic to the stimuli were precisely controlled. The only difference between the

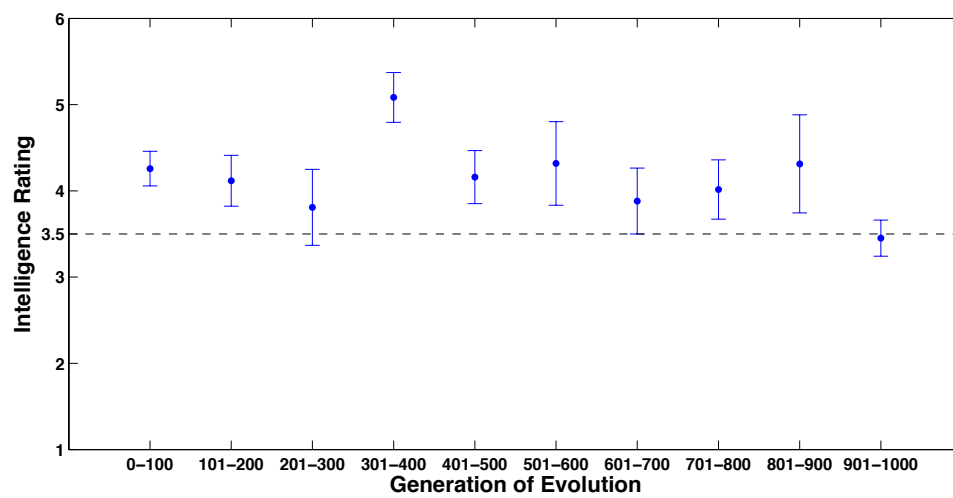


Figure 3.9: Each evolved IMP was sampled from a random generation of a 1000-generation simulated evolution. Here, IMPs are placed into bins according to the generation of evolution from which they were sampled. Whereas subjects consistently rated evolved IMPs to be more intelligent (i.e. gave a rating greater than the chance level of 3.5, shown here with a dashed black line), IMPs sampled from later versus earlier generations of evolution do not appear to be perceived any differently with respect to intelligence.

experience of a subject in Experiment 1a and his counterpart in Experiment 1b was the task instructions.

Results

8 out of 12 subjects produced responses that were significantly positively correlated with their counterparts in Experiment 1a. The average pair of corresponding subjects from Experiments 1a and 1b produced responses that were positively correlated ($M = .401, t[11] = 4.66, p < .001$). As in Experiment 1a, the average subject rated evolved IMPs to be significantly more intelligent than randomly-generated IMPs ($M = 3.78, t[11] = 2.30, p < .05$). However, this effect was not as strong as in Experiment 1a ($t[22] = 2.22, p < .05$), as demonstrated in Figure 3.10. Subjects were therefore sensitive to these task instructions when assessing the intelligence of IMPs; the evolved IMPs seemed intelligent, but not as intelligent as they would have seemed given a description of the demands of the environment more consistent with that of the actual environment in which the IMPs were evolved.

Certain IMP qualities, such as the ability to navigate the environment well, were likely to appear intelligent irrespective of more specific details about the cost structure of the environment, which partially explains the correlation among subject responses across experiments, and why evolved IMPs appeared more intelligent in both cases. While less sensitive to the tendency of the IMPs to collect food, they were not totally insensitive to it (Fig. 3.7). This is consistent with a fascinating account reported by several subjects during debriefing: Even though subjects reported being keenly aware that food was of negligible value, the IMP act of successfully pursuing and collecting food just seemed irresistibly intentional or elegant, and it was difficult to suppress the impression that an IMP exhibiting this behavior was intelligent—especially when the other probe IMP did not seem to be doing anything particularly intentional.

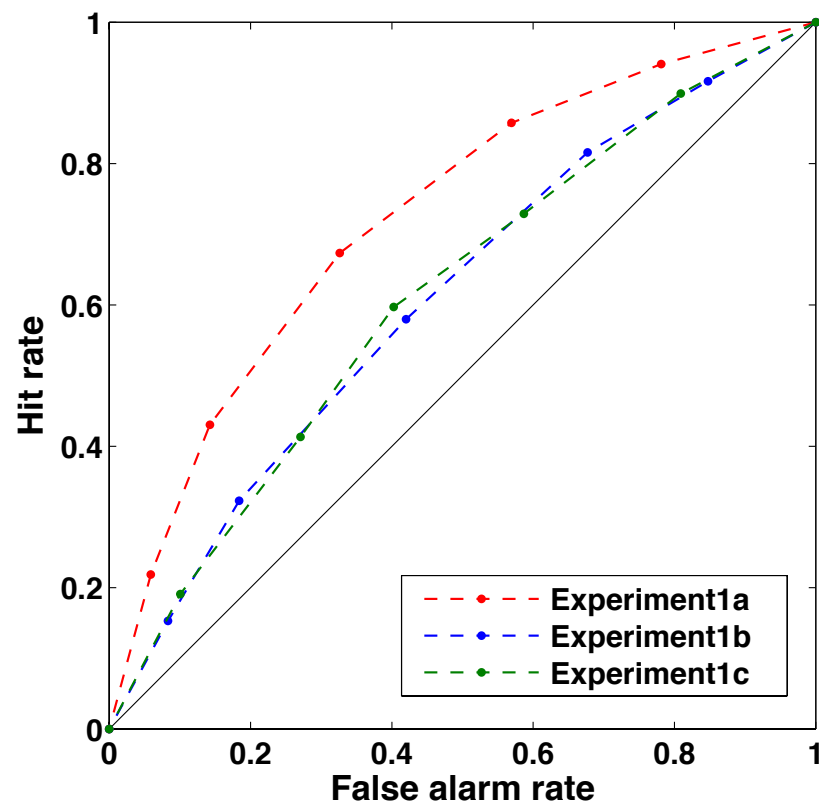


Figure 3.10: ROC curves for Experiments 1a–1c, with false alarm rate plotted on the x-axis and hit rate plotted on the y-axis.

3.5.5 Experiment 1c: “Turing test”

In Experiment 1c, subjects were given an accurate description of what was good or bad for the IMPs, the very same description as was provided in Experiment 1a. As in Experiment 1b, each of 12 subjects saw a set of 24 scenes which were identical to those seen by a subject in Experiment 1a, and in the same order, controlling for all of the low-level perceptual cues intrinsic to the stimuli. Again, the only difference between the experience of a subject in Experiment 1a and his counterpart in Experiment 1c was the task instructions (see Appendix C). The motivation for this experiment was to ask whether evolved IMPs (compared to non-evolved IMPs) would appear to subjects to behave more convincingly as though they were controlled by humans, and to further ask whether subjects’ expectations of a rational agent and an agent controlled by a human were the same.

Subjects were told that some of the IMPs were robots controlled by a computer program, and some were controlled by human players. To make this cover story more plausible, subjects were told that the human player would not be able to see the entire environment from above (as the subject could), but could only see this environment from the perspective of his triangular avatar. Furthermore, the subject was told that seeing objects at a distance was often difficult for the human player, as was moving around the environment fluidly. This cover story was effective; during debriefing, few subjects reported any suspicion that, in fact, neither of the probe IMPs had actually been controlled by a human player (the reactions ranged from “I had some sneaking suspicion all along.” to “My mind is blown!").

