Design Objectives for Cognitive Models of Intention Perception

David Pautler[1,2], Edwin Wirawan[1], Bryan L. Koenig[1,3], Kum-Seong Wan[1]
[1]Computational Social Cognition, Institute of High Performance Computing, A*STAR, Singapore

[2]Intention Perception LLC
[3]National University of Singapore

***Abstract*—Human recognition of intentions in other people, in robots, and in all kinds of animated characters is a capability known as intention perception. There have been a number of computational cognitive models of human intention perception, and they often differ radically in design. Yet, the published research often does not make explicit what the design objectives were. Consequently, it is difficult to compare the models fairly, or to combine their innovations.**

This article reviews three such models and then proposes a set of design objectives that model authors could explicitly adopt or offer alternatives for. The three reviewed approaches are Thibadeau's schema matching [1], Blythe et al.'s fast-and-frugal heuristics [3], and Baker, Tenenbaum and colleagues' probabilistic inference (e.g., [4]). Although each of these approaches has advantages, a consideration of characteristics desirable for computational cognitive models of intention perception suggests that the schema-matching approach, if modified and augmented, has advantages over the other two.

Keywords—intention perception; cognitive model; computational model

I. Introduction

Intention perception is the process of inferring what goals might underlie other people's actions by observing their behavior in context. Intention perception supports a wide range of social skills, such as avoiding collisions in shared spaces, coordinating work and play, and judging moral responsibility. How the human mind perceives intentions is largely unknown. Psychologists suggest at least two mechanisms, simulation involving mirror neurons [5] and mental-state representation in theory of mind [6]. Studies on theory of mind and social perception have used animations in which simple geometric figures move in ways that appear to observers as though they have intentions. A classic of this sort is the Heider and Simmel animation [7]. Despite showing only two triangles, a circle, and a house-like rectangle, it compellingly elicits in naïve human observers the perception of a bully fighting with and chasing two innocent passersby.[1] Researchers who have attempted to build computational models of intention perception often use such animations, which simplify the behavior and context to a tractable level.

Yet, beyond some agreement among researchers about stimulus design, there is little agreement — or even discussion — regarding what about this kind of cognition needs modelling. In particular, some models appear to be **product models** that replicate human judgments but do not propose a mechanism, while others are **process models** and do propose a mechanism. Distinctions like these should be made clearer in the claims about a model. Similarly, what other than replicating human judgments should count as a research success (because replication itself may tell us little about how the cognition is done)?

[1] The animation can be viewed at http://intentionperception.org/resources/heider-and-simmel-1944-an-experimental-study-of-apparent-behavior/

II. Previous Work on Modeling Human Intention Perception

In contrast to the well-developed AI literature on plan recognition (e.g., [8,9]) and activity recognition (e.g., [10,11,12]), this article focuses on the relatively nascent field of computational modeling of human perception of intentions. In the field of psychology and cognitive science, there has been research on intention perception or plan recognition (e.g., [13,14,15]) but little has been published about computational models. A review of the relevant literature found only three computational approaches to modeling human perception of intentions. All three focus on the movement of one or more figures within a 2D space. To facilitate comparison across approaches, we will explain how each addresses a rather simple example in which the computational model perceives that an agent has or had an intention to get to a specific location. Towards the end of the article we compare their fit to a set of eight desirable characteristics for computational models of intention perception we identified during a review of the literature on the competing approaches.

A. The schema-matching approach

The earliest effort used schema-matching [1]. It takes advantage of common recipe-like patterns in which agents enact several observable steps in sequence to fulfill an intention. The schema for each intention is comprised of distinct action steps, and this approach matches these steps of the schema with observations. Thus, this model abductively infers intentions from observed steps that consist of movement patterns and states.

The following is an action schema for an agent that has been observed trying to get to a door ([1], p. 136):

CON:	Went(person,door)
COLLECT:	MoveToward(person,door)
	AND notTouching(person,door)
ACHIEVE:	At(person,door)

This schema can be translated into English as "a person achieved their intention of going to a door if the person was not touching the door, moved towards it, and then touched it.""This schema has three components. CON refers to the intention attributed to the observed agent. COLLECT refers to the observable movement patterns and states that indicate the intention-related action is occurring. ACHIEVE refers to the observable movement patterns and states that indicate when the intention is achieved.

