Recovery Planning for Ambiguous Cases in Perceptual Anchoring

Mathias Broxvall and **Silvia Coradeschi** and **Lars Karlsson** and **Alessandro Saffiotti**

{mathias.broxvall,silvia.coradeschi,lars.karlsson,alessandro.saffiotti}@tech.oru.se

Center for Applied Autonomous Sensor Systems

Dept. of Technology, Örebro University

SE-701 82 Örebro, Sweden

http://www.aass.oru.se

Abstract

An autonomous robot using symbolic reasoning, sensing and acting in a real environment needs the ability to create and maintain the connection between symbols representing objects in the world and the corresponding perceptual representations given by its sensors. This connection has been named *perceptual anchoring*. In complex environments, anchoring is not always easy to establish: the situation may often be ambiguous as to which percept actually corresponds to a given symbol. In this paper, we extend perceptual anchoring to deal robustly with ambiguous situations by providing general methods for detecting them and recovering from them. We consider different kinds of ambiguous situations and present planning-based methods to recover from them. We illustrate our approach by showing experiments involving a mobile robot equipped with a color camera and an electronic nose.

Introduction

Autonomous systems embedded in the physical world typically incorporate two different types of processes: high-level cognitive processes, that perform abstract reasoning and generate plans for actions; and sensory-motoric processes, that observe the physical world and act on it. These processes have different ways to refer to physical objects in the environment. Cognitive processes typically (although not necessarily) use symbols to denote objects, like 'b1'. Sensory-motoric processes typically operate from sensor data that originate from observing these objects, like a region in a segmented image. If the overall system has to successfully perform its tasks, it needs to make sure that these processes "talk about" the same physical objects. We call this the anchoring problem (Coradeschi & Saffiotti 2003).

An important challenge in anchoring is to resolve situations where the sensors detect several objects that are consistent with the symbolic description of a desired object. In order for an autonomous system to function robustly when encountered with such ambiguous situations, it needs to reason and act in a way that allows it to distinguish between the perceived object and determine which one is the correct one. In (Broxvall, Karlsson, & Saffiotti 2004) a simple case of ambiguity in anchoring due to the accumulation of

uncertainty has been investigated. In this paper, we take a more general approach and analyze different types of ambiguity that can make the anchoring process fail. We then propose a general approach to automatically detect and isolate these failures, and to automatically generate a conditional plan to recover from the failure when possible. As an additional contribution, we extend the anchoring framework to handle symbolic descriptions including relations among objects, like "the red ball near the yellow can". Relations are an important aspect in the description and recognition of objects and they contribute to enrich the anchoring process by making more complex anchoring cases possible.

The motivation for using recovery planning is the possibility to handle complex situations in an optimal way in terms of, e.g., the number of observation and movement actions required. Due to the combinatorial nature of the problem with different types of failures, involving different numbers of objects characterized by different properties and (possibly nested) relations, the number of recovery situations and plans grows exponentially and it becomes infeasible to use a hand-coded approach.

There are a number of systems that use planning-based techniques to perform recovery. However, most systems that take this approach (e.g., (Fikes, Hart, & Nilsson 1972; Beetz & McDermott 1997; Pell *et al.* 1998; Seabra-Lopes 1999)) focus on the external state of the world, looking for discrepancies between the observed state and the expected one. In comparison, our work focus on the inability to acquire the perceptual information needed to make a certain decision, in particular about how to anchor a specific symbol. The problem of perception is still unsolved in the general case and unreliable in most practical situations. By performing recovery at a higher cognitive level, we can increase the robustness of the system in face of imperfect perception, and handle cases which are inherently ambiguous even with a perfect perception system.

Several works have addressed the problem of planning for perceptual actions. Perception planning has been studied as a means for gathering better visual information (Kovacic, Leonardis, & Pernus 1998), for achieving safer landmark-based navigation (Lazanas & Latombe 1995), for performing tasks that involve sensing actions (Giacomo *et al.* 1997), and for generating image processing routines (Beetz *et al.* 1998). None of these works, however, deal with the prob-

lem of recovery.

