A Cognitive Robotics Approach to Identifying Perceptually Indistinguishable Objects

John F. Santore and Stuart C. Shapiro
Department of Computer Science and Engineering
and Center for Cognitive Science
University at Buffalo, The State University of New York
201 Bell Hall Box 602000
Buffalo, NY 14260-2000
{jsantore|shapiro}@cse.buffalo.edu

Abstract

This paper describes a cognitively motivated computational theory of identifying perceptually indistinguishable objects (PIOs) based on a set of experiments, conducted with human participants, which were designed to identify the knowledge and perceptual cues that people use for this purpose. Identifying a PIO in this context means connecting sensor data from some physical object either to a new mental level symbol or to the correct preexisting one. We discuss the experiments with people, several base and intermediate cases in the identification process and the knowledge that is needed for the general case. A summary of an algorithm for identifying PIOs is included.

Introduction

People often encounter objects that are perceptually indistinguishable from objects that they have seen before. When this happens, how do they decide whether the object they are looking at is something never before seen, or if it is the same one they encountered before? How should an agent, a person or a robot, identify a perceptually indistinguishable object (PIO) the way that people often can?

Identifying a perceptually indistinguishable object means deciding if the object just encountered is a new, never before seen object, or if it has been previously encountered, which previously perceived object it is. Identifying a PIO is a sub-problem of object *identification* rather than object *recognition*. Object recognition is defined in a computational vision textbook (Jain et al., 1995) as the process of finding and "labeling objects [in the real world] based on known object models", that is object recognition in computer vision is the process of deciding what category an object belongs to. By object identification, we mean deciding which individual object it is, rather than deciding what category of objects it belongs to. When an agent perceives an object, it first uses its object recognition system to decide what category of thing it is, then it uses its object identification routine to choose and anchor a mental concept to the object. The object identification system uses non-perceptual properties and background knowledge to identify the object as being the same one that the agent perceived at some previous time or to identify it as something new that the agent has never thought about before. This identification of objects across time is a necessary part of any solution to the symbol anchoring problem.

Sometimes identifying PIOs seems effortless. Consider the case in which a man has a pile of stamps of the same design in his drawer. He opens the drawer, takes out a stamp and puts it on an envelope and mails the envelope. The next day, he needs to mail another envelope and so needs a stamp. He opens the drawer and takes out a stamp that looks just like the one used the day before. The man never considers whether it is the same stamp, even though there is no perceptual difference between the two stamps. He is able to quickly and easily decide that this is a different stamp.

Sometimes the task is not as easy. Consider the case in which a woman puts her glass down on a counter at a cocktail party. When the woman returns to pick her glass up again and finds more than one glass on the counter, the woman will often find it difficult to decide which glass is hers. Sometimes the woman is not able to decide with enough certainty which glass is hers even after thinking about it.

This paper proposes a cognitively motivated computational theory of how agents, particularly artificial embodied agents (such as robots) can use reasoning to identify PIOs the same way humans do. Others have used the performance of humans at a particular task to design a robot that can do the same task in the same manner (and as well) (Trafton et al., 2004). The scope of the human trials used to formulate this theory is a little larger than that used by (Trafton et al., 2004), but the intention is similar: a theory based on human cognitive methods will allow us to develop an agent capable of doing a task, in this case identifying PIOs, in a manner similar to the way humans do. Basing our algorithm on human performance gains us the twin advantages of both giving us a basis for a computational solution to a very difficult common sense problem, and making our artificial agent easier to work with and interact with because it does the PIO identification task in a way that humans would expect it to.

Let us examine what is required to identify an object in the world. An embodied agent gathers information about its world by observing the world with its sensors and using its effectors to move itself to a better observation point when necessary. From its observations, the agent forms beliefs about the objects in the world. People use these beliefs in conjunction with their common-sense rules about the world to help them identify objects in the world. Identifying PIOs relies entirely on this mechanism since there is no sensory information that will help to distinguish one PIO from another.

The designer of an artificial embodied agent must provide the agent a mechanism for both creating beliefs from observations and using those beliefs to reason. In the remainder of this paper we will assume that a mechanism for reasoning from observations to beliefs, such as (Shapiro, 1998, p138) has been provided. The focus will be on reasoning with beliefs about the world in order to identify PIOs.

