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 plausible computational theory of identifying perceptually indistinguishable objects (PIOs) based on a set of experiments 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, and is part of the solution to the symbol anchoring problem. We discuss several base cases in the identification process, some related intermediate cases and the knowledge that is needed for the general case. 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)?

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. This task is required in both the *Track* and *Reaquire* functionalities that Coradeschi and Saffiotti (2003) define as part of a solution to the symbol anchoring problem. Symbol anchoring is defined as "the process of creating and maintaining the correspondence between symbols and sensor data that refer to the same physical objects." The *Track* functionality assumes that the object being tracked is continuously observed by the agent. This continual observation will certainly make the task of identifying the object and anchoring its description to the appropriate mental symbol easier. The *Reaquire* functionality in particular, however, needs a complete theory of identifying PIOs. *Reaquire* is the case "in which an object is reobserved after some time" (Coradeschi & Saffioti 2003, p 91). In order to know that the same object has been reobserved after some time (rather than a new object that looks just like the old one) the *Reaquire* functionality requires a mechanism for identifying PIOs.

Identifying a PIO is a sub-problem of object *identification* rather than object *recognition*. Object recognition is defined in a computational vision textbook (Jain, Kasturi, & Schunck 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 theory of how agents, particularly artificial embodied agents (such as robots) can use reasoning to identify PIOs as well as humans do. Let us first 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 commonsense 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 & 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 & Rapaport 1992; Shapiro & 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). Mental entities are the denotations of the symbols described by Coradeschi and Saffiotti (2003) as part of the symbol anchoring process.

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 & 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 discuss four base cases in the identification of PIOs, and then introduce the knowledge used in more complex cases. The base cases and the more general knowledge are all based on evidence drawn from protocol analysis-style experiments done with human subjects to see how humans identify PIOs[2]. We will then give 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.

The base cases in the identification of perceptually indistinguishable objects

What makes a base case.

Experiments with human subjects showed that there are four conditions under which human subjects find the identification of perceptually indistinguishable objects to be very easy. We'll call these four conditions the base cases of the identification task. Subjects 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 experiment.

When the computational agent identifies a perceptually indistinguishable object using a base case, it does not form a new mental entity for the object and then try to find an existing entity with an equivalent extension. The agent only creates new entities as needed for

[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.

[2]We are currently preparing a paper describing the results of these experiments.

cognizing information (Maida & Shapiro 1982). The object that the agent is perceiving is either *the one* that it has seen before, or a new object. If the object is *the one*, then the agent ought to use the original mental entity for it and not conceive of something new which the agent then believes is really the same thing in the world. If the object is a new one, a new mental entity is created for the new object in the world that our agent conceives of.

Each of the four base cases is described in its own subsection below.

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.[3] Unlike some of the base case strategies, subjects were conscious that they were using this strategy of simultaneous perceptions while they used it. While counting moving robots, Subject 37 states "I'm trying to see them simultaneously." Subject 4, while doing the same task, is even more explicit when stating "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 distinct PIOs in the world, then the agent can instantly anchor the object to its mental entity as soon as the agent perceives an object with that appearance. The agent can and ought to use its original entity for the object in this case.

Subjects were often aware enough of their use of this assumption of unique appearances to try to verify the assumption when possible. Subject 15, 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 use of a single mental entity for an object believed to have a unique appearance was particularly noticeable when the subject'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 Subject 42 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 subject "He came up that way. How did he come up that way?" The subject clearly seems to be using the same mental entity for both robots and believes that there is only one robot.

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

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 which 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 seeing. Then 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 anchors a new entity to the object description.

Experiments with human subjects support the claim that location is of paramount importance in identifying immobile PIOs. Human subjects find the use of location information so intuitive that they rarely notice it at the conscious level. When human subjects 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, subjects 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 subjects were asked to count glasses; The glasses were immobile in this environment and recognized as such by subjects.

> **Experimenter:** how were you able to distinguish between the glasses even when they looked the same?
>
> **Subject 33:** ah because they are lying in the different rooms. That's why. They are different.

An agent cannot use only an object's location to identify that object as the appropriate mental entity. The agent must still recognize the object as belonging to the same class of (perceptually indistinguishable) objects as the previously encountered object. There are two reasons for this. The first reason is that any object might be destroyed: a house might burn down, a tree might blow down in a hurricane. Since an object might be destroyed, and some other object take its place (perhaps a gazebo in the place of an ancient tree) the object itself needs to be identified as being perceptually indistinguishable from the previously seen object.