In the next section, we give a brief reminder of perceptual anchoring. In section 3 we analyze different types of ambiguity, and explain how the ambiguity can be detected and dealt with. In section 4 we show how the ambiguous situation can be modeled in a planner and a recovery plan generated automatically for those cases that can be resolved. Finally, we demonstrate our technique by presenting a series of experiments run on a mobile robot.

Perceptual Anchoring

Anchoring is the process of creating and maintaining the correspondence between symbols and sensor data that refer to the same physical objects. In our work, we use the computational framework for anchoring defined in (Coradeschi & Saffiotti 2000). In that framework, the symbol-data correspondence for a specific object is represented by a data structure called an **anchor**. An anchor includes pointers to the symbol and sensor data being connected, together with a set of properties useful to re-identify the object, e.g., its color and position. These properties can also be used as input to the control routines.

Consider for concreteness a mobile robot equipped with a vision system and a symbolic planner. Suppose that the planner has generated the action 'GoNear(b1)', where the symbol 'b1' denotes an object described in the planner as 'a green garbage can'. The 'GoNear' operator is implemented by a sensori-motor loop that controls the robot using the position parameters extracted from a region in the camera image. In order to execute the 'GoNear(b1)' action, the robot must make sure that the region used in the control loop is exactly the one generated by observing the object that the planner calls 'b1'. Thus, the robot uses a functionality called **Find** to link the symbol 'b1' to a region in the image that matches the description 'a tall green gas bottle'. The output of Find is an anchor that contains, among other properties, the current position of the gas bottle. While the robot is moving, a functionality called **Track** is used to update this position using new perceptual data. Should the gas bottle go out of view for some time the **Reacquire** functionality would be called to update the anchor as soon as the gas bottle is in view again. More details on perceptual anchoring can be found in (Coradeschi & Saffiotti 2000; 2001; 2003).

Matching

A central ingredient in all the anchoring functionalities is the *matching* between the symbolic description given by the planner and the properties of percepts generated by the sensor system. This is needed to decide which percepts to use to create or update the anchor for a given symbol. Matchings between a symbolic description and a percept can be *partial* or *complete* (Coradeschi & Saffiotti 2001).

Definition Given a percept π and a description σ, we say that π *fully matches* σ if each property in π matches a property in σ and vice-versa. π *partially matches* σ if each property in π matches a property in σ, but some property in σ is not observed in π. Otherwise π *does not match* σ.

For example, consider the description "a gas bottle with a yellow mark". A gas bottle in an image where no mark is visible provides a partial match, since the mark might not be visible from the current viewpoint. Another example of partial match is the case when several cups are identical from the point of view of vision, but they can be distinguished by using the sense of smell. A percept is said to be a complete anchoring candidate for a symbol if it fully matches the symbolic description of the symbol, and a partial anchoring candidate if it partially matches the description.

Definite and indefinite descriptions

The symbolic descriptions used in the anchoring process can be either definite or indefinite. A description is *definite* when it denotes a unique object, for instance "the cup in my office". Linguistically one uses in this case the article "the". An *indefinite* description requires that the object corresponds to the description, but not that it is unique, for instance "a red cup". Definite descriptions are especially challenging when an object is conceptually unique, but its perceptual properties do not characterize it unequivocally, for instance "the cup that I have seen before". This is a common event in the Reacquire functionality when more than one object matches the description of a previously seen object (in Reacquire, descriptions are always definite). An example of this situation is shown later in this paper.

Anchoring with Ambiguities

The matching process described in the previous section provides complete and partial anchoring candidates for a symbol. The anchoring module detects the presence of ambiguity on the basis of the number of complete and partial anchoring candidates, and whether the description involved is definite or indefinite. The following table summarizes the cases that can occur.