The task, in this experiment, was to choose the IMP which was more likely to have been controlled by a human player. Subjects were therefore probed not on the intelligence (i.e. rationality) of the IMPs, per se, but on how convincingly human they appeared to be. It is important to note that most simulation theory accounts of this task would not consider these tasks to be distinct: We are always implicitly asking “What would I do?”, bootstrapping the same internal simulation process. Dennett (1987) presents this as a main difficulty in testing the predictions of his account of

theory of mind, versus simulation accounts; since the self, other human agents, and the abstractly-rational agent should all be expected to behave similarly in similar situations (namely, according to what the observer construes to be the rational choice), how can one find experimental evidence that supports one theoretical account and not the other? I propose that this experiment may provide such an opportunity, as we shall observe whether subjects had different expectations of “intelligent” versus “human” behavior.

Results

In Experiment 1c, 4 out of 12 subjects produced responses that were significantly positively correlated with their counterparts in Experiment 1a. The distribution of correlations was marginally positive ($M = .251, t[11] = 2.06, p = .064$). Furthermore, the average subject rated evolved IMPs (versus non-evolved IMPs) as being more likely to have been generated by a human (mean rating = 3.83, $t[11] = 3.17, p < .01$). This is a noteworthy result in itself; the evolved IMPs passed this quasi-Turing Test, and were more convincingly human than their randomly-generated counterparts. This effect, however, was smaller than that observed in Experiment 1a ($t[22] = 2.20, p < .05$). I therefore conclude that subjects based their decisions on at least partially different criteria.

The magnitude of the effect was quite similar to that exhibited by subjects in Experiment 1b (Fig. 3.10). One, therefore, might be tempted to conclude that subjects in Experiment 1b and 1c were answering on similar bases. Yet only 2 out of 12 subjects produced responses that were significantly positively correlated with their counterparts in Experiment 1b, and the responses of the average pair of subjects from these two experiments were not significantly correlated ($M = .110, t[11] = 1.23, p = .243$). This suggests that while both groups of subjects showed a weak preference for the evolved IMPs, they were likely responding according to different criteria.

Therefore, in Experiments 1a–1c, subjects viewed the same scenes, but employed three dissociable (though overlapping) sets of standards for discriminating

among the IMPs. In Experiments 1a and 1b, this strategy involved the subject invoking his theory (or model) of how an intelligent (or rational) IMP should behave, given different assumed costs and benefits of various behaviors. In Experiment 1c, subjects were tasked with invoking their personal models for how a *human* player would behave in this environment.

At debriefing, most subjects reported that they expected the human player to more effectively collect food and navigate the environment (i.e. exhibiting the behavior of a rational agent, as in Experiment 1a). But others adopted very different strategies. Some assumed that all of the non-probe IMPs were robots (there were always three other agents inhabiting the scene), and therefore chose the IMP which behaved least like these presumed automata. At least one subject reported that he, in fact, expected that the robots would perform better in this environment, and thus selected the less competent IMP. And one particularly distrustful subject indeed reported that he thought the human player would be trying to trick him into thinking he was a robot, and therefore adopted the strategy of selecting the more robotic-seeming IMP!

Subjects' responses in this task were correlated, if only weakly, with responses made only on the basis of perceived intelligence with respect to the demands of the environment. Yet, it seems clear that, regarding subjects' expectations of how other humans behave, they are willing and able to augment their "human" model with a great many considerations, not all of which are directly related to a straightforward conception of rationality, and not all of which are consistently applied subject to subject.

3.5.6 Discussion

In Experiments 1a–1c, subjects saw scenes which had the same perceptual content but could be interpreted differently depending on what subjects were told about the nature of the IMPs and their environment. Subjects always attended to two IMPs, one of which was sampled from simulated evolution, and the other which was randomly-parameterized (non-evolved). In all three experiments, subjects were more likely to select the evolved IMP, but this effect was strongest when the task was to select the more intelligent IMP, and they were given a description of the demands of the environment

that matched those of the environment in which the IMPs were evolved (Experiment 1a). In Experiment 1b, subjects were misled about what behaviors were likely to be good or bad for the IMP, but they still showed a weak preference for the evolved IMPs. The same was true in Experiment 1c, in which subjects selected the IMP which they believed was more likely to have been controlled by a human player. We conclude that the expectations subjects have for the behavior of a human player and a rational agent may be different, and that what is construed as rational behavior will depend on assumptions about the cost structure of the environment.

3.6 Experiment 2: The inference of mental states in evolved (vs. non-evolved) agents

Here I present the most critical experiment of this thesis. I use the IMPs environment to empirically test a prediction that follows from a specific perspective on theory of mind: The more rational the agent, the easier it should be for subjects to understand its mental states and processes. IMPs may not have bona fide minds, but they do have ground truth goal states (they may *attack*, *explore*, *flee*, or *gather*) which change dynamically over the course of a scene and are meant to be analogous with intention. They also have simulated vision, the capabilities of which can be inferred by the subject on the basis of observed behavior. As subjects watch a target IMP which is either evolved or non-evolved, they are asked to continually estimate the goal state of this IMP (by the same method employed in Chapter 2), and after the scene is finished, provide an estimate of how far they think the IMP can see.

3.6.1 Experiment 2: Evolution results

In Experiment 2, subjects' were instructed to make inferences about the intentional states of target IMPs, and their visual ability. The simulated evolution which generated IMPs for this experiment allowed for the IMPs' state transition probabilities to be evolved, along with their vision modules. Memory resolution and path planning resolution were kept constant (at 150 and 25, respectively), at settings for which IMPs

could be expected to navigate obstacles in the environment well, given adequate vision capabilities.

One 1000-generation run of evolution was simulated (populated with 5000 total IMPs). 91.1% of IMPs survived to the subsequent generation, with 25.8% of IMPs successfully collecting at least one piece of food. For some parameters, examination of the traits of IMPs sampled from this evolution revealed similar patterns of preferred parameter settings when compared with the population generated for Experiments 1a–1c. The distribution of vision ray lengths, for instance, was similarly bimodal, and with these modes at the same values. On the other hand, whereas the Experiment 1 IMPs showed a preference for very wide fields of vision (see Fig. 3.2), the newer population of IMPs show a mode at approximately 60–75°, with very wide fields of vision being much less preferred (Figure 3.11).³ These disparate results illustrate that what makes for an adaptive setting along one parameter will depend critically on the possible settings of other parameters—in this evolution, memory and path planning were constant, whereas in the previous evolutions, the memory and path planning resolution parameters could take on any possible value across a wide range.

3.6.2 Methods

Subjects

Ten adult subjects between the ages of 22 and 37 participated in the experiment. Subjects received \$7 for their participation, which lasted approximately 40 minutes.

Stimuli

Each subject viewed a different set of 27 scenes, generated in advance. As in Experiments 1a–1c, these scenes were 60 seconds in duration and contained 5 IMPs, 5 food clusters, and 20 obstacles. The displayed scenes were of the same size as reported in

³In this histogram, the 0–15° bin appears to be the most popular setting for this parameter. However, because of the hard limits set at each end of the allowable range for each parameter during evolution, the resulting histograms tend to artificially “pile up” at both extremes. Thus, this apparent mode is probably spurious in comparison to the mode at 60–75°.