In this approach the schemas are comprised of observable propositional conditions, some corresponding to actions, each of which can only be true or false with no degree of uncertainty. The conditions include movement information (i.e., rotation, speed, and trajectory in relation to agents or objects), being at or near a person or place, touch, agent line of sight, and object states (e.g., the door is open). The model compares the conditions to the data of the incoming animation frame. This comparison is facilitated by (a) manual pre-labeling of conceptually important locations and objects in the animation, such as the identity of the agents, (b) what counts as the door, and (c) what counts as being inside or outside the house. When observations support a condition, it is marked as true and can then support intention-level perception.

Thibadeau [1] developed the approach using only the Heider and Simmel [7] animation. Therefore its schemas and conditions are relevant for that animation, but may not be for other animations. To generalize to other animations, this approach would require pre-coding the important locations and objects in the new animation, integrating them into the schemas, and developing new schemas for intentions absent in the Heider and Simmel animation.

B. The fast-and-frugal heuristics approach

The second approach is from the fast-and-frugal heuristics line of research [2]. This approach assumes that many decision-making mechanisms evolved to be quick and simple, require little information, and yet be reliable for motion patterns consistently observed during the evolutionary history of the species. Blythe et al. [3] specifically evaluate a hypothesized heuristic they label *categorization by elimination* (CBE), in which a decision maker starts with a set of candidate categories, excludes those that fail to match one cue of the stimulus, and iterates this process for different cues until just one category remains. Categories differ on how many cues they require (similar to decision trees in machine learning). This heuristic often identifies an intention without using all cues, making it "frugal".

Blythe et al. [3] compiled their set of intention categories by analyzing major domains of evolutionary fitness. They concluded that detecting the following six intentions are important across many species: pursuing, evading, courting, being courted, fighting, and playing. Given the importance of detecting these intentions, animals may have evolved abilities to detect them from relatively simple and context-free cues such as trajectory, velocity, relative heading and other observed motion cues using heuristics like CBE.

The published research of this approach does not include studies of an intention to get to a location, so we cannot draw an exact comparison with the example for the schema-matching approach. But the intention of pursuing another agent is similar: Pursuit differs from trying-to-get-to-a-static-location in that the goal location is tied to an agent who changes location in response to being chased.

To evaluate CBE, Blythe et al. [3] had pairs of undergraduates interact by each controlling an ant-like animated figure in a series of computer-based tasks. For example, in one of the experimental conditions, participants were instructed to make one's figure pursue another figure, whose controller made it evade pursuit. Other conditions instructed participants to make their figures fight, court or be courted, lead or follow, invade or guard, and play together. The resultant motions were recorded to evaluate CBE's ability to detect intentions. Before using CBE, the relative ability of motion cues must be determined to distinguish the intentions. This is done through first calculating the mean values of motion cues for intentions across animations. For example, pursuit had a mean velocity of 740 pixels/second and a mean relative angle of 345 degrees (i.e., the angle between one agent's heading and the other's position; these values were for a similar set of animations reported in [16], Table 4.) Next Blythe et al. [3] determined that the motion cue best able to distinguish the intentions was absolute velocity, followed by relative angle, then relative velocity, etc. With this preliminary work done, CBE is ready to categorize which intention is displayed in an animation. CBE starts by first evaluating the animation with regard to the most informative motion cue, absolute velocity. If the input animation depicts a typical pursuit, its movements will have high mean velocity. High mean velocities are uncommon for play, courting, and being courted, so CBE eliminates them from the set of candidate intentions. CBE then evaluates the second most informative motion cue, relative angle. Typical pursuit animations have a large relative angle as well, so CBE eliminates fleeing and fighting since they usually have small relative angles, leaving only pursuit. Thus, CBE would indicate for this animation that the agent's intention is pursuit.

Blythe et al. [3] compared the performance of CBE against a neural network and two simplified forms of multiple regression. Benchmarks were human judgments and actual intentions in new animations. For the new animations, CBE categorized intentions correctly for 57% of the animations, doing better than humans (49% correct) but slightly worse than the two types of regression (both had 60% correct) and the neural network (67% correct). CBE, however, used on average only 3.6 cues, about half that used by the neural network and regression, which used all 7 cues. CBE thus performed well despite being relatively frugal.

III. The Probabilistic Inference Approach

In the final approach, a simulated observer calculates the conditional probability that each of a set of intentions could explain an observed sequence of actions of a rational but imperfect agent [4,17,18].[2] The first observation necessary to understand this approach is its rational agent model, which is implemented as a stochastic Markov decision process (MDP). Instead of using MPDs to *determine* movement choices as usually seen in MDP research, in this approach, the observed agent adopts a pre-determined motion path while the simulated observer uses MDP to *explain* the agent's movements.