	# Matches		Definite		Indefinite	
Case	full	partial	result	action	result	action
1	0	0	Fail	Search	Fail	Search
2	0	1+	Fail	Observe	Fail	Observe
3	1	0	Ok	—	Ok	—
4	1	1+	Ok/Fail	-/Observe	Ok	—
5	2+	any	Conflict	—	Ok	—

In cases 1 and 2 no anchoring candidates have been found fully matching the symbolic description. In case 1 the recovery module can try to recover by making a search. In case 2 temporary anchors are created for each of the partial candidates and these anchors are returned to the recovery module. A recovery plan is constructed and executed aiming at observing the missing properties of the object. If the situation is successfully disambiguated, the planner informs the anchoring module about which of the candidate perceived objects should be used for anchoring.

Case 3 is the ideal case where just one complete candidate is present. The anchoring module selects that percept.

In cases 4 and 5 at least one complete candidate for the symbol is present. If the symbolic description is indefinite any one of these complete candidates can be selected for

anchoring the symbol. If the description is definite the presence of several candidates can constitute a problem. In case 4 where a complete candidate and partial ones are present a cautious approach consists of constructing and executing a recovery plan aiming at observing the missing properties of the partial candidates in order to rule them out. However the complete candidate could also be selected, this is why we have ok/fail as result. Case 5 constitutes a serious problem: as the matchings are full, the situation cannot be resolved by getting more perceptual information. Instead, the description has to be considered insufficient, and needs to be made more precise (how to do that is not addressed in this paper).

Finally, we should point to some particularly difficult situations: when important characteristics of the object have changed in an unpredictable way (e.g., the shape has been deformed); and when our percepts are not just uncertain but wrong (e.g., a reflection is seen as a mark). In such cases, we might get mismatches that should have been matches, and vice versa, which leads to an erroneous estimate of the situation and hence does not allow a correct recovery.

Relational properties and ambiguity

An interesting challenge in the treatment of ambiguity is represented by situations where an object can be described not only by its properties, like color and shape, but also by its relations to other objects. By considering relations, we may be able to resolve cases where the known properties of the object are not sufficient to distinguish it from other similar objects. An example is "the green garbage can that is near the red ball and the blue box". We consider the object that needs to be anchored, in the example "the green can", as the *primary object* and the other objects related to it, in the example "the red ball" and "the blue box", *secondary objects*. In this paper we consider in particular binary relations and we allow for descriptions with several relations.

The anchoring process handles relational cases by considering the relation as an additional property of the primary object. In the previous example, being "near the red ball" is an additional property of the object besides being a garbage can and being green. Clearly a relational property has the additional complexity that an anchor needs to be found also for the secondary object. The anchoring process for the secondary object is the same as the one for the primary object: the secondary object can be described as definite or indefinite and it can have complete, partial or no anchoring candidates.

In practice we first consider all possible candidates for the primary object based on the non-relational properties in its description. Then for each of these candidates we try to find anchors for all secondary objects on the basis of their descriptions and their relations to the primary object. A *relational anchoring candidate* is represented by a list $(\pi_0, (\pi_{1,1}\ldots), (\pi_{2,1}\ldots), \ldots)$ containing a candidate percept π_0 for the primary object and for each secondary object, a (possibly empty) list of all candidate percepts satisfying the expected relation.

We say that a relational anchoring candidate is *completely matching* if, by applying the table of cases above, the result of anchoring is "OK" for the primary object candidate (viewed as a single-element list) and for each secondary object candidate list. For instance in the example above if the garbage can, the red ball, and the blue box are seen and the red ball and blue box are near the garbage can, then this constitutes a complete matching. A relational anchoring candidate is *partially matching* if the result of anchoring is "fail" for at least one of the candidate lists, but for no object the result is "conflict". A "fail" indicates that an object is not seen or some of its properties are not perceived, for instance a mark is not visible. The planner can in this case select searching and observation actions. Finally, the presence of a "conflict" indicates that one of the objects was described as unique, but more than one object was found corresponding to the description. For instance, given that the object to anchor was "the green garbage can near **the** red ball" a "conflict" would be present if a garbage can and two red balls both near the can were detected. In this case additional observation actions would not help and a more accurate description would be needed.