Our agent's beliefs and reasoning are based on an intensional representation (Maida and Shapiro, 1982). Intensional representations model the sense (Frege, 1892) of an object rather than the object referent, itself. The terms of our representation language, SNePS (Shapiro and Rapaport, 1992; Shapiro and the SNePS Implementation Group, 2002), denote mental entities. Some such entities are propositions; others are abstract ideas; others are the agent's "concepts" or "ideas" of objects in the world. This is important for the task of identifying PIOs, because before the identification task is complete, the agent may have two mental entities, e_1 and e_2, that it might or might not conclude correspond to the same object in the world. It is in a similar situation as George IV, who "wished to know whether Scott was the author of Waverly" (Russell, 1905, p 108).

We will use "object" to refer to an object in the world and "entity" to refer to a mental entity that is the denotation of a SNePS term. The task is "identifying perceptually indistinguishable objects", because the agent has perceived an object in the world that might or might not be the same as a previously perceived object in the world. Its task is to decide whether the entity e_2 (think of "the author of Waverly") corresponding to the newly perceived object is coreferential with an entity e_1 (think of "Scott") that corresponds to a previously perceived object.

When an agent wants to identify an object, it must accomplish two things. First the agent must identify what *kind* of object it is sensing. The agent should use its sensors and its understanding of what things "look like"[1] to those sensors to identify the type or kind of thing that it (the agent) is looking at. The agent must then reason about what *actual* object it is looking at. A simple solution, and one easy to implement, might be to assume that all objects that look the same, are in fact the same object, but this is clearly not the case. A better solution, discussed in (Shapiro and Ismail, 2003), is whenever an agent looks for an object with certain properties, it conceives of a new entity with only those properties. When the agent finds a real world object that has those properties, it should recognize if it already has a mental entity corresponding to the object it just found. If it does have such an entity, then it should adopt a belief that the object looked for is the same as the one that was found. This approach has two drawbacks. First it sidesteps the issue of how the agent reasons about object identity. Second, even though the agent may now correctly believe that the two entities refer to the same object in the world, there are times when a new entity is unnecessary. It would be better to simply use the agent's original entity if, at the time of the second sighting, the agent can instantly reason that the object is the same as one it has seen before.

In the remainder of this paper, we will first briefly describe an experiment with humans upon which our theory is based, then we will discuss four base cases and intermediate cases in the identification of PIOs, and then introduce the knowledge used in more complex cases. We will then sketch an algorithm for identifying a currently perceived object as being the same or different from any PIOs that the agent has been encountered previously. Finally some conclusions drawn from the work so far are discussed.

Human performance as a base for the theory

In order to understand how humans identify PIOs (and how well) we conducted an experiment with 68 human participants (Santore et al., 2004). The experiment was designed as a protocol analysis experiment (Ericsson and Simon, 1984) in which the participants were asked to describe what they were doing and why they were doing it as they performed the tasks in the experiment. The utterances were recorded to tape and later transcribed as completely as possible including things like false starts and other disfluencies.

The participants interacted with a video-game like environment giving them a first person view of the world. (See figure 1 for a participants view of the experiment.) This is the same first person view of the world that our agent, a simulated robot, uses to interact with

[1]Perceiving could be done using any sense, but in this paper we will often use "looking" as a euphemism for any type of perceiving.

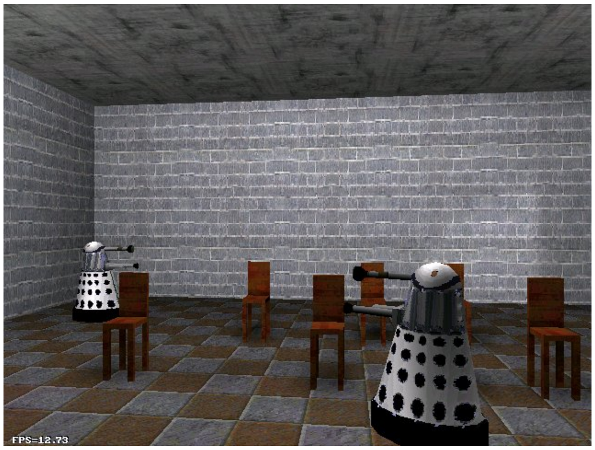

Table 1: A sample view of what a participant in the experiment might see.

the world. Each participant was randomly assigned two tasks in which they had to identify PIOs. Participants were assigned one “counting task” and one “following task”. In a counting task, participants counted either moving robots, or stationary glasses. In a following task, participants followed a tour guide, either a robot or a person, through a virtual suite of rooms in which several distractors were also wandering.

