Outcome or Strategy? A Bayesian Model of Intelligence Attribution

Marta Kryven (mkryven@uwaterloo.ca)
Department of Computer Science, University of Waterloo

Tomer Ullman (tomeru@mit.edu)
Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology

William Cowan (wmcowan@cgl.uwaterloo.ca)
Department of Computer Science
University of Waterloo

Joshua B. Tenenbaum (jbt@mit.edu)
Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology

Abstract

People have a common-sense notion of intelligence and use it to evaluate decisions and decision-makers. One can attribute intelligence by evaluating the strategy or the outcome of a goal-directed agent. We propose a model of intelligence attribution, based on inverse planning in Partially Observable Markov Decision Processes (POMDPs) in a probabilistic environment, inferring the most likely planning parameters given observed actions. The model explains the agent's decisions by a combination of probabilistic planning, a *softmax* decision noise, prior knowledge about the world and forgetting, estimating the agent's intelligence by a proxy measure of efficiently optimising costs and rewards. Behavioural evidence from two experiments shows that people cluster into those who attribute intelligence to the strategy and those who attribute intelligence to the outcome of the observed actions. People in the strategy cluster attribute more intelligence to decisions that minimise the agent's overall cost, even if the outcome is unlucky. People in the outcome cluster attribute intelligence to the outcome, judging low-cost outcomes as a sign of intelligence even if the outcome is accidental and make neutral judgements before they observe the result. Our model explains human intelligence judgements better than perceptual cues such as the number of revisits or moves.

Keywords: Theory of mind; Intelligence attribution; Social cognition; Bayesian inference; Partially Observable Markov Decision Processes; Inverse planning

Introduction

In everyday life people make fast, intuitive judgements of intelligence. For example, it is more intelligent to rush to catch an infrequent bus than to run towards a subway that departs every 3 minutes, and solving a puzzle from scratch is more intelligent than looking up the answer in a key. Existing qualitative accounts of intelligence explain in general terms why people might describe actions as intelligent or unintelligent. According to Dennett's *rationality principle*, people expect intentional agents to act efficiently in order to achieve their desires, given their beliefs about the world (Dennett, 1989). People may describe behaviour as intelligent if it agrees with their expectations of what the agent should do, and as stupid if the behavior can be explained by a failure of attention, overconfidence or loss of control (Aczel, Palfi, & Kekecs, 2015).

One way to attribute intelligence is to judge the agent's efficiency in achieving a goal, echoing the rationality principle. Efficiency means achieving a rewarding goal at a minimal cost. So, an observer attributing intelligence needs to understand the costs and rewards of a situation and how to plan under uncertainty to maximise the likelihood of success. Evaluating the specific action of an agent seeking a goal requires the observer to ask 'Is this the best goal for that agent?' and 'Is this the best way to achieve that goal?'. The particular end-goal might even be negative, but the behavior could still be conceded as intelligent and thus preferable. Children as young as two understand and value competence, preferring agents who can perform an action on the first attempt, even if the agent is mean (Jara-Ettinger, Tenenbaum, & Schulz, 2015).

However, fully evaluating an agent's planning procedure can be computationally hard, and observing an agent's outcome can provide a shortcut to evaluating its intelligence. In effect, the observer might think 'If they achieved their goal, they must be smart', even though the goal was achieved by sheer luck or prior knowledge. Adults under conditions of cognitive load (Olson et al., 2013) and 7-year old children perceive lucky persons as more likable (Olson, Banaji, Dweck, & Spelke, 2006; Olson, Dunham, Dweck, Spelke, & Banaji, 2008), which suggests that lucky others may also be seen as more intelligent.

In this paper we develop a formal model of intelligence attribution based on Bayesian inverse planning (Baker, Saxe, & Tenenbaum, 2011). We describe two behavioral experiments that validate the model, considering the role of rational expectations and mental short-cuts. The next section gives an informal sketch of the model, followed by formal modelling and behavioral experiments to evaluate the model.

Computational Framework

To describe intelligent behaviour formally, consider a simple context in which it can be observed: a 2-D animation of geometric shapes. Most adults describe the actions of agents in such displays by constructing a narrative about the actors' mental states (Heider & Simmel, 1944). Such 2-D animations were previously used to inform computational models of goal attribution (Baker et al., 2011), preference attribution (Jara-Ettinger, Baker, & Tenenbaum, 2012) and social behaviour (Ullman et al., 2009). Building on these studies we propose a model of intelligence attribution in the context of a rational agent looking for an object in a 2-D world.