Once each relational candidate has been classified this way, the entire situation is classified according to the table of cases based on the number of complete and partial relational anchoring candidates. In case of ambiguity the recovery module is invoked, which devises a plan to acquire additional information about primary and/or secondary candidate objects according to the matching results of the objects.

The above concepts are illustrated in the following example. This example will also be used throughout the next section.

Example The robot is presented with the symbolic description "g1 is a garbage can near a red ball with a mark" and given the task to go near g1. To do this, the robot needs to anchor the symbol g1. Consider a situation where a garbage can and a red ball are seen, but no mark is visible. Applying our table of anchoring cases we find that this corresponds to case 3, one fully matching percept (ok), for the primary object and case 2, one or more partial matches (fail), for the secondary object. This implies that the entire relational candidate is a partial match. Applying the case table to our singleton set of relational candidates, we find that we have case 2, a partial match (fail). Thus, to be sure that the observed garbage can is the requested one, the red ball has to be observed further to test if it has any marks visible from other viewpoints.

Notice that we allow cases when the secondary object is not initially found due to, for instance, occlusion. The framework also allows for nested relations, for instance "the red ball near the blue box touching the green can". An additional case we have not considered yet is when the relational property per se is not observable, for instance it cannot be established if two objects are touching from the current view.

Recovery Planning for Anchoring

We propose an approach to actively recover from the recoverable cases above by automatically analyzing and encoding the ambiguous situation as a planning problem and generating a conditional recovery plan. In practise we use a re-

covery module using a conditional possibilistic/probabilistic planner called PTLplan (Karlsson 2001). PTLplan searches in a space of belief states, where a belief state represents the agent's incomplete and uncertain knowledge about the world at some point in time. A belief state can be considered to represent a set of hypotheses about the actual state of the world, for instance that a certain gas bottle has a mark on it or has not a mark on it. Actions can both have causal effects that change properties in the world, and observation effects that lead to a splitting up of a belief state into several new and more informative belief states. The latter leads to conditional branches in the plan.

A recovery situation in anchoring typically occurs when the robot is executing some higher-level plan and encounters one of the ambiguous but recoverable cases above. Such a situation is handled in five steps:

1. The problematic situation is detected and classified as above, and the top-level plan is halted.

2. The recovery module formulates an initial situation (belief state) by considering the properties of the requested object and of the perceived objects, and generating different hypotheses for which of the objects corresponds to the requested object. It also formulates a goal that the requested object should be identified if present.

3. The recovery module calls the planner with the belief state and the goal as input, and a plan is returned.

4. The plan is executed, and either the requested object is found and identified and can be anchored, or it is established that it cannot be identified.

5. If recovery was successful, the top-level plan is resumed.

We now present steps 2, 3 and 4 in more detail.

Formulating the initial situations and goals

When there are one or more partially matching anchoring candidates, the agent needs to figure out which of them actually corresponds to the requested object s. Thus, the recovery module formulates a set of hypotheses that consist of the different ways s can be anchored, based on the known properties of s and its secondary objects and the observed properties of the perceived objects.

1. One starts with a combined description d for the requested object and its related objects. For each anchoring candidate $a_i = (\pi_{i,0}, (\pi_{i,1,1} \ldots), (\pi_{i,2,1} \ldots), \ldots)$, a description d_i of the perceived objects of a_i is computed.

   Example The robot Pippi is looking for g1, described as $d =$ (and (shape g1 = garbage-can) (near g1 b1 = t) (shape b1 = ball) (mark b1 = t)). Two anchoring candidates are found: $a_1 =$ (pi1 (pi2)) and $a_2 =$ (pi3 (pi4)) . Their descriptions are $d_1 =$ (and (shape pi1 = garbage-can) (near pi1 pi2 = t) (shape pi2 = ball) (mark pi2 = t f)) and $d_2 =$ (and (shape pi3 = garbage-can) (near pi3 pi4 = t) (shape pi4 = ball) (mark pi4 = t f)). Note that "= t f" denotes that the particular property can be both true and false, i.e. it's unknown.