In order to count objects, a person must be able to identify them. One cannot get a successful count of objects if one cannot reason about the identity of objects. Likewise when following an object, one must be able to identify the object at the various stages of the following process. When the object to be followed and distractor objects have the same appearance, it is even more important to identify the tour guide at all times. The requirement to identify the tour guide in order to follow it was even more apparent when a participant lost the tour guide for a few moments. If the tour guide entered a room before the participant, the participant had to identify the tour guide in order to follow it through the remainder of its journey.

The strategies of the successful participants (those who gave the correct count or followed the tour guide to the end of its route) were used as the basis of the computational theory. The strategies considered are those directly stated by participants such as participant 37’s “I’m trying to see them simultaneously” or participant 48’s statement “I was just moving fast in and out of the rooms before the robots can actually move out of the rooms.” Strategies suggested over the course of a longer utterance were also used in the development of the theory, though space limits prevent reproducing them here.

Base Cases in the Identification of PIOs

What makes a base case.

The experiment with human participants described above showed that there are four conditions under which human agents find the identification of perceptually indistinguishable objects to be very easy. We’ll call these four conditions the base cases of the identification task. Participants in the experiment actively tried to put themselves into a position where they could use one or more of these base cases to identify the PIOs in the simulated world of the experiment.

When the computational agent identifies a perceptually indistinguishable immobile object using a base case, it does not form a new mental entity for the object and then try to find (or “remember” to use a commonsense colloquialism) an existing entity with an equivalent extension. The agent only creates new entities as needed for cognizing information (Maida and Shapiro, 1982). The object that the agent is perceiving is either *the one* that it has seen at this location before, or a new, never-before-perceived object. If the object is *the one* that it has seen here before, then the agent ought to use the original mental entity for it and not conceive of something new that the agent believes is really the same thing in the world. If the object is a newly encountered one, a new mental entity is created for the newly encountered world object that the agent conceives of.

Human participants support the claim that new entities are not created when using base cases to identify PIOs. As an illustrative example, participant P55 is using a base case to identify PIOs when he says: "ahhh ok, it seems to me I’ve already seen this room and counted this one glass here." Upon seeing the glass, P55 immediately identifies it as *the one* he has seen before. Contrast this to P33 who is not using a base case to identify his PIO in the following utterance: "Where did the robot go? I think *this* is *the one*." In this statement, P33 indicates that he has one entity for the robot he is looking at now (vocalized as *“this”* in the transcript) and a separate entity for the robot he is looking for (vocalized as “*the one*” in the transcript.)

Note that these base cases are all based on the beliefs of the agent rather than facts about the world. For example when using the immobile objects base case, it is the agent’s belief that the object is immobile that allows the case to be used. Each of the other base cases is based on the agent’s beliefs about the objects that it is seeing. A last characteristic of base cases of identifying PIOs is that so long as the agent’s beliefs that the base case holds is correct, the base case is a nearly foolproof method of identifying the PIO.

Base Case 1: Simultaneous perceptions.

If an agent perceives two perceptually indistinguishable objects in its sensory field at the same time, the agent can trivially conclude that the two are not the same object.[2] Unlike some of the base case strategies, participants were conscious that they were using this strategy of simultaneous perceptions while they used the strategy and discussed its use while they used it. While counting moving robots, P37 states “I’m trying to see them simultaneously.” P4, while doing the same task, is even more explicit when she states “The same two robots at the same time, so I know that there are at least two robots here.”

Base case 2: Objects with a unique appearance.

If the agent believes that an object has a unique appearance and there are no other PIOs in the world, then the agent can instantly identify the object. The agent has only one entity for an object of this appearance and believes there is only one such object, so the agent should immediately use this entity to refer to the object whenever the object is encountered. Thus, like other base cases, the agent can and ought to use its original entity for the object in this case.