The second reason not to rely on location alone is that location is only sufficient to anchor immobile objects. So an agent must recognize the object as being of a class of objects that are immobile in order to take advantage of location as the distinguishing factor for anchoring an object.

The use of the original entity seems to be supported by the human subject data in the immobile object case as well. While performing the glass-counting task, no subject 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 subjects often clearly seem to have more than one entity for an object and can talk about both entities. For example when following a robotic tour guide in a suite with several perceptually indistinguishable distractors, Subject 30 briefly loses the tour guide robot and then makes the following statement "Where did the robot go? I think this is the one" In contrast, Subject 55 makes the following typical statement while counting immobile objects: "ahhh ok, it seems to me I've already seen this room and counted this one glass here."

Earlier it was noted that the use of a single entity is contingent on an agent correctly identifying its current location. Our subjects were vulnerable to mistaking one room for another if the two looked similar. Kuipers and his colleagues (Kuipers & Byun 1991; Kuipers & Beeson 2002) call this sort of mistake "perceptual aliasing" and have discussed the problem and a solution for robotic agents. When our subjects fell victim to perceptual aliasing, use of location information to identify immobile objects was fallible. Sometimes subjects would notice the possible aliasing such as Subject 20 while counting glasses who 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." Subjects who fell victim to perceptual aliasing and never realized it, generally failed at the counting and identification tasks.

Base Case 4: Continuous viewing.

Pollock (1974) has discussed reidentification of objects, a subproblem of identifying PIOs. 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 subjects 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 (Coradeschi & Saffioti 2003) *Track* functionality.

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 subjects 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. Subject 7, 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." Subject 23, 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[4] that the base cases described in section 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. In looking at our experimental data we found evidence of non-perceptual cases that are similar to the base cases. Each perceptual base case, had one non-perceptual simple case which can be closely identified with the base case. We 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.

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 Subject 18 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.

Prior to this excerpt, Subject 18 has seen only one robot, a silver-gray robot. As he enters another room, Subject 18 sees a "multi-colored" robot as well as a silver-gray

[4] Our thanks to the anonymous reviewer who did so.

robot. In order to identify this silver-gray robot as a new, never before seen robot, Subject 18 looks back toward the place where he last saw a silver-gray robot. When he sees a silver-gray robot in the previous location as well, Subject 18 assumes (correctly in this case) that the robot seen in the current room is different than 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. If all of these conditions hold, the agent can determine with a very high degree of 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. The agent must believe that an object is unique in the current context. For example if 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 you see the second twin. Of course the simultaneous perceptions base case described above will trump a belief that an object has a unique appearance.

Subjects seemed to use this assumption that unique objects can be effortlessly identified as the same mental entity when they could. Sometimes the assumption of uniqueness of appearance would be limited to a single room. Subject 12, while following a robotic tour guide in a suite of rooms with PIOs as distractors says “I’m stuck, ok but there is only one robot so I can follow it.” Subject 23, doing the same task, says something similar “There aren’t any other robots in this room so it’s a little easier to follow.”

Intermediate Case 3: Stationary Objects.

The next intermediate case is related to the base case of immobile objects. Stationary objects are those objects which 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 can not 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. subject 31 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?
>
> **Subject 31:** Mmm I guess I just kind of based it on the fact that they would be stationery 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: Continuously ‘Perceived’ Objects

It is well known (Johnson 1998) that young children will identify object that are briefly occluded as the original objects. Our subjects overwhelmingly did likewise. Though subjects may have briefly lost sight of the focus object by looking away or having the object occluded, the subjects 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, subjects were not even aware that they had lost sight of the object in question.

Identifying PIOs in general.

While identifying PIOs is trivial and intuitive when one of the base cases can be applied, and only a little more difficult in one of the intermediate cases, when one of the base cases does not hold, the task can be much harder. An agent usually requires several more pieces of knowledge to identify mobile objects which are not continuously viewed. 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 subjects 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? [Subject 8 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? [Subject 6 asked in a retrospective why subject 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? [Subject 18 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 subjects 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 subject believes might be the PIO that the subject is currently looking at. [Subject 25 counting robots with two distinct groups of perceptually indistinguishable robots: "I am entering the third room ...[5] 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?" [Subject 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 [Subject 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. [Subject 23 following a robot "So I just cut in front of that robot, in order to keep following mine."] Successful subjects like subject 23 would often keep some awareness of nearby PIOs and act to avoid occlusion of their focus object by other PIOs.