2. Next, for each d_i two extra sets of descriptions d_i^+ and d_i^- are computed as follows.

a) Certain properties for certain perceived objects π_{ijk} in the candidate will be unspecified in d_i, implying a partial match. We can constrain all unspecified properties for such a π_{ijk} in terms of a formula d_{ijk}^+, such that $d_i \wedge d_{ijk}^+$ forms a description for π_{ijk} that matches the one in d. By constraining any of the properties not to match in terms of a formula d_{ijk}^-, one can likewise make it a mismatching description.

Example For d_1 we get $d_{1,0}^+ = \top$ and $d_{1,1,1}^+ =$ (mark pi2 = t), and $d_{1,0}^- = \bot$ and $d_{1,1,1}^- =$ (mark pi2 = f), and similarily for d_2.

b) If there are several objects $\pi_{ij1}, \ldots, \pi_{ijk}$ partially matching related object number j in a candidate, one can combine those to get a match by making at least one of them match[1]:

$$d_{i,j}^+ = \bigvee_{1 \le l \le k} d_{ijl}^+.$$

One can make a mismatch by making all of them mismatch:

$$d_{i,j}^- = \bigwedge_{1 \le l \le k} d_{ijl}^-.$$

Example $d_{1,1}^+ = d_{1,1,1}^+$ and $d_{1,1}^- = d_{1,1,1}^-$.

c) If there is no candidate percept for a specific related object, we assume that it has not been seen yet, as discussed below.

d) Each d_i^+ is a description of candidate i that completely matches d, by making the primary perceived object $\pi_{i,0}$ and one perceived object for each related object completely matching:

$$d_i^+ = d_{i,0}^+ \wedge \bigwedge_j d_{i,j}^+.$$

Likewise, each d_i^- contains the different ways in which either the primary perceived object or all candidate perceived objects for one related object can mismatch:

$$d_i^- = d_{i,0}^- \vee \bigvee_j d_{i,j}^-.$$

Example We get $d_1^+ = \top \wedge$ (mark pi2 = t), and $d_1^- = \bot \vee$ (mark pi2 = f).

3. Each hypothesis then consists of a matching description for one candidate and mismatching descriptions for the remaining ones. To each hypothesis is also added the statement (anchor $s = \pi_{i,0}$) denoting that s should be anchored to the object anchored by $\pi_{i,0}$:

$$h_i = d_i^+ \wedge \bigwedge_{j \ne i} d_j^- \wedge (\text{anchor } s = \pi_{i,0})$$

There is also one hypothesis that no object matches: $\bigwedge_j d_j^- \wedge$ (anchor $s =$ f). Finally, if the recovery module takes a cautious approach and wish to ascertain that no more than one object is matching (in particular in case 3),

---

[1] In the following, we divert from the Lisp-style syntax for logical formulae in order to enhance readability.

it might also add hypotheses consisting of $d_i^+ \wedge d_j^+$ for each pair of candidates, and (anchor s = f).

Example We get the following (incautious) set of hypotheses:

h1: (mark pi2 = t) $\wedge$ (mark pi4 = f) $\wedge$ (anchor b1 = pi1)
h2: (mark pi2 = f) $\wedge$ (mark pi4 = t) $\wedge$ (anchor b1 = pi3)
h3: (mark pi2 = f) $\wedge$ (mark pi4 = f) $\wedge$ (anchor b1 = f)

In addition, each of the two first hypotheses can be subdivided further into three different hypotheses regarding from where the mark can be detected: (mark-visible-from pi2 = r1_1) etc. To the above is added information about the topology of the room and other relevant background information.

The goal is achieved once a specific action (anchor *s x*) has been performed. It represents the decision to anchor the symbol *s* to some specific perceived object *x* (or to no object at all, if *x* = f). This action has as a precondition that *x* is the only remaining anchor for *s*: (nec (anchor *s* = *x*)). Thus, all other candidate anchors have to be eliminated before the anchor action is applied.