Participants were often aware enough of their use of this assumption of unique appearances to try to verify the assumption when possible. P15, when counting robots when there were two groups of perceptually indistinguishable robots, says “And I see the clown dalek here. aaand the little black and white one I don’t.. annd a clown here - is that the same clown?”

The belief that an object is unique is subject to being “overruled” by a case of simultaneous perceptions. If an agent believes that an object has a unique appearance, but then sees two PIOs with that appearance, the agent will have to put aside the belief that there is only a single object with this appearance. For example P9 in the tour guide following task ended up following a distractor after never seeing more than one robot at a time. When P9 sees another robot, he makes the following statement: “So which one am I supposed to follow? There are two robots now....”, Indicating that he clearly abandons his belief in the unique appearance of the robot. That P9 had that belief is born out in the retrospective interview done a few moments later with the following exchange:

> **Experimenter:** What strategies did you use and why did you choose the robot that you chose to follow?
>
> **P9:** Well I had no clue that it’s a different robot. If I had known that there were more than one robot, I probably would have been more careful but I didn’t know.

The use of a single mental entity for an object believed to have a unique appearance was particularly noticeable when the participant’s assumption that an object has a unique appearance turned out to be incorrect. While trying to follow a robotic tour guide who turns into a room on the left of a corridor P42 says “I can’t catch up with you. Where are you going?!” And then a second later as a second robot emerges from a room on the right of the corridor a little further from the participant “He came up that way. How did he come up that way?”

Base case 3: Immobile objects.

Immobile objects are defined here as those objects which cannot move or be moved. We’re also including those objects which humans expect cannot be moved, even if such an object might be moved by using a rarely used technique. For example, people do not expect things like houses and other buildings, or even large trees, to be moved intact from one place to another, even though it is possible.

Since the location of an immobile object does not change, location is the most important feature that allows an agent to identify immobile PIOs. In order to identify an immobile PIO, the agent must first recognize what kind of object it is perceiving. Next, the agent needs to reason, or realize that objects of this kind are immobile. Then the agent cognizes the location of the object. At this point the agent can identify the object. Either the agent knows about an immobile object of this kind at this location, in which case it now identifies the current object using the entity that denotes that previously seen object, or the agent has never encountered one of this kind of object at this location, in which case the agent identifies the object as a newly encountered object and creates a new entity to refer to the object which has that description and is at that location.

The human-participant experiment supports the claim that location is of paramount importance in identifying immobile PIOs. Human participants find the use of location information so intuitive that they rarely notice it at the conscious level. When human participants were asked to discuss what they were doing and why while counting immobile PIOs, they never mentioned the object’s location as being important during the task, even if they were clearly using location information. However, when asked in a retrospective interview, participants were able to articulate that location information was what they were relying on. The following exchange is representative. It was taken from a retrospective interview following an experimental task in which participants were asked to count glasses; The glasses were immobile in the simulated environment and recognized as such by participants.

> **Experimenter:** how were you able to distinguish between the glasses even when they looked the same?
>
> **P33:** ah because they are lying in the different rooms. That’s why. They are different.

[2] We are ignoring the use of illusions with mirrors and other deliberate attempts to make a single object appear to be multiple objects in this dissertation.

The use of the original entity is supported by the human-participant data in the immobile object case as well. While performing the glass-counting task, no participant who was sure about what room he/she was in expressed doubt about the identity of a glass. The glass was either referred to as the same one seen previously or it was referred to as a new glass. This contrasts with mobile objects, where participants often clearly seem to have more than one entity for an object and can talk about both entities. To reuse the example from subsection above in more context, when following a robotic tour guide in a suite with several perceptually indistinguishable distractors, P30 briefly loses the tour guide robot and then makes the following statement "Where did the robot go? I think this is the one"

The use of a single entity is contingent on an agent correctly identifying its current location. Our participants were vulnerable to mistaking one room for another if the two looked similar. Kuipers and his colleagues (Kuipers and Byun, 1991; Kuipers and Beeson, 2002) call this sort of mistake "perceptual aliasing" and have discussed the problem and a solution for robotic agents. When our participants fell victim to perceptual aliasing, use of location information to identify immobile objects was fallible. Sometimes participants would notice the possible aliasing, such as when P20, while counting glasses, says "I'm just, just curious to whether or not this is the same room. So I'm going to go back and retrace that, my steps." Participants who fell victim to perceptual aliasing and never realized it generally failed at the identification and thus the counting tasks.