Consider an agent (for concreteness, a mouse) looking for a goal (a treat) in a maze. The mouse is familiar with the layout of the maze: it knows which rooms are big and which are small, it knows which rooms are close and easily accessible and which are far away. The mouse knows there is a treat in the maze, but does not know where it is. The treat is equally likely to be anywhere. What should an intelligent mouse do to find the treat? It should plan efficiently. And when seeing the mouse running through the maze, how does one know if it is intelligent? By comparing its behavior to the behaviour expected of an agent that plans efficiently.

We formalise a rational agent's decision-making process as probabilistic planning in a Partially Observable Markov Decision Processes (POMDP). POMDP assumes that the agent acts sequentially to maximise the reward and minimise the cost of each action, given its beliefs about the world (Figure 1). After each action the agent updates its beliefs based on observations caused by the previously chosen action.

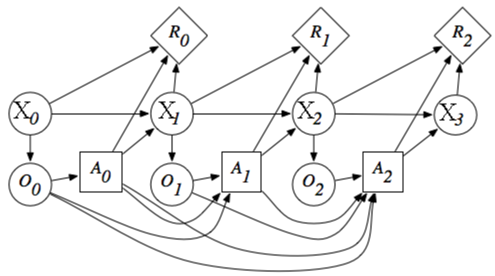

Figure 1: (a) POMDP is a sequential process described by beliefs X_t, observations O_t, rewards R_t and actions A_t at time step t. Arrows indicate a causal relationship.

A viewer observes decisions made by an agent and judges the agent's intelligence by inferring the parameters guiding the agent's planning. The observer estimates the agent's intelligence by a measure of agent efficiency, so that the more optimal agents are judged as more intelligent. We define the agent optimality rank (Table 1) based on how well the corresponding planing strategies optimise the long-term costs and rewards over randomly-generated mazes with a randomly placed goal. Thus, this optimality rank is defined for the model problem and not as a universal cost-minimising strategy. Prior knowledge indicates a non-uniformly distributed prior belief about the goal's location. Since the agent supposedly does not know where the goal is, we use such prior knowledge (with correct information) to represent luck. Decision noise is captured by a $softmax$ parameter, set such that actions are chosen with a probability proportional to their reward. To model the observer, we use Bayesian inverse inference on the POMDP by integrating the likelihood of the observed actions with the prior over a set of possible POMDP parameters (Baker et al., 2011).

In principle, this model can describe an evaluation of any behaviour encoded in discrete time and space and can admit a variety of cost functions, reward functions and discount rate models. For concreteness, we limit the planning part of the model to three possible reasons for deviating from an optimal search strategy: prior knowledge about the world, decision noise, and forgetting. While this model is not meant to fully capture the complexities of human intelligence attribution, it provides a computational ideal-observer benchmark to test against experimental data and perceptual-based metrics.

Table 1: Agent optimality rank

Rank	Description
6	Optimal
5	Suboptimal with decision noise or Suboptimal or with prior knowledge
4	Suboptimal with decision noise and prior knowledge
3	Suboptimal with decision noise and forgetting
2	Suboptimal with prior knowledge, forgetting, and decision noise
1	Suboptimal and seemingly random

The Gridworld and POMDP in Detail

Formally, a model world (a maze) is described by discrete time, $0 \leq t \leq T$, and a grid of cells, $W = \{w(i, j)\}$, $w(i, j) \in \{wall, empty, goal\}$. A single cell contains a goal: $\exists(i_g, j_g) \rightarrow w(i_g, j_g) = goal$. The agent acts based on its belief about the world, which is a set of probabilities $X_t = \{P(\mathbf{x_s})_t\}$. Here $X = \{\mathbf{x_s}\}$ is the set of all possible world states such that $\mathbf{x_s}$ represents a world with the goal in cell s. X_0 encodes the set of the agent's initial beliefs.