An algorithm

Assumptions

Let us suppose that our agent has seen an object with appearance d at location p at time t. The agent needs to identify the object. Either the object is a new one, or it is one that the agent has seen before, in which case the agent already has an entity representing the object. When the agent already has one or more entities representing objects with description d, the agent must recognize which entity, if any, represents the same PIO using reasoning.

We assume for simplicity that the object's speed, if known, will be constant. If the agent doesn't know the speed of the object, it will probably not be able to decide if the object it is perceiving has been previously encountered or not.

Helper Functions

The PIO identification functions assume that several support functions are available.

Believe-equiv(E_1, E_2) **effect:** *creates a new belief that E_1 and E_2 refer to the same object in the world.*

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

NotAutoMobile(Class) **returns:** true *if the agent believes that members of Class cannot move on their own.*

ClassOf(Desc) **returns:** *the class of objects that "look like" Desc.*

ContinuouslyPerceivedPIO(Desc, ESet, P) **returns:** *an entity from the set of entities ESet. The returned entity corresponds to the object with the description Desc that the agent has perceived as being the same from some previous place to the place* P. *This function is similar to* ContinuouslyViewedPIO *(see below) however, there is no requirement that the object be viewed continuously, rather, if sight of the object is lost for a few seconds, the tracking mechanism will continue to track where it 'ought' to be and continue to accept the object perceived there as the one being tracked. If there is no such entity then the function returns* null, *the non-entity.*

ContinuouslyViewedPIO(Desc, ESet, P) **returns:** *an entity from the set of entities ESet. The returned entity corresponds to the object with the description Desc that the agent believes it has viewed continuously from some previous known place, to the place P. If there is no such entity then the function returns* null, *the non-entity. This function will rely on a FINST-like* (Pylyshyn 1989) *mechanism for tracking continuously viewed objects*. Sandewall (2002) *has implemented such a system using* tracker objects *and* activity demons *to supervise the tracker objects.* ContinuouslyViewedPIO *will return true if there is a tracker object that has continuously viewed the object at* P *from some previous spot.*

CouldReach(E, P, T) **returns:** true *if the agent believes that the object corresponding to entity E could have arrived at the place P by time T.*

Disallow(MSet, P_1,P_2,T_1,T_2) **returns:** true *if the agent believes that any element of the set of motivations, MSet, of an entity would disallow the entity from being at location P_2 and time T_2, given the agent's belief of a previous encounter at time T_1 and location P_1; otherwise the function returns* false.

HeadedToward(E, P) **returns:** true *if the object corresponding to entity E was headed in the direction of the place* P *when the agent last observed the object; otherwise returns* false.

Immobile(Class) **returns:** *true if the agent believes that members of Class are immobile and* *false otherwise.*

IsUnique(Desc) **returns:** true *if the agent believes that there is only one object in the world that has the appearance Desc; otherwise the function returns* false.

IsUniqueInContext(Desc) **returns:** true *if the agent believes that there is only one object in the current context that has the appearance Desc; otherwise the function returns* false.

Loc(E) **returns:** *the location where the entity E was last encountered.*

MakeEntity(Desc, Loc) ***returns:*** *a new mental entity to represent the object with the description Desc, which is believed to be at location Loc.*

MotivationsOf(E) **returns:** *a set containing any motivations that the agent believes the entity E has.*

OtherKnownPIOs(E_1, E_2) **returns:** true *if the agent knows about other entities that are perceptually indistinguishable from the two entities E_1 E_2 and which are not coreferential with either; otherwise returns* false.

OtherPossMover(Class) **returns:** *a possibly empty set of all of the entities that the agent knows are in the area that could have moved something of the class* Class.