Base Case 4: Continuous viewing.

Pollock has discussed reidentification of objects, a subproblem of identifying PIOs (Pollock, 1974). He notes that an object under continuous observation can be reidentified at a later time as being the same object, in particular, that "continuity of appearance is a logical reason for reidentification."

Continuous viewing of an object also appeared in the human-participants trials as a base case for identifying PIOs. Continuous viewing, like location, is used to identify an object as being the same as a perceptually indistinguishable object seen earlier (Pollock's reidentification). This ease of identification of object while under continuous observation seems to be implicitly assumed in Coradeschi and Saffiotti's *Track* functionality(Coradeschi and Saffioti, 2003).

More concretely, the continuous viewing case applies if an agent views an object at position p_1 and later observes an object that is perceptually indistinguishable at position p_2. If the agent has continuously viewed the object as it moves from p_1 to p_2, the agent may assume with great certainty that the object it is currently seeing at p_2 is the same object that it originally saw.

Human participants tried to use this base case as often as possible when asked to follow a virtual robotic tour guide through a suite of rooms that also contained several perceptually indistinguishable robots serving as distractors. Like the simultaneous-perceptions case, participants were aware enough of this strategy to report it while performing the task. P7, after an early bit of difficulty, says "And I am following him very closely. And I am not going to lose sight of him this time." P23, is also very specific about using continuous viewing: "So I'm just staying, uh, close to this robot keeping my eye on him."

Intermediate cases of PIO identification.

What makes an intermediate case.

It has been pointed out[3] that the base cases described above represent primarily perceptual cases of identifying PIOs and that there were likely to be simple cases that do not rely on purely perceptual mechanisms for the identification of PIOs. When we examined the human performance data collected from the experiment, we saw evidence of non-perceptual cases that are similar to the base cases. In fact, for every perceptual base case, there is at least one non-perceptual simple case which can be closely identified with the base case. We will call these associated non-perceptual cases "intermediate cases". They are so named because they are between the largely perceptual base cases and a mostly cognitive general PIO identification mechanism. Like the base cases, intermediate cases are chosen based on the beliefs of the agent, not something that actually occurs in the world. Therefore, like the base cases, the intermediate cases might lead the agent to make an incorrect identification if the belief that triggered the use of an intermediate case was erroneous.

Intermediate Case 1: rapid perceptions

The first intermediate case is related to the base case of simultaneous perceptions. In that case, seeing multiple objects at once was sufficient to assure that there are multiple objects in the world. In the rapid perceptions case, on the other hand, the objects (usually two of them) are not perceived at the same time, but rather in rapid succession, with no PIO encountered between the two perceptions. As in the case of simultaneous perceptions, the rapid perception case is used to prove to the agent that two objects are not the same.

Participants in the experiment sometimes used rapid perceptions to disprove a hypothesis of unique appearance, as P18 does in the following transcript excerpt.

> Going into the next room, there is a multi-colored robot, and one who looks like the last one. I'm turning back, that robot is still in the other room so I know that these are two distinct robots.

[3]Our thanks to the anonymous reviewer of another paper who did so.

Prior to this excerpt, P18 has seen only one robot, a silver-gray robot. As he enters another room, P18 sees a "multi-colored" robot as well as a silver-gray robot. In order to identify this silver-gray robot as a new, never before seen robot, P18 looks back toward the place where he last saw a silver-gray robot. When he sees a silver-gray robot in the previous location, P18 assumes (correctly in this case) that the robot seen in the current room is different from the one he looked back to see.

In order to take advantage of this rapid-perceptions case, an agent must see an object Ω, then must turn at least as fast as objects of type Ω can move, turning no more than 180°, and must see another object that looks like Ω. The agent must turn at least as fast as the object can move because if the agent turns more slowly, there is the chance that the object will be able to move to the new position before the agent views it. For example, if the agent turns so slowly that the object can move ahead of the agent's gaze, then this intermediate case does not hold. Likewise, if the agent turns more than 180°, then the object Ω could move around behind the agent and be seen again when the agent stops turning. In this case, the agent cannot use the intermediate case of rapid perceptions. However, if all of the conditions hold, the agent can determine with great confidence that there are two different PIO objects in the world.