The agent knows its own location at time t, L_t and sees 180^0 of visual field. The results are observation probabilities $O_t = P(W|X_t, L_t)$. A cell s is visible if the four rays cast from each of the corners of L_t to the corresponding corners of cell s do not intersect walls. If s is visible then $O_t(s) = 1$, otherwise $O_t(s) = 0$. Unseen cells are portrayed as dark, and cells previously seen are patterned (Figure 2). The agent moves one grid cell at a time, choosing among available deterministic actions $A(L_t) = \{a_i\} \in \{N, S, W, E\}$. An action is available if it does not lead into a wall. The agent updates its beliefs contingent on observations using standard Bayesian updating: $X_{t+1} = P(W|A_t, X_t) \propto P(O_t|A_t, X_t)P(A_t|X_t)P(W)$

Calculating Rewards

At time t, the agent calculates the value of each action $Q(a, L_t, X_t)$ and the reward function $R : X_t \times A_t \mapsto \mathbb{R}$ and chooses an action A_t with a likelihood proportional to its reward. Strictly optimal planning in POMDPs is computationally intractable, and in our example the solution is approximated by one step lookahead, or by direct design, on an assumption that the space of beliefs remains unchanged after taking one action (Hauskrecht, 2000):

$$\begin{aligned} Q(a, L_t, X_t) = &\sum_{\mathbf{x_s}} P(\mathbf{x_s})_t \rho(\mathbf{x_s}, L_t + a) \\ &+ \gamma \max_{a_i \in A(L_t + a)} \{Q(a_i, L_t + a, X_t)\}, \end{aligned} \tag{1}$$

where γ is a discount factor, and ρ is proportional to the square of the distance traveled to reach each cell: $\rho(\mathbf{x_s}, L_t + a) \propto \frac{1}{||(L_t+a)-(i,j)||^2}$.

Thus, $\rho(\mathbf{x_s}, L_t + a)$ represents the value of being in a location $L_t + a$ in the world $\mathbf{x_s}$ and $\sum_{\mathbf{x_s}} P(\mathbf{x_s})_t \rho(\mathbf{x_s}, L_t + a)$ is the value of an action a given current beliefs. The second term in equation (1) describes the discounted value of subsequent actions, assuming actions are chosen optimally.

Finally, the reward function is defined as:

$$R(Q_t(a_i)) = \frac{exp(Q_t(a_i)/\tau)}{\sum_j exp(Q_t(a_j)/\tau)}, \quad (2)$$

where τ is a $softmax$ decision noise. As $\tau \to 0$ the agent deterministically chooses the action with the highest value, and as $\tau \to \infty$ the agent acts at random. For our experiments we either set $\tau = 0$ (optimal) or used a low level of $\tau = 0.01$ such that actions were chosen with a probability proportional to $Q(a_i, L_t, X_t)$ but with most actions optimal (decision noise conditions). A forgetting parameter $0 \leq f \leq 1$ regresses the agent's beliefs toward the mean after each step so that the agent gradually forgets whether a previously observed cell is empty and must re-check already visited locations. Furthermore, prior knowledge about the world is used to to simulate luck: an agent with correct prior knowledge finds the goal sooner.

Experiments

Experiment 1

The first experiment tests whether people attribute intelligence differently to optimal and to suboptimal actions.

Participants. 12 participants recruited from the University of Waterloo, 4 females and 8 males, median age 27.5. Both experiments received ethics clearance from a University of Waterloo Research Ethics Committee and from an MIT Ethics Review Board.

Stimuli. 30 animated movies of a mouse looking for food in a maze were shown in two blocks on a computer screen using Psychtoolbox (Brainard, 1997). The movies were computer-generated by solving a POMDP on one of 9 mazes with two levels of forgetting (on, off), prior knowledge (a correct prior, or a uniformly distributed prior belief) and decision noise (on, off). We varied the appearance of the mouse, the layout of the maze, the textures of the maze and the POMDP parameters on each trial. In every movie the mouse found the treat. The location of the food was counterbalanced so that in one half of the mazes the mouse could find it equally quickly by optimal planning or by luck. In another half of the mazes, the goal was placed so that an optimal agent had to search exhaustively, while a lucky agent with prior knowledge could get finish in fewer steps.