How do simulated observers use MDPs to explain another agent's movement choice? The simulated observer assumes that the observed agent used MDP to choose its movement and also that, at each time-point in the animation, the observed agent assigns to each possible action an expected utility value that reflects how much that movement will help the agent achieve its intention of getting to a goal destination. This generally boils down to the simulated observer assuming that the observed agent prefers the shortest path to its goal destination. A critical caveat is that Baker et al. [4] used *stochastic* MDPs. This means that the simulated observer assumes that, after assigning utility values to moves, the observed agent chooses where to move *randomly*—with the probability of each move being proportional to its utility value.[3] This randomness allows leeway for multiple intentions to each be a plausible explanation for a single observed action or action sequence. Some intentions might be better explanations than others for the observed action(s), however, and Bayesian inference enables comparison via likelihood estimates across intentions regarding how well each explains the observed action(s). Another critical caveat is that the modeler must select the hypotheses, or by default every location that the agent might occupy is a candidate intended destination.

Consider the following example. The modeler selects two locations, A and B, as the potential destinations of interest. A simulated observer is trying to figure out if an observed agent intends to get to location A or B. Baker and colleagues [4] decided that the observer would have no other information before the agent moves; therefore, they set equal *prior* probabilities for the two intentions: 0.5 probability that the agent is trying to get to A, and 0.5 probability that the agent is trying to get to B. That is, the observer believes it is equally likely the agent is trying to get to A or B. As the agent made a move in a direction towards A but perpendicular to B, the observer uses stochastic MDPs to calculate the two probabilities. As such, the observer's MDP calculations might indicate that an agent with the intention of getting to A would take the observed action with a probability of 0.9 (the probabilities for other possible actions will sum to 0.1). For the intention of getting to B, MDP calculations might indicate the agent would select this move with a probability of 0.2 (the probabilities for other moves for reaching B will sum to 0.8). This completes the stochastic MDP component of intention perception for this observed movement.

The simulated observer then uses Bayesian inference to update its beliefs about the overall probabilities that the agent has for each of the two intentions. These beliefs correspond to *posterior* probabilities, which are calculated separately for each intention. The *posterior* probability for each intention is equal to the observer's beliefs about that intention before seeing the action (its *prior* probability) multiplied by the probability that the agent would take that action if the agent had the intention (its conditional probability calculated above as an MDP) divided by the overall probability of the agent taking that action (across all intentions):

$$P(\text{intention}|\text{action}) = P(\text{intention}) * P(\text{action}|\text{intention}) / P(\text{action})$$

The following is the *posterior* probability calculation for the intention of getting to location A: $P(\text{A}|\text{action}) = (0.5)(0.9)/((0.5)(0.9)+(0.5)(0.2)) = 0.82$. And the following is for getting to location B: $P(\text{B}|\text{action}) = (0.5)(0.2)/((0.5)(0.9)+(0.5)(0.2)) = 0.18$. The observer thus combines its beliefs before the action with the information acquired from the action. For each subsequent action, the observer repeats this process of calculating posterior probabilities for each intention. This cumulative posterior probability reflects the simulated agent's degree of belief about how well each intention explains the whole sequence of actions.

IV. Identifying Design Objectives

Clearly the three models differ greatly in design. One reason for their differences is that the probabilistic inference model was designed as a product model and does not claim to represent a mental mechanism (according to personal communication with its authors), while the schema-matching and heuristics-learning models do. The two process models differ in what they assume to be given in the mental environment (i.e., dimensions of movement and related statistics in the case of heuristics-learning, and logical predications about the room in the case of schema-matching) and whether the model should provide a mechanism for learning its content in addition to applying it. Pursuing a product model as a step toward a process model is a common strategy, and we do not mean to criticize such a strategy but only to encourage that it be stated explicitly.

In contrast to the goals of AI plan recognition research (i.e., to be efficient and accurate), the goals of computationally modeling human cognition can be quite different, similar to the

[2] This approach gets complex at times, so our discussion simplifies some aspects for the sake of clarity. Please see original works for more details.