Intermediate Case 2: Locally Unique Objects

An agent can often easily identify an object without the object being truly unique in the world, or even believed to be by the agent. It is only necessary for the agent to believe that an object is unique in the current context. For example, suppose you know identical twins, but one of them is in the army posted abroad for the next six months. If you see someone that looks like these twins in town tomorrow, you can immediately assume that this person is the second twin. As with the unique items base case discussed above, the simultaneous perceptions base case described above will trump a belief that an object has a a locally unique appearance. In such a case, the agent must put aside the belief that an object has a locally unique appearance. Continuing the above example, if you knew that one of the twins was stations abroad, but you saw two people that looked just like the twins, you must then put aside the assumption that the twins' appearance is locally unique. You may realize that something unusual happened and ask the twin about his sudden return, but you must give up the assumption of of locally unique appearances.

Participants seemed to use this assumption that locally unique objects can be effortlessly identified as the same thing seen previously (using the same mental entity) when they could. Sometimes the assumption of local uniqueness of appearance would be limited to a single room. For example, P12, while following a robotic tour guide in a suite of rooms with PIOs as distractors says "I'm stuck, okay but there is only one robot so I can follow *it*." P23, doing the same task, says something similar "There aren't any other robots in this room so *it*'s a little easier to follow." In both cases, the participants thought that the robot that they were following was the only object with the robotic appearance in the room. When entering the room they see only one object with that appearance and so they automatically identify this robot as the one they have been following.

Intermediate Case 3: Stationary Objects.

The next intermediate case is related to the base case of immobile objects. Stationary objects are those objects that cannot move themselves and are not easily moved by a breath of air. A helium-filled balloon is not a stationary object, even though it cannot move itself. On the other hand, many of the objects that we come into contact with in our daily lives are stationary: Lamps, computers, textbooks, and similar objects are all stationary objects. Their position will not change (or at least people do not expect it to change) unless there is an animate object to move the stationary object. P31 explicitly pointed this out in a retrospective after counting glasses in task 1 of the experiment:

> **Experimenter:** What strategies did you use to do this task?
>
> **P31:** Mmm I guess I just kind of based it on the fact that they would be stationary throughout the rooms and there was nobody else in there.

In the absence of a mover, stationary objects can be treated just like immobile objects; that is, location becomes the paramount criterion for identifying the object. The lack of another agent capable of moving a stationary object is something that a PIO identifying agent must reason about.

Intermediate Case 4: Continually 'Perceived' Objects

It is well known (Johnson, 1998) that young children will identify objects that are briefly occluded as the original objects. The participants in our experiment overwhelmingly did likewise. Though participants may have briefly lost sight of the focus object by looking away or having the object occluded, the participants nonetheless knew where the object was and looked for it "where it ought to be" when they viewed the object again. Most of the time, participants were not even aware that they had lost sight of the object in question.

Identifying PIOs in general.

While identifying PIOs is trivial when one of the base or intermediate cases can be applied, when one of these cases does not hold, the task can be much harder. An agent usually requires several more pieces of knowledge to identify PIOs in the general case. If people need to identify an object as the mental entity e, experiments show that they use knowledge of how rare or common they believe objects that look like e are. They will also use their beliefs about how fast the objects like e can move and the time between the time, t_1, that the agent last encountered an object it thinks might have been e and the time, t_2, that the agent sees e itself. Humans will also use the motivations of the object being identified if they can infer any.

Humans participants seem to use general beliefs formed from observations of the world. The most salient is information about the class of objects to which the PIOs being identified belong. These include things like: how fast or slow do objects of this kind move? [P8 while counting moving robots: "I think that's the guy I counted already because, ah well he- uh couldn't have moved that fast"] Has an object of this kind ever been known to change speed? [P6 asked in a retrospective why participant chose to follow a particular robot: "It's possible that it changed speeds, but it didn't really appear to do so throughout the game"] Have I ever identified more than one object that is perceptually indistinguishable from this one? [P18 while counting robots in a condition with two distinct groups of perceptually indistinguishable robots: "Because I thought maybe the multicolored robot had traveled, into that last room that I just searched, but it looks like there are two multi colored robots."]