If no candidate has been found for the primary object (case 1) the recovery module hypothesize that the object is somewhere but is not visible from the current position. Therefore, the initial situation consists of a number of hypotheses concerning from what position the object can be found, each of the form (visible-from *s* = *pos*) where *pos* = f signifies that the object is nowhere around. The planning goal is formulated as (exists (?x) (nec (visible-from *s* = ?x))), which means that the agent has determined from what place the object is visible. A missing secondary object in an anchoring candidate is treated in a similar way, but here the relation involved is also part of the hypothesis.

Generating the recovery plan

After the initial situation and the goal have been established, plan generation starts, using the initial belief state and goal and the set of available actions. The following action, for instance, is for looking for marks (and other visual characteristics) on objects.

```
(ptl-action
 :name (look-at ?y)
 :precond ( ((?p) (robot-at = ?p)) ((?y) (perceived-object ?y)) )
 :results (cond
          ((and (mark ?y = t) (mark-visible-from ?y = ?p))
           (obs (mark! ?y = t)))
          ((not (and (mark ?y = t)
                     (mark-visible-from ?y = ?p)))
           (obs (mark! ?y = f))))
 :execute  ((aiming-at me ?y)(anchor-find ?y)))
```

In short, the precond part states that the action requires a perceived object ?y and a current position ?p. The result part states that if ?y has a mark, and if the robot looks at ?y from the position from which the mark is visible, then the robot will observe the mark (and thus know that there is a mark), and otherwise it will not observe any mark. The obs form is the way to encode that the agent makes a specific observation, and different observation results in different new belief states. In this case, there would be one belief state where the agent knows there is a mark, and one where it knows

Table 1: Experimental results

	Experiments	#Anchors	Success
a	Find 2 odors Find 3 odors Find 4 odors Find 5 odors	11 15 21 25	82% 80% 76.3% 76%
b	Reacquire 2 gas bottles Reacquire 3 gas bottles Reacquire 4 gas bottles	15 10 10	87% 80% 90%
c	Find can near (occluded) ball Find can near ball with mark	10 15	80% 93%
d	Multiple recoveries	24	79%

there isn't a mark on that side. If the agent keeps making observations, it can ideally eliminate anchoring hypotheses (signified by (anchor *s* = *x*)) until only one remains. It can then perform the action (anchor *s x*). Recall that the goal is to have done an anchor.

PTLplan is a progressive planner, so it starts from the initial belief state and adds actions until a belief state satisfying the goal is reached. When an action results in several new belief states with different observations, the planner inserts a conditional branching in the plan and continues planning for each branch separately. In order to search more efficiently, the planner can also eliminate belief states that invalidate a given temporal logic formula.

The following plan is generated for looking for marks on a red ball pi2 from three different positions, starting from a fourth position:

```
((move r1_2) (look-at pi2)
 (cond
   ((mark! pi2 = f) (move r1_3) (look-at pi2)
    (cond
      ((mark! pi2 = f) (move r1_4) (look-at pi2)
       (cond
         ((mark! pi2 = t) (anchor g1 pi1) :success)
         ((mark! pi2 = f) (anchor g1 f) :fail)))
      ((mark! pi2 = t) (anchor g1 pi1) :success))
   ((mark! pi2 = t) (anchor g1 pi1) :success)))
```

Note how a conditional branching follows after each application of look-at: the first clause "(mark! pi2 = t/f)" of each branch is the observation one should have made in order to enter that branch, and the subsequent clauses are actions.

Plan execution

The anchoring plan is then executed: the actions such as (look-at pi2) are translated into executable perceptive and movement tasks (see field :execute in the definition of look-at above). The anchor action has a special role: it causes the symbol of the requested object to be anchored to a specific perceived object. The robot can then continue performing the task in its top-level plan that was interrupted.