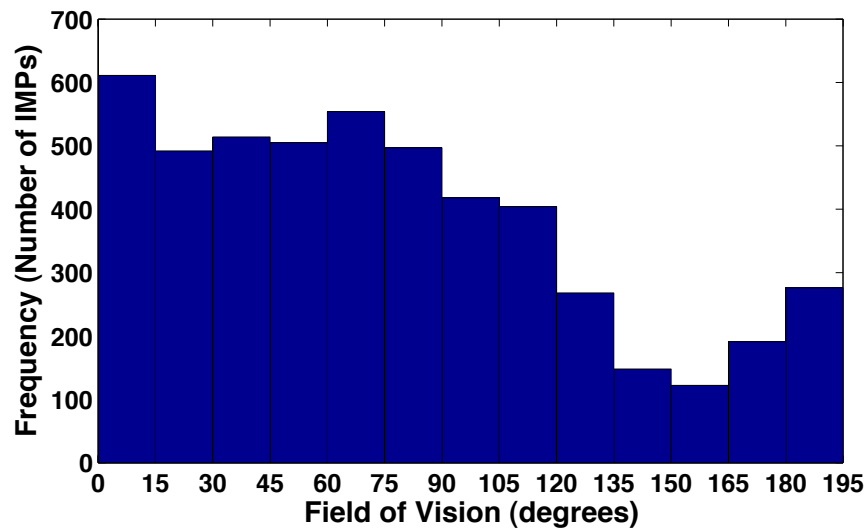


Figure 3.11: A histogram of the frequency of all possible fields of vision for the IMPs, over 1000 generations of evolution performed to generate stimuli IMPs for Experiment 2.

the previous experiments.

Procedure

The first two scenes served as training for the subject, and were each populated with 5 randomly-parameterized IMPs. In these training scenes, each IMP's true goal state was reflected in its color, and changed with varying frequencies over the course of the scene. Subjects were invited to guess which of the four possible IMP states—*attack*, *explore*, *flee*, or *gather*—corresponded to which colors. In general, they were accurate in their initial assessments, but these training scenes provided the opportunity to get a concrete sense of what the behaviors corresponding to these four goal states looked like, and what was meant by the “state” of the IMP.

These training scenes also provided an opportunity for the subject to become comfortable with the demands of the task, as they practiced making keypresses and using the mouse. After these two training scenes, the subject saw 25 scenes (the first of which was treated as practice and not analyzed), in which one target IMP was colored blue, and the other 4 were colored black. Subjects were instructed to pay attention to

the blue 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 state⁴.

At the end of a scene, the display would freeze at the last frame, cuing the subject to make one additional judgment. The subject was instructed to estimate how far he believed the target IMP could see, and was told that this distance could range from not extending past the tip of the IMP's "nose" (i.e. a blind IMP) all the way to a maximum distance equivalent to the length of the environment's diagonal—This was, in fact, the true range of possible vision lengths for the IMPs.

The subject indicated his estimate with two mouse clicks (first the left button, then the right button). Subjects were advised that an intuitive method for performing this task was to first click the center of the target agent, and then make a second click at a location as far from the agent as it could see. However, if the subject believed that, due to the position of an IMP in the environment and a high estimated vision length, this second click would necessarily fall outside of the environment, then clicking any two arbitrary points in the environment would suffice. After the subject provided two mouse clicks, the experiment moved on to the next trial.

In half of the 24 randomly-shuffled trials, the blue target IMP was randomly parameterized. In the other half of trials, the IMP was sampled from a random generation of simulated evolution. Whether the IMP was evolved or non-evolved thus served as the primary independent variable, and we measured the effect of this variable on the inferences subjects made about the IMPs' goal states and vision capabilities.

3.6.3 Results: Goal state inference

The average subject correctly identified the goal state of non-evolved (randomly-parameterized) IMPs 33.4% of the time, if one considers the beginning portion of a

⁴This aspect of the subject's task was identical to that of the previous experiments reported in Chapter 2.

trial, during which the subject had not yet responded, to be “incorrect.” If one throws out these portions of trials, average subject accuracy was 36.0%. In either case, accuracy was only slightly (though significantly) above chance performance of 25% ($t[9] = 4.33, p < .01$ or $t[9] = 5.37, p < .001$, depending on the scoring criterion).

For the IMPs sampled from simulated evolution, the subjects made correct inferences 39.9% (or 43.2%, by the alternate criterion) of the time, performing significantly better than when observing the non-evolved IMPs ($t[9] = 3.27, p < .01$ or $t[9] = 3.51, p < .01$).

Table 3.1: Confusion matrix, hit rates, false alarm rates, and d' for each of the IMPs' goal states, for non-evolved (randomly-parameterized) IMPs (pooled across subjects). Mean proportion of IMP time spent in each state is in parentheses. Overall response rates are included in the bottom row.

Actual State	Subject Response					HR	FAR	d'
	None	Attack	Explore	Flee	Gather			
Attack (.23)	.048	.112	.544	.108	.188	.112	.073	.236
Explore (.34)	.090	.073	.610	.054	.174	.610	.512	.248
Flee (.17)	.063	.092	.548	.112	.186	.112	.080	.184
Gather (.27)	.059	.062	.462	.090	.328	.328	.181	.466
	.068	.082	.545	.086	.220			

Table 3.2: Confusion matrix, hit rates, false alarm rates, and d' for each of the IMPs' goal states, for IMPs sampled from the simulated evolution (pooled across subjects). Mean proportion of IMP time spent in each state is in parentheses. Overall response rates are included in the bottom row.

Actual State	Subject Response					HR	FAR	d'
	None	Attack	Explore	Flee	Gather			
Attack (.18)	.057	.145	.431	.101	.266	.145	.076	.373
Explore (.32)	.109	.071	.529	.108	.184	.529	.361	.428
Flee (.12)	.065	.135	.364	.151	.285	.151	.089	.317
Gather (.38)	.049	.062	.327	.066	.496	.496	.227	.737
	.071	.089	.414	.096	.330			

These results are consistent with my hypothesis that, when observing an agent whose behavior better conforms to subjects' expectations of rationality, the intentions of this agent will be easier to infer. However, there are at least two plausible alternate explanations for why accuracy would be better for these evolved IMPs: 1) Certain states are intrinsically easier to discriminate than other states, and these states occur

more frequently in the evolved case, and 2) The distribution over the possible IMP states has less uncertainty (i.e. lower entropy) in the evolved case, and therefore the state of the IMP will be inherently easier to predict on the basis of a modified response strategy (i.e. responding more frequently for *a priori* more likely states, without an increase in actual discrimination for these states).

These are reasonable hypotheses, as their premises are both empirically true. The goal state for which subjects' discrimination was best is the *gather* state, and IMPs were in this state for a larger proportion of time in the evolved IMP case. And, the distribution over the possible goal states had a higher entropy for the randomly-parameterized IMPs (1.95) than the evolved IMPs (1.87). However, neither of these alternate hypotheses predict better *discrimination* of the individual states for the evolved IMPs. Better accuracy for the evolved IMPs would wholly result, under these hypotheses, from either a greater tendency toward easier-to-discriminate states, or from a modified response strategy. As shown in Tables 3.1 and 3.2, however, discrimination for *each* of the four IMP goal states was better when subjects were observing the evolved IMPs, a result which cannot be explained under these alternate hypotheses.