[3] The policy itself has a "fuzziness" parameter, beta, which "determines the agent's level of determinism. At high values of beta, agents rarely deviate from the optimal path to their goals, but at low beta values, agents' behavior is noisy, becoming a random walk at beta = 0." ([3]; Appendix p.2)

difference between engineering and reverse-engineering.[4] We identified seven such characteristics of computational cognitive models of intention perception [20], and we add one here. In the following paragraphs, we briefly describe these characteristics and explain how in principle the schema-matching technique, together with some augmentations, can achieve them. The desired characteristics mostly emphasize cognitive and experiential aspects of intention perception instead of considerations such as efficiency. Further justifications for these objectives are provided in [20]. In the next section, we explain why the heuristics and probabilistic inference approaches have difficulty achieving these characteristics.

The first objective is that the system should not be artificially limited to inferring intentions but also *be capable of inferring physical causes*, just as people find that some nonliving things (e.g., leaves blown by wind) can move as if alive but later are realized not to be. We should not artificially restrict the set of candidate hypotheses to intention types, so that we do not force the system into correct answers when it otherwise might not select them. Second, the system should entertain *multiple* competing ascriptions, not force itself to adopt one of many before evidence clearly supports it (unless the system can backtrack and unless there is evidence that humans show processing delays where backtracking would occur). Third, it should form hypotheses *as the action unfolds*, as people do, not only when the action finishes. Fourth, the animations should contain *representative richness* from real environments, such as obstacles. Fifth, the system's inference rules should reflect *folk theories of psychology and physical causality*, not scientific theories. Sixth, the system should allow for new types of goals and physical causes to be added through *incremental augmentation*, not recomputation over all prior samples. Seventh and most importantly, inference rules should use *known psychological cues* that people use when interpreting similar stimuli (e.g., spatial context [21]. Finally, a newly-recognized objective is that *the model, not the modeler, should formulate its hypothesized intentions* (and causes), although this ability must necessarily be limited to the types of intentions (and causes) that the modeler encodes.

V. EVALUATING THE MODELS ON THESE OBJECTIVES

The first objective we identified above is that the model should be able to infer physical causes such as collisions that involve no agent whatsoever. This rules out an MDP approach because physical causes cannot be assigned utility values. It is also problematic for the heuristics approach, because incorporating each new attribution requires revisiting all samples and adjusting the discriminating ranges of motion statistics for all attributions.

The probabilistic approach meets the second and third objectives—to manage multiple alternate hypotheses as the action unfolds—as well as the schema-matching approach does. The heuristics approach cannot satisfy these objectives because it must wait until the action completes in order to gather and average its motion statistics.

Our fourth objective—representational richness—is consistent with the ecological perspective of fast-and-frugal heuristics, but overfit of its motion statistics to one environment can limit generalization to other environments [18]. The probabilistic inference approach can accommodate rich environments, and one of the seminal findings on spatial context as a cue arose from this work [17].

Our fifth objective, to use folk theories of psychology and physical causality rather than scientific ones, applies only to models that represent intermediate inferred states. The heuristic approach does not use such states. Although MDPs cannot be used to infer physical causes, the probabilistic inference approach uses other Bayesian techniques to do so [22]; yet, so far no model of this kind has integrated the inference of intentions and physical causes.

As already mentioned, incorporating new types of intention in the heuristics approach requires revisiting all samples and adjusting the discriminating ranges of motion statistics for all attributions. This need for recomputation makes it difficult for the approach to satisfy our sixth objective of gradual learning. The probabilistic inference approach might be able to do so because related techniques have been developed (e.g., [23]), but their successful implementation for intention perception has yet to be demonstrated.

Our seventh objective was for the model to use the same motion cues as people do. It might be problematic for utility-oriented approaches when psychological cues lack clear utilitarian interpretations. For example, the animation technique of adding a slight wiggle to a figure's motion might make the figure appear more lifelike and thus more capable of pursuing an intention through its movement, but wiggles violate a utilitarian strategy of taking the shortest path.

The eighth objective is that the model rather than the modeler should select the candidate attributions. This is difficult for the probabilistic inference approach, at least when applied location-based intentions. The reason is that every possible location in the space must be considered a candidate destination unless a mechanism is added for identifying candidates. Schema-matching has an advantage in that it includes a mechanism for identifying candidates (by constraining features of candidates) without assuming every location might be a candidate.

In contrast, the schema-based approach is consistent with our objectives, if the implementation picks up features of the figures and their motions, and if it incorporates known psychological cues when computing the likelihoods of competing attributions.