Human participants also use information from observations of the specific objects being identified. Beliefs formed from these observations include beliefs about where and when the agent last encountered a PIO that the participant believes might be the PIO that the participant is currently looking at. [P25 counting robots with two distinct groups of perceptually indistinguishable robots: "I am entering the third room ...[4] I can find the third robot, but I guess this is the same one as the first one but the room is different"] Another belief formed about the object itself is the answer to the question: "Does the object appear to have a particular purpose or motivation? and if so what is it?" [participant 10 following a tour guide "There are a total of three robots in here now. But... and they seem to be moving randomly."] The direction or trajectory that the object is moving in is important when an agent is trying identify a PIO only a relatively short time after encountering another PIO [participant 18 following a robot "He hasn't changed directions, so I can still tell which one is him"] It is also important for the agent to have some awareness of where other PIOs are in the area to make sure that it doesn't get unnecessarily confused if the object it is focusing on moves too close to one of the others. [participant 23 following a robot "So I just cut in front of that robot, in order to keep following mine."] Successful participants like participant 23 would often keep some awareness of nearby PIOs and act to avoid occlusion of their focus object by other PIOs.

[4] When sequences of two or more dots appear inside of a quote from a participant, it indicates that the participant gave a noticeable pause at that point. The number of dots indicates the length of the pause.

An algorithm

In this section we will give an English gloss of our PIO identification algorithm which is omitted here to save space. For the complete algorithm see (Santore, 2004).

To identify an object *O* with description *D*, just encountered by agent *A*, *A* should first decide if it has ever seen anything that looks like D before. If not, then *O* is a newly encountered object. If *A* has seen something that looks like *D*, then *A* then checks the base and intermediate cases of identifying PIOs in an interleaved order. If there is only one thing with description *D* visible to *A* now, and *A* believes that there is only one thing e_1 that looks like *D* or at least that there is only one thing here that looks like *D*, and *A* has only seen one such object before, *A* assumes that it has encountered e_1 again. Otherwise, if objects that look like *D* are immobile or stationary and without a mover, *A* must decide if it has seen an e_1 at the same location before, if so, then the object is e_1, else the object is something new. Otherwise, if *A* believes that it has continuously viewed or has continually perceived an e_1, as it traveled to the place that A now sees *O*, *A* believes that *O* is e_1. If none of these base or intermediate cases hold, then for each location that *A* currently sees an *O* with description *D*, *A* should create a new mental entity e_2. With each e_2, *A* should consider if that e_2 actually refers to the same object in the world as some previously conceived entity e_1.

We make a simplifying assumption at this point: that a moving object will move at a constant speed. When trying to decide if e_2 refers to the same entity as e_1 *A* first considers the rate of movement of each. If the rates are not the same then e_1 and e_2 refer to different objects. If *A* doesn't know the rate of movement, *A* cannot make an informed decision about the identity of e_1 and e_2. Next *A* checks to see if the distance that e_1 could have traveled is less than the shortest path (that *A* knows about) between the place it last saw e_1 and the place *A* sees e_2. If so, then e_1 and e_2 refer to different objects. Next *A* should consider if it believes that the motivations and capabilities that it believes e_1 has would disallow e_1 from being in the place that e_2 is currently being encountered. If so e_1 and e_2 refer to different objects. At this point, A should consider if the possible range of e_1 is larger than an environment specific constant ("it could be almost anywhere by now"). If so, then *A* cannot decide with certainty if e_1 and e_2 refer to the same object. Otherwise *A* should decide if e_1 and e_2 are coreferential given that only a short distance could be traveled.

When there is only a short possible distance to travel, the agent can now make the closed world assumption.

Since an object moves at a constant speed, any other object that might be mistaken for the object being identified will also move at the same speed, and thus be restricted to a small travel distance of its own. So *A* now checks to see if it knows of any other PIOs except for e_1 and e_2 that could reach the place where e_2 is now. If not, then it assumes that e_1 and e_2 are coreferential. Otherwise, if e_1 was headed toward the place e_2 is now seen but no other PIOs were, then *A* assumes that e_1 and e_2 are coreferential. If there is another PIO headed in the same direction, *A* can't be sure if e_1 and e_2 are coreferential. If none of the above cases hold, the agent will assume that e_1 and e_2 are not coreferential.

Conclusions and Future Work.