To further illustrate the point, the evolved IMPs are less likely to be in either the *attack* or *flee* states, but subjects are actually slightly *more* likely to respond "attack" or "flee" when observing them, and are also more accurate in these responses. Improved performance of this nature cannot be attributed to mere sensitivity to base rates, but reflects truly enhanced discrimination.

Another alternate hypothesis might be that evolved agents might be more likely to remain in a given state (rather than change state), and that this might make the task of inferring the state easier for subjects. That is, if by the time a subjects is prepared to make a response after observing the behavior of an IMP, the state of the IMP has already changed, accuracy might be severely limited by reaction time, and this would be exacerbated when observing an IMP whose state changes more frequently. There is little evidence for the premise of this alternate hypothesis, however. Empirically, the average evolved IMP remained in its state from one frame to the next 97.6% of the time, while the average non-evolved IMP did so 97.31% of the time. For the average

subject, the difference between evolved and non-evolved IMPs along this metric was negligible, and not statistically significant ($t[9] = 0.58, p = .571$).

3.6.4 Results: Vision length estimation

Intentions and beliefs are two types of mental states which may be inferred from an agent's behavior. When subjects were probed about their subjective impression of IMP goal states, this task served as a functional definition for the inference of intention ("What are the IMPs trying to do?"). Similarly, it was hoped that asking subjects about how far they believed the IMPs could see would be an analogous task to the inference of beliefs ("What is the quality of IMPs' knowledge of their environment?"). After viewing each scene, the subject was instructed to estimate how far the target IMP was able to see using two mouse clicks.

For the non-evolved IMPs, the actual lengths of IMP vision rays were drawn randomly from a uniform distribution, from 0 to the length of the diagonal of the environment. Empirically, these non-evolved IMPs had an average vision ray length of 21.1 IMP lengths (a distance approximately 79% of the width of the environment and 57% of the diagonal). The distribution of evolved IMPs' vision lengths was strongly bimodal, and the mean setting of this parameter was only slightly higher (22.9 IMP lengths) than in the non-evolved case (see Fig. 3.12 for both empirical cumulative distributions).

The most striking empirical result was subjects' vast underestimation of the actual IMP vision lengths. On average, IMPs' vision rays extended 22.0 IMP lengths, but subjects, on average, estimated that IMPs could only see a distance of 4.5 IMP lengths. Furthermore, there was no correlation between these estimates and the actual IMP vision lengths—if anything, the overall correlation was weakly negative ($r[238] = -.099, p = .128$). Additionally, while the underlying ground truth distributions for this IMP parameter were uniform and bimodal, respectively (for the non-evolved and evolved IMPs) subjects' responses were gamma-distributed under both conditions (Fig. 3.12)

From this overall pattern of data, it is clear that the way the IMP vision actually

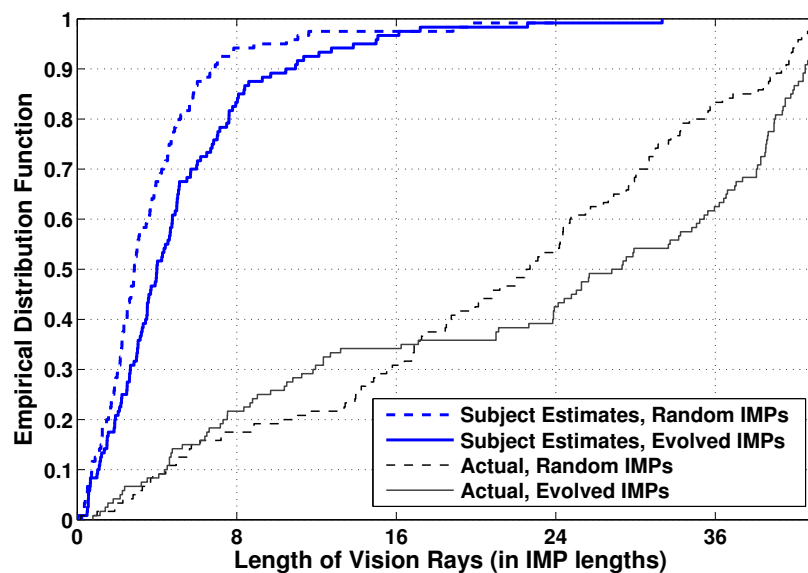


Figure 3.12: Empirical cumulative distribution functions for estimated and actual IMP vision ray lengths in Experiment 2.

works was completely opaque to subjects, and that subjects were (at best) completely insensitive to the underlying parameter governing IMP vision ray length when making their estimates. Perhaps they came to the experiment with strong prior expectations that shorter IMP vision lengths were more plausible. Or, perhaps subjects intuited that if an IMP could “see” a certain distance, this meant that it could, at minimum, detect objects directly in front of it at that distance. This is not necessarily true for all IMPs, as their vision 1) loses resolution as large distances, as the space between rays increases, and 2) requires two adjacent rays to return the same color in order to detect any obstacle, food, or agent.

Another expectation subjects might have had about IMP behavior was that if an IMP could see an object at a certain distance, it would react to that object, and in a rational way. But due to the cognitive architecture of the IMPs—they are finite state automata which transition among goal states probabilistically—they will often turn toward or away from an object in the environment randomly. This can hardly be what subjects expect from a rational agent: If the subject assumes a rational IMP, and there

is food in front of an IMP at a certain distance, and the IMP does not seem to react to it, the subject very well may infer that the IMP *could not see it*. To the subject, this may seem a more plausible inference compared to the alternative—and true—account: the IMPs' decision to *attack, explore, flee, or gather* is actually generated by a partially stochastic process.

In the experiments presented in Chapter 2, IMPs had slightly-different decision making schemes (see Appendices A and B), transitioning among goal states according to four separate transition probability tables, each reflecting a qualitatively different circumstance the IMP might encounter. It should be noted that subjects' overall accuracy for inferring the goal states of IMPs was higher in this previous study, perhaps reflecting that their decision making was more intuitive to subjects, although we cannot be sure of this explanation as this was not the only difference between the two virtual environments and their resident IMPs.

Any behavior can be rationalized, i.e. construed as rational given some underlying set of beliefs, a set of possible actions, and a utility function. But if the set of possible actions is known, the utility function is given, and the behavior is assumed to be rational, then it may be the estimate of the beliefs which will change depending on agent's level of rationality. Given an observed behavior that appears to be irrational given a veridical set of beliefs, one must either infer that the agent is behaving irrationally, or that the agent holds non-veridical beliefs about the environment.

A reasoning process like this may explain why subjects so grossly underestimated IMP vision length, because this scenario occurs frequently due to the stochastic nature of IMP behavior. Yet, for the evolved IMPs, which are more rational than their randomly-parameterized counterparts, and more likely to pursue food if they can see it, this scenario should play out less frequently. Therefore subjects may, on average, infer that the evolved IMPs can see further.