Each movie was labeled according to the most likely POMDP settings estimated by the inverse-planning inference. A movie was labelled *optimal-lucky* if the optimal planner and the prior knowledge planner were equally likely and *optimal-unlucky* if the optimal planner was most likely. Labels *prior knowledge*, *decision noise*, *noise-forgetting*, *noise-forgetting-prior* and *noise-prior* accordingly represented combinations of parameters. A control condition *suboptimal-control* showed an inefficient, highly forgetful and random agent. Each of the conditions occurred four times and the control condition occurred twice. The full set of stimuli used in this and the following experiment can be downloaded from `http://cgl.uwaterloo.ca/~mkryven/`

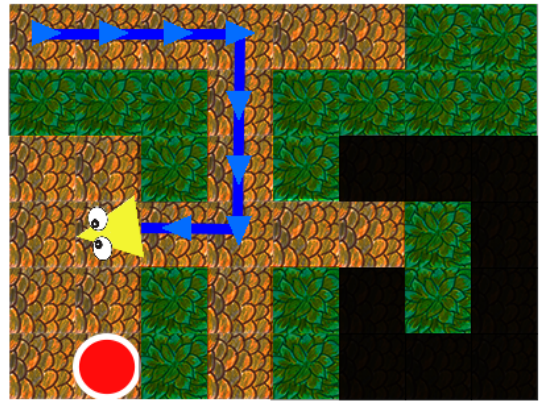

Figure 2: An example of an *optimal-lucky* condition. The agent makes an (optimal) decision to go into the room on the left (reader's perspective). Incidentally, it finds the goal quickly. Dark squares indicate that the agent has not yet seen the area.

Method. Each participant read the instructions on the computer screen and viewed four familiarisation examples followed by the 30 stimuli in two blocks. After viewing each movie the participant recorded his or her rating into a provided Likert answer sheet on a scale from 1 (less intelligent) to 5 (more intelligent).

Results. There was no main effect of block. A two factor ANOVA of rating for *participant* $\times$ *condition* shows a main effect of participant (mean rating 3.44, standard deviation 0.46, $p < .0001$) and of condition ($p < .0001$, Table 2). There was no significant difference between the *optimal-lucky* and *optimal-unlucky* conditions ($p = 0.99$) or between the *optimal-lucky, optimal-unlucky* and *decision noise* conditions ($p = .08$) indicating that participants did not penalise occasional inefficient moves. However, the difference between the *optimal-lucky, optimal-unlucky* and the *prior knowledge* conditions ($p = .001$, difference = 0.8) indicates that participants attributed intelligence to the agent's strategy more than to the agent's outcome.

According to a Scheffe Post-Hoc test, suboptimal conditions were rated differently depending on the cause inferred by the model. Thus, the *decision noise* condition was rated higher than *noise-forgetting* ($p < .0001$, difference = 1) or *noise-forgetting-prior* ($p < .0001$, difference = 2), and *prior knowledge* condition was rated higher than *noise-forgetting* ($p = .0002$, difference = 0.9) or *noise-forgetting-prior* ($p <$ .0001, difference = 1.8). The *suboptimal-control* was rated lower than all other conditions ($p < .0001$). Pearson correlations between the agent optimality rank and human rating is .73 and regression of optimality and rating ($r^2 = 53.7\%$), showing a good fit of our model to the data.

Table 2: Mean ratings and Std. err. of the mean by condition

Condition	Rating	SEM
optimal-lucky	4.7	0.11
optimal-unlucky	4.6	0.11
decision noise	4.1	0.11
prior knowledge	3.9	0.11
noise-prior	3.5	0.11
noise-forgetting	3	0.11
noise-forgetting-prior	2.1	0.11
suboptimal-control	1.6	0.15

The agent's strategy can thus help explain human attributions of intelligence. But can perceptual cues, the path length or the number of cells revisits, explain it equally well? Assuming that a move is scored either for moving from one cell to another or turning, ANOVA for *participant* $\times$ *moves* $\times$ *revisits*, shows significant effects of participant ($p < .0001$), moves ($p = .001$) and revisits ($p < .0001$). Table 3 shows mean ratings for different averaged levels of moves and of revisits, where move levels were obtained by splitting the trials into four bins of equal size and calculating the average number of moves per bin. Pearson correlations between the number of moves and ratings, $-.42$, number of revisits and ratings, $-.49$, and multiple regression of moves, revisits and rating ($r^2 = 31.5\%$), calculated over individual trials, show that our model provides a closer fit to the data than the perceptual metrics alone.