PIO-AtLoc(Class, ESet, Loc) **returns:** *If an object*[6] *of Class has been seen at location Loc before, then the entity (from the set of entities ESet) denoting that object is returned; otherwise null (the non-entity) is returned. Note: The agent may still fall victim to perceptual aliasing (Kuipers & Byun 1991; Kuipers & Beeson 2002), and so if the agent is mistaken in its beliefs about the location, then the agent is likely to reach the wrong conclusion about the identity of the object.*

RateOf(E) **returns:** *the speed or rate of movement of the entity E*

RemoveEquivs(E, ESet) **effect:** *removes from the set of entities ESet, all entities that are coreferential with entity E.*

ShortestKnownPathBetween(P_1, P_2) **returns:** *the length of the shortest route (that the agent knows about) between the two positions P_1 and P_2. This function is done at a subcognitive level. The path is calculated by using a simple path planning algorithm like the one proposed in (Lozano-Perez & Wesley 1979). Once the path is available, the length of the path is calculated from the lengths of the path's straight-line segments.*

Time(E) **returns:** *the time when the entity E was last encountered.*

[6] *Only one object of a particular class can be at a single location at any one time. One might think of a cabinet and the fancy china plate inside it as being at the same location, however you cannot have two plates at the same location.*

PIO Identification Functions

Below is an algorithm based on human subjects experiments for recognizing if a currently perceived object is new or is the same object as a PIO that was seen earlier. The algorithm is given below as four functions written in pseudo-code. The four functions presented are Recognize, IdentifyNonMovingObjects, IsSame and IsSameShortD.

The function Recognize takes an object description, a set of positions of all of the current sightings of objects with that description, the time of sighting, and a (possibly empty) set of entities representing objects with this description that have already been encountered. The function returns a set containing mental entities corresponding to each object currently seen. Due to the simultaneous perception base case, there will be one entity returned for every element in the set of places.

```
function Recognize(D, P, T, E)
 returns a set of entities
 inputs:
 D: description of an object in the world,
 P: set of places that agent currently
  perceives objects that have description D,
 T: the time of the perception, and
 E: set of previously created entities that
   have description D
eSet ←{}
First check the case of no PIOs - first time the
agent sees something with description D
  if |E| = 0{
    for each (P_i ∈P)
      eSet ← eSet + MakeEntity(D, P_i)
    return eSet}
Next check the base case of unique objects and
the intermediate case of those unique in the
current context.
  if (|P| = 1 & |E|=1) &
  (IsUnique(D) ⋁ IsUniqueInContext(D))
    return E
Now check the base and intermediate cases
arising from objects that are not moving.
  if NotAutoMobile(ClassOf(D))
    return IdentifyNonMovingObjects(D, P, E)
Next check the base case of continuously
viewed objects
  for each (P_i ∈P) {
    E_n ←ContinuouslyViewedPIO(D,E, P_i)
if not the base case, consider the intermediate
case of continuously perceived objects
```

```
    if(En = null)
      En ←ContinuouslyPerceivedPIO(D,E, Pi)
    if En != null{
      eSet ←eSet + En
      P ←P - Pi}}
```
For each remaining non-base case object, make a new entity and reason if that entity is the same as something seen before.
```
  for each (Pi ∈P)
    En ←MakeEntity(D, Pi)
    for each(Em ∈E){
     if IsSame(Em,En,Loc(Em),Pi,Time(Em),T)
         {Believe-equiv(Em, En)
         RemoveEquivs(En, E) }
      eSet ←eSet + En}
  return eSet
```

function IdentifyNonMovingObjects(D, P, E)
returns: a set of entities
inputs:
D: description of an object in the world,
P: set of places that agent currently perceives objects that have description D,
E: set of previously created entities that have description D
In this function, the agent considers those objects which cannot move themselves.
```
  if Immobile(ClassOf(D)) ⋁
  ∄ OtherPossMover(ClassOf(D)){
    for each (Pi ∈P){
      En ←PIO-AtLoc(ClassOf(D),E,Pi)
      if(En !=null)
        eSet ←eSet + En
      else
        eSet ←eSet + MakeEntity(D, Pi)}
    return eSet}
```

The function IsSame checks to see if two entities represent the same object. Three return values are possible. A return value of *true* means that the agent believes that $e_1 = e_2$. A return value of *false* means that the agent believes that $e_1 \neq e_2$. A return value of *unknown* means that the agent does not believe it has enough information to decide whether $e_1 = e_2$. If the agent must make a decision with an answer of unknown, it will probably have to select randomly.