I have already established that the actual IMP vision ray length did not have any positive correspondence with subjects' estimates. For the average subject, the correlation between actual IMP vision ray length and the subjects' estimates was actually

negative for randomly-parameterized IMPs ($M = -.184, t[9] = 2.63, p < .05$). The average correlation for the evolved IMPs was also weakly negative ($M = -.077$), though not significantly so ($t[9] = 0.78, p = .454$). Therefore, if one were to observe a strong difference between subjects estimates of non-evolved versus evolved IMP vision lengths, this difference would be unlikely to be attributable to overall actual differences in programmed IMP vision ray length, which were quite small to begin with (evolved IMPs could only see, on average, 8.5% further). Subjects were insensitive to this IMP parameter.

Subjects estimated that evolved IMPs could see 37.2% further than their non-evolved counterparts. The average subject estimated the vision length of an evolved IMP to be 5.17 IMP lengths, significantly longer than for random IMPs (3.77 IMP lengths; $t[9] = 3.28, p < .01$). As mentioned above, in both cases, subjects' estimates were approximately gamma-distributed; overall, the best gamma fit for random and evolved IMPs, respectively, obeyed the same shape parameter k (1.58 vs. 1.60), but differed on scale parameter θ (2.39 vs. 3.24)—Because the mean of a gamma distribution is $k\theta$, the difference in means between the two conditions largely reflects the difference in this θ parameter.

In summary, this analysis revealed two main findings: 1) Subjects could not accurately estimate the actual length of IMP vision rays, consistently underestimating them, and 2) Subjects tended to believe the evolved IMPs could see further than randomly-parameterized IMPs. I argue that these results can be explained by subjects' rationalization of IMP behavior; when an IMP behaved irrationally (a more common occurrence in the case of non-evolved IMPs), subjects may have attributed this to shortcomings in their visual capability.

3.6.5 Discussion

Evolved IMPs can be considered rational (relative to their non-evolved counterparts) given the context of their environment niche. They also *appear* to be more intelligent to subjects; Experiment 1 established that the evolved status of the IMP was a robust perceptual variable. So, given that the evolved IMPs both behave more rationally

and convey the subjective impression of behaving more rationally, I asked how their enhanced rationality would influence subjects' more specific inferences about their intentions and beliefs. Experiment 2 revealed two robust effects: Subjects were better able to infer the goal states of evolved IMPs, and estimated that evolved IMPs could see further.

3.7 General Discussion and Conclusion

As some minimal level of rationality is critical to the concept of agency, it follows that a theory of mind—a means of understanding the mental states of the observed agent, and their behavioral consequences—should also reflect an expectation of agent rationality. But how should one define the agent rationality presumed by the inferential system? For all of the creatures found in nature—and also some found in artificial systems, such as our IMPs—it is quite difficult, if not impossible, to prescribe some normatively rational strategy on the basis of game theory or decision theory. The strategy space is simply too high-dimensional, and the full set of relevant environmental variables to consider is both ever-changing and exceedingly difficult to discern in advance with any confidence. In many known cases, a normative model of agent decision making can also be plainly inaccurate as a descriptive theory of agent behavior, an empirical fact to which an enormous heuristics and biases literature can testify (Kahneman et al., 1982; Johnson-Laird, 1983; Gigerenzer and Goldstein, 1996). Furthermore, when subjects fail to conform to normative models of rationality, it can often be argued that it is the normative model that has betrayed its own shortcomings, either by lacking robustness or efficiency, or making unsound assumptions about the nature of the problem (Cosmides and Tooby, 1994; Gigerenzer, 2008).

For these reasons, I have operationally defined agent rationality in a manner consistent with Dennett (1987): Behaviors that have been selected by evolutionary pressure can be considered approximately rational, and more adaptive strategies are more rational.

Perhaps, for the IMPs, there does exist one setting in their 22-parameter program which is provably optimal. I cannot derive this optimal IMP, but what I can do is simulate evolution—the “master of high-dimensional trial and error” (Taleb, 2012, p. 349)—to generate a pool of IMPs that is more rational than their non-evolved counterparts. Experiment 1a demonstrated that subjects *agree* that these IMPs are more intelligent, thereby demonstrating some level of consistency between our model of evolved rationality and subjects’ expectations of rational agency. Experiment 2 then demonstrated that, as predicted, subjects are also better able to make sense of the behavior of more rational (evolved) IMPs. Agent rationality—as we have defined it—results in better discrimination of intentions.

These experiments demonstrate that human intuitions about the mental processes generating an agent’s behavior can be rigorously studied, and suggest that these intuitions feature a notion of rationality. However, one important question remains unanswered: How does one access an accurate generative model of agent behavior? Dennett (1987) suggests that because evolved agents are approximately rational most of the time, this allows the cognitive apparatus 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; 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 (as in Experiment 2b) and the nature of the agent itself (Experiment 1c). 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.

3.7.1 Future Directions

In this study, I employ a definition of rationality that is based on evolutionary fitness, but the experimental approach I have used (the IMPs environment) opens the door for

testing a wide range of hypotheses related to competing conceptions of rationality. For instance, imagine if two sets of IMPs were created, one of which exhibited behavior prescribed by a normative theory, and another whose strategies were determined by a more naturalistic evolutionary process. Which set of IMPs would exhibit behavior that better matches human intuitions and expectations about agency? The modular nature of IMPs, and the ability to directly manipulate their cognitive and perceptual capabilities, invites novel experimentation of this kind.

4. Conclusions

The mind itself can be considered a model, a simplified abstraction with value derived not from its veridicality, *per se*, but from its power to explain and predict phenomena. Human beings rely critically on their understanding of this abstraction, because the behavior of an entire class of objects in the world—agents—cannot be effectively explained or predicted according to alternate models like physics (whether naïve or rigorous). If one knew the physical state of every subatomic particle in the agent’s body and its surrounding environment, and all of the relevant physical rules governing their interaction, one could physically model the agent, in principle. But modeling a physical process in this manner, to the precision required to predict any interesting agent behavior, is well beyond the epistemological and computational limits of the brain. Indeed, human evolution has selected against this inefficient approach, because as a matter of fact we do not attempt to perform such an analysis when interacting with objects construed to be agents. We instead rely on our theory of mind.¹

Past experiments in the domain of theory of mind, dating to Heider and Simmel (1944), have frequently relied on stimuli generated according to the intuitions of either the experimenters or their subjects. This may result in too close a correlation between the nature of the independent and dependent variables—stimulus generation and subject response are both direct reflections of human intuitions about the subjective content of scenes, and therefore what is manipulated and what is being measured may be, at least to an extent, the same thing. The IMPs, of course, are programmed according to the intuition of the programmer. But once the play button is hit in the simulation environment, what will be displayed to subjects in the resulting scene is not known in advance. And in Chapter 3, wherein the rationality of the IMPs was the independent variable, it was not my subjective intuition that manipulated this variable. A simulated evolution process produced the more rational IMPs.

¹A bona fide “theory of mind” is sometimes construed to necessarily include a cognitively complex metarepresentative ability, cf. Leslie (1984)’s “second order” theory of mind. I use the term less strictly, in a manner which would also encompass relatively more reflexive or perceptual mechanisms (in response to motion cues), and the inferences made by non-human species with respect to the intentions or dispositions of other animals.