[4] The aim of computational cognitive models is scientific reverse-engineering of the mind in order to develop a mechanistic theory to explain observed human performance. For example, unlike in software engineering, there is low priority on making the mechanism tractable for CPU architectures and real-time interaction; instead, the corresponding, moderately-high priority is to design the mechanism so that it responds similarly to circumstances in which the observed human performances occurred, usually according to the pace of time in a simulation. Computational cognitive models like ACT-R [19] adhere to the same priorities.

VI. COMPARISONS WITH AI PLAN RECOGNITION

Our ability to "perceive" intentions is similar to the functionality that AI plan recognition systems aim for but with three differences: 1) the system inputs are constrained so that they must be derivable from visual access to the world, 2) the processing must use plausible cognitive mechanisms, and 3) "ground truth" is human judgments about what intentions are most likely from the evidence, which may not match the planning agent's actual intentions.

From an AI perspective, the three approaches to computationally modeling intention perception each have their own advantages and disadvantages. Thibadeau's [1] schema-matching model uses a technique similar to that in AI plan recognition known as abductive inference, or "running the planner in reverse", first pioneered in the 1970s [24,25]. AI plan recognition systems generally use a stream of atomic symbols as input, as did Thibadeau's model. Such tokenized inputs enable the use of category grammars to represent schemas or plan structures, and when carefully structured, these grammars are tractable [8,9]. But when using non-tagged animations to simulate visual access to the world, the inputs are a stream of frames, each a structure of shapes with positions, sizes, colors, and so on. Matching these non-atomic structures and enforcing constraints on their contained values across frames is beyond what categorical grammars can do.

We do not know of any AI plan recognition work using a technique similar to fast-and-frugal heuristics, although the technique is known in machine learning as "top-down induction of decision trees," and the type of tree in this case is a "classification tree" [26].

Just as Tenenbaum's group has applied Bayesian inference to cognitive models of intention perception in order to rank competing goal attributions, several AI plan recognition researchers use Bayesian inference to rank competing hypotheses [9,27]. From a modeling perspective, we identify two issues with using the technique. First, Bayesian inference applied to inverse planning requires a prior probability value for every possible state that might lead to the inferred goal state. When the goal state involves a predication like "be at a certain position", then priors for nearby states can plausibly be estimated using Euclidean distance from that end position. But predications that are not based on metric space, such as "to catch (another moving agent)" or "to help" have less obvious and perhaps less plausible methods of providing estimated priors. Second, as already mentioned, psychological research has revealed several visual cues and other influences that suggest intentionality (or agency), and some resist the rational interpretation that an MDP approach would require. For example, a figure moving toward an object but facing away from it is judged to be less likely to be an agent than a featureless circle with no obvious orientation but the same movement path [28]. It makes intuitive sense that an agent would want to face a target if approaching it purposefully, but creating a MDP-like utility-oriented story of why this would be so seems quite complicated.

These issues suggest that probabilistic inference models may have difficulty scaling to model the various visual cues identified in the literature or representing how these cues influence confidence in competing attributions as events unfold over time.

To us, the best match for these objectives is the schema-matching technique. In particular, implementing it with an incremental parser [20] allows the system to generate hypotheses as the action unfolds, unlike the fast-and-frugal approach that analyzes only completed sequences. Also, using grammatical rules [20] allows for combinatorial richness in matching a practically unlimited set of potential sequences, which the probabilistic approaches so far used for intention perception cannot do.

REFERENCES

[1] R. Thibadeau, "Artificial perception of actions," Cognitive Science, vol. 10, 1986, pp.117–149.

[2] G. Gigerenzer, P. M. Todd, and the ABC Research Group, Simple heuristics that make us smart, New York: Oxford University Press, 1999

[3] P.W. Blythe, P.M. Todd, and G.F. Miller, "How motion reveals intention: Categorizing social interactions," in Simple heuristics that make us smart, G. Gigerenzer, P. M. Todd, and the ABC Research Group, Eds. New York: Oxford University Press, 1999, pp. 257–286.

[4] C.L. Baker, R. Saxe, and J.B. Tenenbaum, "Action understanding as inverse planning," Cognition, vol. 113, 2009, pp. 329–349.

[5] V. Gallese and A. Goldman, "Mirror neurons and the simulation theory of mind-reading," Trends in Cognitive Sciences, vol. 2, 1998, pp. 493–501.

[6] R. Saxe, "Against simulation: The argument from error," Trends in Cognitive Science, vol. 9, 2005, pp. 174–179.