function IsSame(E1, E2, P1, P2, T1, T2)
returns *three-value-boolean:*
inputs:
E1, E2*:Two entities,*
P1, P2*:the place where each was perceived,*
T1, T2 *:the time of each perception.*
```
  rate1 ← RateOf(E1)
  rate2 ← RateOf(E2)
```
Since the assumption is that an object's speed is constant, if the rates of speeds differ, then the objects differ.
```
  if (rate1 != rate2)
    return false
```
if the agent doesn't know the speed of the objects in this non-base case situation, the agent cannot make a good decision about the identity of the objects.
```
  if rate1 = unknown
    return unknown
```
next check to see if an object with the known speed could have traveled from its previously known position to the currently perceived position. If not, the two must be different objects.
```
  possibleRange ← rate1 * (T2- T1)
  if (possibleRange <
  ShortestKnownPathBetween(P1, P2)
    return false
```
If the agent knows a motivation of E1 which would prevent E1 from being at the place where E2 is currently perceived, then assume the two are different.
```
  if Disallow(MotivationsOf(E1),P1,P2,T1,T2)
    return false
```
If the distance that E1 could have traveled between sightings is greater than some environment-specific large distance then the agent can't decide
```
  if (possibleRange > LargeD)
    return unknown
    else
```
Otherwise use a rule for identifying objects that might have only traveled a small distance.
```
    return IsSameSmallD(E1,E2,P2,T2)
```

function IsSameSmallD(E1, E2, P2, T2)
returns *three-value-boolean:*
inputs:
E1, E2*: Two entities,*
P2*:the place where the second was perceived,*
T2*: the time of the second perception.*
If the agent knows about PIOs other than the two entities in question then if none of the others could reach the place of the second perception, P2, by the time T2 then assume that E1 = E2
```
  if OtherKnownPIOs(E1, E2)
    if ∄(other-known-PIO) suchthat
    CouldReach(other-known-PIO, P2, T2)
      return true
```
Otherwise if the object represented by the first entity was headed toward the place of the second perception and there are no other PIOs known to be headed in that location E1 = E2
```
  if HeadedToward(E1, P2)
    if ∄(other-known-PIO) suchthat
```

HeadedToward(other-known-PIO, P_2)
return true
if there was another PIO headed in the same location, then the agent isn't sure.
else return unknown
If none of the above, there is no evidence that the two entities represent the same object and there has only been enough time to travel a short distance so assume that the two are not the same.
return false}

Conclusions and Future Work.

We are currently implementing this theory in a simulated embodied robotic agent. The base case of identifying immobile objects is currently implemented and we are working on the other base cases and expect to begin implementing the intermediate cases and the general algorithm soon. We still need to formalize and implement some of the the support functions that we have assumed, such as Disallow.

This paper has described a computational theory and algorithm of identifying an object which is perceptually indistinguishable from one seen before as either being the same as the one before, or as being different. The theory is built using the results of experiments with human subjects who did the same task. It is 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., eds. 1985. *Readings in Knowledge Representation*. San Mateo, CA: Morgan Kaufmann.

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

Feigl, H., and Sellars, W., eds. 1949. *Readings in Philosophical Analysis*. New York: Appleton-Century-Crofts.

Feigle, H., and Sellers, W., eds. 1949. *Readings in Philisophical Analysis*. New York: Appleton Centuray Crofts.

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

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

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

Kuipers, B., and Beeson, P. 2002. Bootstrap learning for place recognition. In *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence (IJCAI-02)*, 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., ed. 1992. *Semantic Networks in Artificial Intelligence*. Oxford: Pergamon Press.

Lozano-Perez, T., and Wesley, M. A. 1979. An algorithm for planning collision-free paths among polyhedral objects. *Communications of the ACM* 22(10):560–570.

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

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

Pylyshyn, Z. 1989. The role of location indexes in spatial perception: A sketch of the finst spatial-index model. *Cognition* 32(1):65–97.

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

Sandewall, E. 2002. Use of cognitive robotics logic in a double helix architecture for autonomous systems. In Beetz, M.; Hertzberg, J.; Ghallab, M.; and Pollack, M. E., eds., *Advances in Plan-Based Control of Robotic Agents*, Lecture Notes in Computer Science, 226–248. New York: Springer Verlag.

Shapiro, S. C., and Ismail, H. O. 2003. Symbol anchoring in a grounded layered architecture with integrated reasoning. *Robotics and Autonomous Systems*.

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.

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

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