In summary, participants did not attribute more intelligence to lucky agents, and judged agents with an efficient decision-making strategy to be more intelligent. The participants attributed the highest intelligence to the approximately rational agents, which either chose optimally or deviated from the optimal decision within a margin explainable by a *softmax* decision noise. Moreover, the inferred cause of sub-optimality matters: random agents are judged least intelligent of all, and forgetful agents as less intelligent than agents with decision noise.

Table 3: Mean ratings and Std. err. of the mean for number of moves and revisits

Moves	Rating	SEM
14	4	0.76
19	3.2	0.14
23	3.1	0.38
31	3.7	0.37

Revisits	Rating	SEM
0	3.9	0.55
2	4.2	0.27
4	3.6	0.18
9	2.2	0.25

Experiment 2

In the second experiment we investigated how do people form a judgement of intelligence. May people decide after observing just one decision, or do they accumulate evidence over time? We used the same mouse in a maze scenario as in Experiment 1, except that on half of the trials the movie stopped after the mouse chose one of the rooms but before the treat was found. In the other half of the trials the movie played until the mouse found the treat.

Participants. 32 participants were recruited via Amazon Mechanical Turk, 2 were discarded for failing to answer questions. The analysis thus included 30 participants (11 females, median age 34).

Stimuli. 40 animated movies were generated by solving POMDP on 8 different mazes. We varied the appearance of the mouse, the layout, textures and orientation of the maze, the location of the goal, and the POMDP settings on each trial. There were 19 complete trials, with each condition (*optimal-unlucky, prior knowledge, decision noise* and *decision noise with prior knowledge*) occurring 4 times and *optimal-lucky* occurring three times. Another 19 trials were incomplete, 12 showing a suboptimal decision (*suboptimal-incomplete* condition) and 7 showing an optimal decision (*optimal-incomplete* condition). Two control trials showed highly forgetful, inefficient mice.

Method. Participants read the instructions on a computer screen in a web browser and viewed 4 familiarisation examples followed by the 40 animated movies shown in two blocks. After viewing each movie, the participant selected a rating from a Likert scale between 1 (less intelligent) to 5 (more intelligent). At the end of the Web survey participants were asked: *How did you make your decision?*

Table 4: Mean ratings and Std. err. of the mean by condition

Condition	Rating	SEM
optimal-lucky	4.5	0.08
prior knowledge	4.3	0.07
optimal-unlucky	4.1	0.07
noise-prior	3.9	0.07
decision noise	3.8	0.07
optimal-incomplete	3.5	0.06
suboptimal-incomplete	3.3	0.05
control	1.2	0.1

Results. There was no main effect of block. A two-factor ANOVA of rating for *participant* $\times$ *condition* shows a main effect of participant ($p < .0001$, mean 3.63, standard deviation 0.45) and a main effect of condition ($p < .0001$, Table 4). Main effects of age ($p = .02$) and gender ($p = .0006$) indicate that older participants (difference between groups split by the median age 0.24; median ages 26 and 43) and females (difference 0.3) were more generous. ANOVA of rating over incomplete conditions for *participant* $\times$ *condition* shows a small difference between intelligence attributed to optimal and suboptimal decisions (difference = 0.17, $p = .009$).

In agreement with Experiment 1, the *control* condition was judged as least intelligent ($p < .0001$). However, there was no difference between the *optimal-lucky* ($p = .73$), *optimal-unlucky* ($p = .19$) and the *prior knowledge* conditions. Moreover, *optimal-lucky* and *optimal-unlucky* agents were rated as different ($p = .0.004$, difference 0.47) indicating that online participants judged lucky and efficient agents equally, and optimal agents as more intelligent when they were lucky. Pearson correlation between the agent optimality rank and ratings is .55, regression of optimality and rating $r^2 = 34.6\%$.

The relationship between ratings and perceptual cues was analysed only on full trials, as incomplete trials are always shorter. ANOVA of rating for *participant* × *moves* × *revisits* shows significant main effects of participants ($p < .0001$), moves ($p = .009$) and revisits ($p < .0001$). The mean ratings for number of moves and revisits are shown in Table 5. Pearson correlations between the number of moves and ratings, $-.31$, and between number of revisits and ratings, $-.59$. Multiple regression of moves, revisits and rating over individual trials ($r^2 = 36.4\%$), suggests that the ratings in Experiment 2 can be explained by the perceptual cues as well as by the agent's efficiency.