This approach also enables new experimental directions in the study of intentionality, only a few of which have been exploited in this dissertation. 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 and Feldman, 2012) in which subjects' interaction with agents within the virtual environment—in addition to their judgments—may shed light on underlying cognitive mechanisms. But perhaps most importantly, 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.

5. Appendices

5.1 Appendix A: IMP programming, Chapter 2

The experiment was programmed in MATLAB using the Psychtoolbox libraries (Brainard, 1997; Pelli, 1997; Kleiner et al., 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, I linearly interpolated the locations and orientations of objects in the scene between frames, and each scene was sped up 15x. Thus, an originally rendered simulation consisting of ~ 300 frames over 900 seconds was transformed into a displayed scene of ~ 1500 frames over 60 s (a change from approximately 0.33 to 25 frames per second). During the experiments, subjects' keyboard responses were recorded after the presentation of each frame: ~ 25 Hz.

5.1.1 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 quickly shrinks and then disappears.

5.1.2 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 Figure 2.1b). This retina is modeled as a series of rays cast 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 $[r_m \dots r_n]$ 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. 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.

5.1.3 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*, 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 cleared (see Figure 2.1d). This dynamic updating is primarily used for food objects and other agents because obstacles are represented as static elements in a binary map.

5.1.4 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, 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 5.1 and 5.2, respectively.

5.1.5 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).

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.

5.2 Appendix B: IMP programming, Chapter 3

The virtual environment was programmed in Java and presented as an applet to subjects. Frames of scenes could be generated and displayed nearly in real time (that is, much quicker than in the MATLAB version presented in Chapter 2), but would often display unevenly, owing to the variable amount of computation necessary for

Table 5.1: Transition probability matrices for IMPs used in Chapter 2, Experiment 1.

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

Table 5.2: Transition probability matrices for IMPs used in Chapter 2, Experiment 2.

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

generating each frame. Therefore, scenes were rendered offline. Each scene consisted of 3600 frames presented to subjects at a constant 60 frames per second. However, if scenes were generated without being rendered on the computer screen—i.e. during evolution—they could often be simulated at speeds up to 200-500 frames per second (on a 3.10 GHz processor), with the path planning resolution of the IMPs (and to a lesser extent, memory resolution) being the main limiting factor in terms of both time and memory complexity.

A scene can be displayed in any size window; in these experiments, we used a 950×950 pixel window. If the square window is of width W , then an IMP has length $.0375W$, a “small” obstacle is a $.05W \times .05W$ red square, and a “large” obstacle is a $.1W \times .1W$ red square. A “cluster” of 5 food objects initializes as a gray regular octagon with circumradius $.0281W$. As food objects (each of circumradius $.0118W$) are pulled from the cluster by IMPs, this octagon shrinks linearly with the number of food objects removed from it, until it is of the same radius as a single food object and may itself be completely pulled away by an IMP.

The IMPs’ vision, memory, and path planning modules work in the same manner as those of the previously described IMPs (see Appendix A), except that the settings of these modules’ parameters are not, in this case, set by the experimenter, but rather by the evolutionary algorithm. One difference in the programming of the Java IMPs, versus the MATLAB IMPs, is that when the Java IMPs “grab” food, they pull or “tow” this food toward their target cache location rather than pushing it. But the primary difference in the way the Java IMPs are programmed lies in their decision making module, described here:

5.2.1 Decision making

In order to conditionalize their behavior somewhat intelligently, the MATLAB IMPs utilized in Chapter 2 relied on having four different goal state transition probability tables, each applied to one of four qualitatively different circumstances in which an IMP might find itself. By contrast, the Java IMPs only employ one transition table, but this table’s respective columns are multiplied by different constants depending

on the IMP's circumstances and the content of its memory. That is, the probability of transitioning to a particular goal state will vary over the course an IMP's stint in the environment.

The precise algorithmic rules governing this procedure are denoted below, but the general empirical consequences are that the IMP will default to the *explore* state if there are no other IMPs or food objects in its memory, that the IMP will be biased against transitioning to the *explore* state if there are other IMPs or food objects in its memory, and that IMPs will only be likely to *flee* other IMPs if they are nearby. This is why even randomly-parameterized IMPs do not, on average, spend equal time in each goal state (see Table 3.1).

Each IMP has *attack*, *explore*, *flee*, and *gather* "brain" modules in its program. Every time the decision making module is queried (which happens once every n frames, depending on the setting of another evolvable parameter), each brain returns a potential target location in the environment (according to the same rules described in Appendix A), and an "urgency" value between 0 and 1.0.

The previous goal state (G_t) of the IMP is known, and indexes one row of the IMP's goal state transition table. Each of the four entries in this row vector corresponds to the probability of transitioning to one of the four goal states at time $t + 1$, and each of these four probabilities is multiplied by the urgency returned by its corresponding brain module. After this multiplication, the row vector is renormalized to sum to 1.0, and the IMP transitions to one of the four goal states randomly according to this new vector of probabilities.

- **Attack brain:** If the IMP memory contains no traces of other IMPs, returns an urgency of 0. Otherwise, returns an urgency of 1.0.
- **Explore brain:** If G_t is *explore*, and the previous target for exploration has not yet been reached by the IMP, returns an urgency equivalent to the maximum urgency outputted by all brains at time t . Otherwise, returns an urgency of 0.6.
- **Flee brain:** If the IMP memory contains no traces of other IMPs, returns an urgency of 0. If the distance to the nearest other IMP in memory, d_n , is less than

or equal to d_c (where d_c is two IMP lengths), returns an urgency of 1.0. If $d_n > d_c$, returns an urgency of d_c/d_n .

- **Gather brain:** If the IMP is “grabbing” food, returns an urgency of 0.8. If the IMP memory contains no traces of food, returns an urgency of 0. Otherwise, returns an urgency of 1.0.

5.2.2 Evolution

After each generation of simulated evolution, our algorithm generated a population of 5 new IMPs. Each surviving IMP from the previous generation could spawn each of the new generation’s IMPs with probability h_i/h_t , where h_i was the “health” value of the surviving IMP, and h_t was the total health of the previous generation of IMPs. Every parameter of each new IMP was then randomly “mutated”, with the deviation from the present parameter value being randomly drawn from the following normal distributions, denoted here as $N(\mu, \sigma^2)$:

- **Vision resolution:** $N(0,9)$, rounded to the nearest integer.
- **Vision field of vision:** $N(0,.04)$, in radians.
- **Vision ray length:** $N(0,9)$
- **Memory resolution:** $N(0,25)$, rounded to the nearest integer.
- **Path planning resolution:** $N(0,9)$, rounded to the nearest integer.
- **Frames per new decision:** $N(0,25)$, rounded to the nearest integer.
- **State transition probabilities:** $N(0,.0025)$ for each of the 16 probabilities in the 4×4 table. The rows are then renormalized to sum to 1.

Because IMPs that “died” during a generation were ineligible to reproduce in the next generation, the evolution could not continue under the (rare) circumstance where none of the 5 IMPs in the population survived. When this happened (typically once over the course of 1000 generations, but sometimes never), the evolution would

simply replay the previous generation of IMPs, but with food, obstacles, and IMPs placed in newly-randomized initial locations.

5.3 Appendix C: Experimental instructions, Chapter 3

5.3.1 Experiment 1a

In this study, you will observe a series of 60-second scenes in which triangular “agents” interact with one another and their environment.