[7] F. Heider and M. Simmel, "An experimental study of apparent behavior," The American Journal of Psychology, vol. 57, 1944, pp. 243–259.

[8] D.V. Pynadath and M.P. Wellman, "Probabilistic state-dependent grammars for plan recognition," Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence, 2000, pp. 507–514.

[9] C.W. Geib and R.P. Goldman, "A probabilistic plan recognition algorithm based on plan tree grammars," Artificial Intelligence, vol. 173, 2009, pp. 1101–1132.

[10] L. Liao, D. Fox, and H. Kautz, "Location-based activity recognition," Proceedings of the Neural Information Processing Systems, 2005.

[11] A. Hoogs, and A.G.A. Perera, "Video activity recognition in the real world," Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence, 2008, pp. 1551–1554.

[12] D.L. Vail and M.M. Veloso, "Feature selection for activity recognition in multi-robot domains," Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence, 2008, pp. 1415-1420.

[13] D. Newtson, "Foundations of attribution: The perception of ongoing behavior," in New Directions in Attribution Research, Volume 1, J.H. Harvey, W.J. Ickes, and R.F. Kidd, Eds., Hillsdale, NJ: Lawrence Erlbaum Associates, 1976, pp. 223–247.

[14] D.E. Rumelhart, "Notes on a schema for stories" in Representation and Understanding: Studies in Cognitive Science, D.G. Bobrow and A. Collins, Eds., New York: Academic Press, 1975, pp. 211–236.

[15] C.F. Schmidt, "Understanding human action: Recognizing the plans and motives of other persons," in Cognition and Social Behavior, J. Carroll, and J. Payne, Eds., Hillsdale, NJ: Lawrence Erlbaum Associates, 1976, pp. 47–67.

[16] H.C. Barrett, P.M. Todd, G.M. Miller, and P.W. Blythe, "Accurate judgments of intention from motion cues alone: A cross-cultural study," Evolution and Human Behavior, vol. 26, 2005, pp. 313–331.

[17] C.L. Baker, N.D. Goodman, and J.B. Tenenbaum, "Theory-based social goal inference," in Proceedings of the Thirtieth Annual Conference of the Cognitive Science Society, 2008, pp. 1447–1452.

[18] T.D. Ullman, C.L. Baker, O. Macindoe, O. Evans, N.D. Goodman, and J.B. Tenenbaum, "Help or hinder: Bayesian models of social goal

inference," Advances in Neural Information Processing Systems, vol. 22, 2010, pp. 1874-1882.

[19] J.R. Anderson, M. Matessa, and C. Lebiere, "ACT-R: A theory of higher level cognition and its relation to visual attention," Human-Computer Interaction, 12(4), 1997, pp. 439-462.

[20] D. Pautler, B.L. Koenig, B.K. Quek, and A. Ortony, "Using modified incremental chart parsing to ascribe intentions," Behavior Research Methods, vol. 43, 2011, pp. 643–665.

[21] P.D. Tremoulet and J. Feldman, "The influence of spatial context and the role of intentionality in the interpretation of animacy from motion," Perception and Psychophysics, vol. 68, 2006, pp. 1047–1058.

[22] J. Hamrick, P.W. Battaglia, and J.B. Tenenbaum, "Internal physics models guide probabilistic judgments about object dynamics," Proceedings of the 33rd Annual Conference of the Cognitive Science Society, 2011, pp. 1545-1550.

[23] A. Stolcke and S. Omohundro, "Inducing probabilistic grammars by Bayesian Model merging," Grammatical Inference and Applications, 1994, pp. 106─118.

[24] R. Wilensky, "PAM: a program that infers intentions," Proceedings of the 5th international joint conference on Artificial intelligence, vol. 1, 1977.

[25] C.F. Schmidt, "Understanding human action: Recognizing the plans and motives of other persons," in Cognition and Social Behavior, J. Carroll, and J. Payne, Eds., Hillsdale, NJ: Lawrence Erlbaum Associates, 1976, pp. 47–67.

[26] J.R. Quinlan, "Induction of decision trees," Machine Learning vol 1, Kluwer Academic Publishers, 1986, pp. 81-106.

[27] P. Singla and R.J. Mooney, "Abductive Markov logic for plan recognition," Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence, 2011, pp. 1069–1073.

[28] P.D. Tremoulet and J. Feldman, "Perception of animacy from the motion of a single object," Perception, vol. 29 no. 8, 2000, pp. 943-952.