We are implementing this theory in a simulated embodied robotic agent. Most of the base and intermediate cases are currently implemented as well as most of the general algorithm. We still need to formalize and implement some of the the support functions that we have assumed, such as deciding if an agent's motivations disallow it from being at a spot at a given time.

This paper has described a human-based computational system for the perception-based task of identifying an object which is perceptually indistinguishable from one seen before. The theory is built using the strategies that were found by doing experiments with human participants who did the same task. As a theory based on human performance, it is both cognitively plausible and designed to produce the same successes and failures as humans performing the same task.

References

Brachman, R. J. and Levesque, H. J., editors (1985). *Readings in Knowledge Representation*. Morgan Kaufmann, San Mateo, CA.

Coradeschi, S. and Saffioti, A. (2003). An introduction to the anchoring problem. *Robotics and Autonomous Systems*, 43(2-3):85–96.

Ericsson, K. A. and Simon, H. A. (1984). *Protocal Analysis*. MIT Press, Cambridge, MA.

Feigl, H. and Sellars, W., editors (1949). *Readings in Philosophical Analysis*. Appleton-Century-Crofts, New York.

Feigle, H. and Sellers, W., editors (1949). *Readings in Philisophical Analysis*. Appleton Centuray Crofts, New York.

Frege, G. (1892). *On Sense and Nominatum*, pages 85–102. Reprinted in (Feigle and Sellers, 1949).

Jain, R., Kasturi, R., and Schunck, B. (1995). *Machine Vision*. McGraw-Hill, New York.

Johnson, S. P. (1998). *Object perception and object knowledge in young infants: a view from studies of visual development*, pages 211–239. Printed in (Slater, 1998).

Kuipers, B. and Beeson, P. (2002). Bootstrap learning for place recognition. In *Proceedings of the Eighteenth National Conference on on Artificial Intelligence (AAAI-02)*, pages 174–180, San Francisco, CA. Morgan Kaufmann.

Kuipers, B. and Byun, Y.-T. (1991). A robot exploration and mapping strategy based on a semantic hierarchy of spatial representations. *Journal of Robotics and Autonomous Systems*, 8:47–63,.

Lehmann, F., editor (1992). *Semantic Networks in Artificial Intelligence*. Pergamon Press, Oxford.

Maida, A. S. and Shapiro, S. C. (1982). Intensional concepts in propositional semantic networks. *Cognitive Science*, 6(4):291–330. Reprinted in (Brachman and Levesque, 1985, pp. 170–189).

Pollock, J. (1974). *Knowledge and Justification*. Princeton University Press, Princeton.

Russell, B. (1905). On denoting. *Mind*, 14(56):479–493. Reprinted in (Feigl and Sellars, 1949).

Santore, J. F. (2004). *Identifying Perceptually Indistinguishable Objects*. PhD thesis, Universit at Buffalo: The State University of New York, Buffalo, NY. Forthcoming.

Santore, J. F., Segal, E., and Shapiro, S. C. (2004). Human identification of perceptually indistinguishable objects. In preparation.

Shapiro, S. C. (1998). Embodied Cassie. In *Cognitive Robotics: Papers from the 1998 AAAI Fall Symposium, Technical Report FS-98-02*, pages 136–143. AAAI Press, Menlo Park, CA.

Shapiro, S. C. and Ismail, H. O. (2003). Symbol anchoring in a grounded layered architecture with integrated reasoning. *Robotics and Autonomous Systems*, 2–3:97–108.

Shapiro, S. C. and Rapaport, W. J. (1992). The SNePS family. *Computers and Mathematics with Applications*, 23(2-5):243–275. Reprinted in (Lehmann, 1992, pp. 243–275).

Shapiro, S. C. and the SNePS Implementation Group (2002). *SNePS 2.6 User's Manual*. Department of Computer Science and Engineering, University at Buffalo, The State University of New York, Buffalo NY.

Slater, A., editor (1998). *Preceptual Development: Visual, auditory, and speech peception in infancy*. Psychology Press, East Sussex, UK.

Trafton, J. G., Shultz, A. C., Perzanowski, D., Bugajska, M. D., Adams, W., Cassimatis, N. L., and Brock, D. P. (2004). Children and robots learning to play hide and seek. *Cognitive Systems Journal*.