Table 5: Mean ratings and Std. err. of the mean for number of moves and revisits

Moves	Rating	SEM
10	4.4	0.11
20	3.6	0.11
27	3.5	0.11

Revisits	Rating	SEM
0	4.1	0.13
1.5	4.2	0.1
3.5	3.8	0.07
7	2.1	0.09

Did the online participants in Experiment 2 prefer lucky agents? To test the hypothesis that people may use more than one way of attributing intelligence we used a *k-means* (Lloyd, 1982) and a Gaussian Mixture Model with a Bayesian Information Criterion over a 3-dimensional space of correlations of individual ratings with moves, revisits and optimality rank. Two clusters were preferred over three or one. People in the bigger cluster (19 participants) based their rating on the agent's strategy, and people in the smaller cluster relied on perceptual cues.

ANOVA of rating for *participant* × *condition* over each cluster shows that our model fits the judgements of people in the strategy cluster, but not in the outcome cluster (Table 6). Participants in the Outcome cluster rated *optimal-lucky* agents higher than to the *optimal-unlucky* ones ($p < .0001$, difference = 1.54) and were indifferent between optimal and suboptimal decisions on incomplete trials ($p = .5$). In contrast, people in the Strategy cluster were indifferent between *optimal-lucky* and *optimal-unlucky* agents ($p = .92$) and on incomplete trials rated *optimal-incomplete* agents higher than the *suboptimal-incomplete*($p = .00002$, difference 0.32).

The discrepancy between model and data reveals two different styles of attributing intelligence: attributing intelligence to efficient strategies, or to shorter paths (Figure 3 and

Table 6: Correlations of ratings with perceptual features and with optimality over invividual trials

Metric	Strategy Cluster	Outcome Cluster
Pearson, number of revisits	-.50	-.73
Pearson, number of moves	-.12	-.62
Pearson, Optimality	.65	.40
r^2, revisits, moves	35%	55%
r^2, optimality	44.3%	21.8%

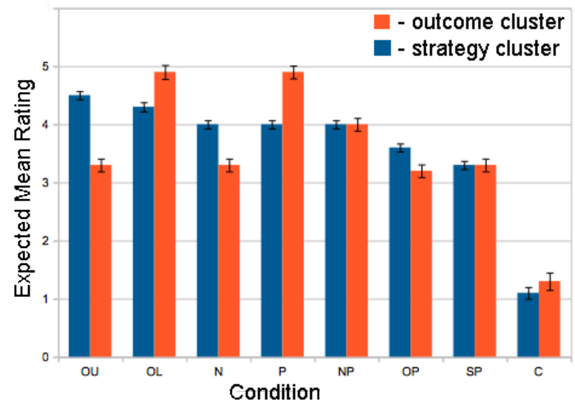

Figure 3: Comparing ratings between the two clusters *OL-optimal-lucky*, *OU-optimal-unlucky*, *N-decision noise*, *P-prior knowledge*, *NP-noise-prior*, *OP-optimal-incomplete*, *SP-suboptimal-incomplete*, *C-control*.

Table 7). Indeed, the participants' answers to the verbal question support this conclusion. We divided the participant answers to *'How did you make your decision?'* into two groups: **outcome** and **strategy**. For example, we coded a response as **strategy** if it said: *'Based on if the mouse checked every nook and cranny.'* and as **outcome** if they said *'Based on how long it took for the rat to find the treat'*. Two independent raters agreed on 27 out of 30 participants, coding 8 of them as **outcome** and 19 as **strategy**. The remaining 3 were randomly assigned to either group. Importantly, the 8 participants independently agreed on as **outcome** by both raters (using verbal measures) were also the 8 participants identified as belonging to the outcome cluster using K-Means clustering.

Discussion

Our model proposes a formal account of intelligence attribution and an apparatus for generating quantitative predictions. The model predicts that people perceive some suboptimal decisions as more intelligent than others, depending on the inferred causes of suboptimal planning, in agreement with behavioural data.

Participants were clustered into two groups: those in the Strategy group attribute intelligence based on partial results, while those in the Outcome group do not decide until they

Table 7: Mean ratings of conditions and Std. err. of the mean of people in the Strategy (S) and the Outcome (O) clusters

Condition	S	SEM	O	SEM
optimal-unlucky	4.5	0.08	3.3	0.11
optimal-lucky	4.3	0.07	4.9	0.12
decision noise	4	0.07	3.3	0.11
prior knowledge	4	0.07	4.9	0.11
noise-prior	4	0.07	4	0.11
optimal-incomplete	3.6	0.04	3.2	0.006
suboptimal-incomplete	3.3	0.04	3.3	0.006
control	1.1	0.1	1.3	0.14

have seen the consequences[1]. Both groups, however, see random actions as less intelligent than those that can be explained by causal inference.