In this environment, red squares are obstacles. The agents cannot move through them.

It is beneficial to the agents to collect “food.” This food is round and gray and can be found at various locations in the environment. For an agent to get the full benefit of this food, it must bring it to a particular location, by “towing” it back there. This location is different for each agent, and is the same location as its starting point in the scene. Food disappears when the agent brings it to this point.

If the agent is attacked and hit head on by another agent, this is damaging to its health. But an agent can also inflict this damage on another agent if it so chooses, which may ultimately result in that other agent “dying” and becoming food.

If two agents hit each other head on, this hurts both agents, but to a lesser degree.

Task instructions: There will be five agents in each scene, but you only need pay attention to the red agent and the blue agent. After the scene is done, you should indicate which of those two agents you think behaved more intelligently over the course of the scene, on a scale of 1 (definitely the red agent) to 6 (definitely the blue agent). So that you can rest your hands comfortably at the bottom row of the keyboard, “1”, “2”, and “3”, will correspond to “Z”, “X”, and “C”, respectively. “4”, “5”, and “6” will correspond to “,” “.” and “/”. The experiment will move on to the next scene when you press a key.

5.3.2 Experiment 1b

In this study, you will observe a series of 60-second scenes in which triangular “agents” interact with one another and their environment.

In this environment, red squares are obstacles. The agents cannot move through them.

It is beneficial to the agents to collect “food,” but only slightly so—it is usually more trouble than it’s worth. This food is round and gray and can be found at various locations in the environment. For an agent to benefit from this food, it must bring it to a particular location, by “towing” it back there. This location is different for each agent, and is the same location as its starting point in the scene. Food disappears when the agent brings it to this point.

If the agent is attacked and hit head on by another agent, this is extremely damaging to its health. But an agent can also inflict this damage on another agent if it so chooses, and hitting another agent head on is very beneficial to its own health. Also, inflicting damage on another agent may ultimately result in that other agent “dying” and becoming food.

If two agents hit each other head on, this does not help or hurt either agent.

Task instructions: There will be five agents in each scene, but you only need pay attention to the red agent and the blue agent. After the scene is done, you should indicate which of those two agents you think behaved more intelligently over the course of the scene, on a scale of 1 (definitely the red agent) to 6 (definitely the blue agent). So that you can rest your hands comfortably at the bottom row of the keyboard, “1”, “2”, and “3”, will correspond to “Z”, “X”, and “C”, respectively. “4”, “5”, and “6” will correspond to “,” “.” and “/”. The experiment will move on to the next scene when you press a key.

5.3.3 Experiment 1c

In this study, you will observe a series of 60-second scenes in which triangular “agents” interact with one another and their environment.

In this environment, red squares are obstacles. The agents cannot move through them.

It is beneficial to the agents to collect “food.” This food is round and gray and can be found at various locations in the environment. For an agent to get the full benefit of this food, it must bring it to a particular location, by “towing” it back there. This location is different for each agent, and is the same location as its starting point in the scene. Food disappears when the agent brings it to this point.

If the agent is attacked and hit head on by another agent, this is damaging to its health. But an agent can also inflict this damage on another agent if it so chooses, which may ultimately result in that other agent “dying” and becoming food.

If two agents hit each other head on, this hurts both agents, but to a lesser degree.

Task instructions: Some of the agents you will see in these video scenes were controlled by human players, and some are robots. A human player does not see the environment from your perspective (looking in on the scene from above), but rather from the perspective of his/her triangular character, which can make navigation difficult. He/she cannot necessarily see things well at a distance, for example, and often must be quite close to something to recognize it. Controlling the movements of one’s character can also be challenging.

There will be five agents in each scene, but you only need pay attention to the red agent and the blue agent. After the scene is done, you should indicate which of those two agents you think is more likely to have been controlled by a human player, on a scale of 1 (definitely the red agent) to 6 (definitely the blue agent). So that you can rest your hands comfortably at the bottom row of the keyboard, “1”, “2”, and “3”, will correspond to “Z”, “X”, and “C”, respectively. “4”, “5”, and “6” will correspond to “,”, “.” and “/”. The experiment will move on to the next scene when you press a key.

Bibliography

- Anderson, J. R. (1989). A rational analysis of human memory. In Roedinger, H. L. and Craik, F., editors, *Varieties of memory and consciousness*, pages 195–210. Lawrence Earlbaum Associates, Hillsdale, NJ.
- Baker, C. L., Saxe, R., and Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition*, 113:329–349.
- Baker, C. L., Tenenbaum, J. B., and Saxe, R. R. (2006). Bayesian models of human action understanding. *Advances in Neural Information Processing* 18.
- Baker, C. L., Tenenbaum, J. B., and Saxe, R. R. (2007). Goal inference as inverse planning. *Proceedings of the Twenty-Ninth Annual Conference of the Cognitive Science Society*.
- Barrett, H. C., Todd, P. M., Miller, G. F., and Blythe, P. W. (2005). Accurate judgments of intention from motion cues alone: A cross-cultural study. *Evolution and Human Behavior*, 26:313–331.
- Bicchieri, C. (2009). *The Handbook of Philosophy of Economics*, chapter Rationality and Indeterminacy, pages 159–188. The Oxford References Library of Philosophy, Oxford University Press.
- Blythe, P., Todd, P. M., and Miller, G. F. (1999). How motion reveals intention: categorizing social interactions. In Gigerenzer, G., Todd, P. M., Miller, G. F., and the ABC Research Group, editors, *Simple Heuristics That Make Us Smart*, pages 257–286. Oxford University Press, New York, NY.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10:443–446.
- Braitenberg, V. (1984). *Vehicles*. MIT Press, Cambridge, MA.
- Burnham, K. P. and Anderson, D. R. (2002). *Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach*. Springer, New York.
- Carnahan, J., Li, S., Costantini, C., Tourè, Y. T., and Taylor, C. E. (1997). Computer simulation of dispersal by *Anopheles gambiae* s.l. in West Africa. *Artificial Life V*, 5:387–394.
- Chaitin, G. J. (1970). To a mathematical definition of “life”. *ACM SICTACT News*, 4:12–18.
- Cherniak, C. (1981). Minimal rationality. *Mind*, 90:161–183.
- Cliff, D. and Miller, G. F. (1996). Co-evolution of pursuit and evasion II: Simulation methods and results. In Maes, P., Mataric, M., Meyer, J., Pollack, J., and Wilson, S., editors, *From Animals to Animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*. MIT Press Bradford Books.
- Cosmides, L. and Tooby, J. (1994). Better than rational: Evolutionary psychology and the invisible hand. *The American Economic Review*, 84(2):327–332.
- Crick, C. and 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*. MIT Press, Cambridge, MA.
- Dittrich, W. H. and Lea, S. E. G. (1994). Visual perception of intentional motion. *Perception*, 23:253–268.
- Feldman, J. and Tremoulet, P. D. (2008). The attribution of mental architecture from motion: Towards a computational theory. Technical Report TR-87, Rutgers University Center for Cognitive Science, Piscataway, NJ.
- Gao, T., McCarthy, G., and 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., and Scholl, B. J. (2009). The psychophysics of chasing: A case study in the perception of animacy. *Cognitive Psychology*, 59:154–179.
- Gao, T. and 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 Chalupa, L. M. and Werner, J. S., editors, *The Visual Neurosciences*. MIT Press, Cambridge, MA.
- Gelman, R., Durgin, F., and Kaufman, L. (1995). Distinguishing between animates and inanimates: not by motion alone. In Sperber, D., Premack, D., and Premack, A. J., editors, *Causal Cognition: A Multidisciplinary Debate*, chapter 6, pages 151–184. Oxford University Press, New York, NY.
- Gergely, G., Nádasdy, Z., Csibra, G., and Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, 56:165–193.
- Gigerenzer, G. (2008). Why heuristics work. *Perspectives on Psychological Science*, 3(1):20–29.
- Gigerenzer, G. and Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103(4):650–669.
- Goldman, A. (2006). *Simulating Minds: The Philosophy, Psychology and Neuroscience of Mindreading*. Oxford University Press, New York.
- Heal, J. (1996). Simulation, theory, and content. In Carruthers, P. and Smith, P. K., editors, *Theories of Theories of Mind*. Cambridge University Press, New York.
- Heider, F. and Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology*, 57:242–259.
- Johnson, S. C. (2000). The recognition of mentalistic agents in infancy. *Trends in Cognitive Development*, 16:637–656.
- Johnson-Laird, P. N. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Harvard University Press, Cambridge, MA.

- Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, 58(9):697–720.
- Kahneman, D., Slovic, P., and Tversky, A., editors (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press, Cambridge, UK.
- Kleiner, M., Brainard, D., and Pelli, D. (2007). What’s new in Psychtoolbox-3? In *Perception 36 ECVF Abstract Supplement*.
- Knill, D. C. and Richards, W. (1996). *Perception as Bayesian inference*. Cambridge University Press, Cambridge, UK.
- Kuhlmeier, V., Wynn, K., and Bloom, P. (2003). Attribution of dispositional states by 12-month-olds. *Psychological Science*, 14(5):402–408.
- Leslie, A. M. (1984). Infant perception of a manual pick-up event. *British Journal of Developmental Psychology*, 2:19–32.
- Leslie, A. M. (1987). Pretense and representation: The origins of “theory of mind”. *Psychological Review*, 94(4):412–426.
- Leslie, A. M. (1994). ToMM, ToBy, and agency: core architecture and domain specificity. In Hirschfeld, L. A. and Gelman, S. A., editors, *Mapping the Mind: Domain Specificity in Cognition and Culture*. Cambridge University Press.
- Malle, B. F., Moses, L. J., and Baldwin, D. A. (2001). Introduction: The significance of intentionality. In Malle, B. F., Moses, L. J., and Baldwin, D. A., editors, *Intentions and intentionality: Foundations of social cognition*, pages 1–24. MIT Press, Cambridge, MA.
- Mann, R., Jepson, A., and Siskind, J. M. (1997). The computational perception of scene dynamics. *Computer Vision and Image Understanding*, 65(2):113–128.
- McAleer, P. and Pollick, F. E. (2008). Understanding intention from minimal displays of human activity. *Behavior Research Methods*, 40(3):830–839.
- Mitchell, M. and Forrest, S. (1993). Genetic algorithms and artificial life. *Artificial Life*, 1:267–389.
- Nichols, S. and Stich, S. (1998). Rethinking co-cognition: A reply to Heal. *Mind & Language*, 13(4):499–512.
- Nolfi, S. and Floreano, D. (1998). Coevolving predator and prey robots: Do “arms races” arise in artificial evolution? *Artificial Life* 4, 4:311–335.
- Nowak, M. A. and Sigmund, K. (2004). Evolutionary dynamics of biological games. *Science*, 303:793–799.
- Pantelis, P. C., Baker, C. L., Cholewiak, S. A., Weinstein, A., Wu, C., Tenenbaum, J. B., and Feldman, J. (2013). Inferring the intentional states of autonomous virtual agents. *Manuscript submitted for publication*.

- Pantelis, P. C., Cholewiak, S. A., Ringstad, P., Sanik, K., Weinstein, A., Wu, C., and Feldman, J. (2011). Perception of intentions and mental states in autonomous virtual agents. In Carlson, L., Hölscher, C., and Shipley, T., editors, *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, pages 1990–1995, Austin, TX. Cognitive Science Society, Cognitive Science Society.
- Pantelis, P. C. and Feldman, J. (2012). Exploring the mental space of autonomous intentional agents. *Attention, Perception, & Psychophysics*, 74(1):239–249.
- Pautler, D., Koenig, B., Quek, B., and 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.
- Reynolds, C. W. (1994). Competition, coevolution and the game of tag. In *Artificial Life IV: Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems*. MIT Press.
- Searle, J. R. (1984). *Minds, brains and science*. Harvard University Press, Cambridge, MA.
- Shao, W. and Terzopoulos, D. (2007). Autonomous pedestrians. *Graphical Models*, 69:246–274.
- Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1):99–118.
- Smith, J. and Price, G. (1973). The logic of animal conflicts. *Nature*, 246:15–18.
- Stewart, J. A. (1982). *Perception of animacy*. PhD thesis, University of Pennsylvania.
- Stich, S. and Nichols, S. (2003). Folk psychology. In *The Blackwell Guide to Philosophy of Mind*, pages 235–255. Oxford: Basil Blackwell.
- Taleb, N. N. (2012). *Antifragile: Things that gain from disorder*. Random House, New York.
- Tauber, S. and Steyvers, M. (2011). Using inverse planning and theory of mind for social goal inference. In Carlson, L., Hölscher, and Shipley, T., editors, *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, Austin, TX. Cognitive.
- Tremoulet, P. D. and Feldman, J. (2000). Perception of animacy from the motion of a single object. *Perception*, 29:943–951.
- Tremoulet, P. D. and 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., and Tenenbaum, J. B. (2009). Help or hinder: Bayesian models of social goal inference. In *Advances in Neural Information Processing Systems*, volume 22, pages 1874–1882.
- Williams, E. W. (2000). *Causal reasoning by children and adults about the trajectory, context, and animacy of a moving object*. PhD thesis, 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 Langton, C., editor, *Proceedings of the Artificial Life III Conference*, pages 263–298, Reading, MA. Addison-Wesley.
- Zacks, J. M., Kumar, S., Abrams, R. A., and Mehta, R. (2009). Using movement and intentions to understand human activity. *Cognition*, 112:201–216.
- Zacks, J. M. and Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, 127:3–21.

Vita

Peter C. Pantelis

2010 M.S. in Psychology, Rutgers University

2007 B.A. in Psychology, Political Science, University of Pennsylvania

2011-2013 Teaching assistant, Department of Psychology, Rutgers University

Journal Publications

Pantelis, P. C., Baker, C. L., Cholewiak, S. A., Sanik, K., Weinstein, A., Wu, C., Tenenbaum, J. B., & Feldman, J. Inferring the intentional states of autonomous virtual agents. *Submitted*.

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

Pantelis, P.C., van Vugt, M. K., Sekuler, R., Wilson, H. R., & Kahana, M. J. (2008). Why are some people's names easier to learn than others? The effects of similarity on memory for face-name associations. *Memory & Cognition*, 36(6), 1182–1195.