Why did people differ in intelligence attribution? Our experiments used simple mazes that most adults can easily solve. While people in the Strategy cluster expected the agent to do what a human would do, people in the Outcome cluster did not. According to Gardner's theory of multiple intelligences (Gardner, 2011) as skills in different domains people in the Outcome cluster may attribute the agent's luck to an invisible skill (a sense of smell). Alternatively, people may assume that lucky agents must be intelligent based on a belief in a just world (Lerner, 1980). The former implies that people themselves should act rationally when solving a maze, and the latter that people should decide randomly. We plan to address this question in future work.

Another interesting avenue for future work is to investigate how children attribute intelligence. Although children recognise and value competence (Jara-Ettinger et al., 2015), children also prefer lucky agents (Olson et al., 2006, 2008). To our knowledge there are no studies evaluating children's attribution of intelligence, and at what age the abilities to reason about outcome vs. strategy emerge.

One possible criticism of our specific implementation is that it encodes space as a grid of squares and makes a new decision each time the agent moves into a new square, which may not be an accurate representation of how people navigate. An alternative – encoding the maze as a weighted graph where each room is represented by a vertex – may be a better approximation of how people represent space. In addition, more fine-grained rationalistic explanations in terms of variable costs and reward functions may be a better causal model for actions currently attributed to decision noise.

The goal of current research in Artificial Intelligence (AI) is to create intelligent computer applications, such as self-driving cars, automatic trading and intelligent energy-saving appliances designed to take over routine human decision-making. Such applications must be not only be algorithmically correct, but also must interact with people in a way humans understand as intelligent. Thus, to make better AI applications we need to measure intelligence in technical terms, and our model takes a step in that direction.

[1] As Solon advised to Croesus, *'Count no man happy until he is dead'*.

Acknowledgements

We thank the reviewers for valuable comments and feedback, which helped us make this manuscript better. This work was supported by the Center for Minds, Brains and Machines (CBMM), under NSF STC award CCF-1231216.

References

Aczel, B., Palfi, B., & Kekecs, Z. (2015). What is stupid? people's conception of unintelligent behavior. *Intelligence*, *53*, 51–58.

Baker, C. L., Saxe, R. R., & Tenenbaum, J. B. (2011). Bayesian theory of mind: Modeling joint belief-desire attribution. In *Proceedings of the thirty-second annual conference of the cognitive science society* (pp. 2469–2474).

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial vision*, *10*, 433–436.

Dennett, D. C. (1989). *The intentional stance*. MIT press.

Gardner, H. (2011). *Frames of mind: The theory of multiple intelligences*. Basic books.

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

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, 243–259.

Jara-Ettinger, J., Baker, C. L., & Tenenbaum, J. B. (2012). Learning what is where from social observations. In *Proceedings of the thirty-fourth annual conference of the cognitive science society* (pp. 515–520).

Jara-Ettinger, J., Tenenbaum, J. B., & Schulz, L. E. (2015). Not so innocent. toddlers inferences about costs and culpability. *Psychological Science*, *26*(5), 633–640.

Lerner, M. (1980). *The belief in a just world; a fundamental delusion*. New york: Plenum Press.

Lloyd, S. P. (1982). Least squares quantization in pcm. *Information Theory, IEEE Transactions on*, *28*(2), 129–137.

Olson, K. R., Banaji, M. R., Dweck, C. S., & Spelke, E. S. (2006). Children's biased evaluations of lucky versus unlucky people and their social groups. *Psychological Science*, *17*(10), 845–846.

Olson, K. R., Dunham, Y., Dweck, C. S., Spelke, E. S., & Banaji, M. R. (2008). Judgments of the lucky across development and culture. *Journal of Personality and Social Psychology*, *94*(5), 757.

Olson, K. R., Heberlein, A. S., Kensinger, E., Burrows, C., Dweck, C. S., Spelke, E. S., & Banaji, M. R. (2013). The role of forgetting in undermining good intentions. *PLoS ONE*.

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