Experimental Evaluation

To be able to experimentally evaluate the methods described above we have implemented and integrated them with a fuzzy behavior based system, the Thinking Cap (Saffiotti,
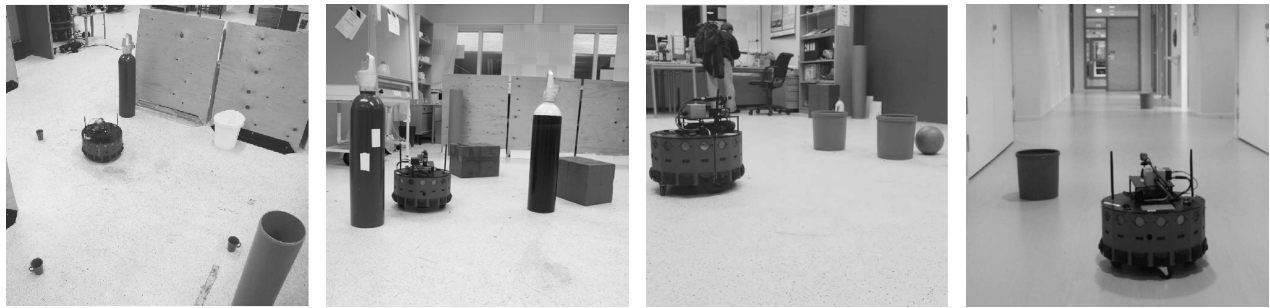
Figure 1: Experimental setups: (a) complementary sensors, (b) displaced object, (c) relations, (d) multiple recoveries.

Konolige, & Ruspini 1995), used for controlling a mobile robot called Pippi. Our primary sensor modality is through a vision system connected to the camera. As a second sensor modality we use a commercially available electronic nose (Cyranose Sciences Inc. 2000) capable of distinguishing between a number of odors. See Loutfi *et al.* (2004) for more information on how the electronic nose can be used together with the PTL planner and anchoring.

We present several experiments that illustrate a variety of ambiguous situations and how they are handled. The results of the experiments are summarized in Table 1. As performance measure we count the number of anchoring operations needed to achieve the top-level goal when performing a number of repetitions using the same setup. Note that we do not count any recursive anchoring operations performed during the recovery phase. We compare the total number of successful operations with the total number of required operations for all repetitions of the respective setups. For the first three scenarios only one top-level anchoring operation is required per setup and the numbers thus also reflects the total number of runs. The last scenario was run a total of eight times and each run required the top-level anchoring of three objects.

Our system has a success rate between 76% and 93% for these experiments, and most failures happened because the perceptual system delivered wrong data. The planner never failed to find a plan, and typically did so in a fraction of a second, with 1.5 s as the longest recorded time.

It should be noted that each one of the scenarios below could potentially be solved by a *ad hoc* procedure, provided that the type of perceptual anomaly encountered were known beforehand. One of the strengths of our approach is that one and the same generic domain model was used to deal with all these situations in a uniform way.

(a) Recovery using complementary sensors

In this first series of experiments we show how ambiguous cases due to partial matching of the description can be resolved by using a complementary sensor modality. The experiments are performed by using a number of cups containing different substances. The cups are scattered throughout a room containing other objects and the task of Pippi is to find a cup that is characterized by both a visual and an odor description (Figure 1 (a)).

When Pippi attempts to anchor an object with the description "the green cup that smells of ethanol", it finds several

objects matching the visual description. In order to recover from this situation, Pippi determined that the odor part of the description ("smells of ethanol") was needed and used its e-nose to sample the different cups.

A number of configurations of the above scenario where tried with 2 to 5 cups containing different substances. The odor part of the description as well as the positions of the cups and other objects where varied. Pippi was given the description of the target object and generated and executed a conditional plan examining all candidate cups until the correct one was found.

The occasional failures were due to either misclassification of odors and/or problems with the vision module.

(b) Reacquiring Displaced Objects

The second series of experiments concerned the resolution of ambiguous situations when an object is to be reacquired. The scenario involved two or more identically colored gas bottles in a room together with other objects such as boxes (Figure 1 (b)). One gas bottle named gb-b was distinguished by a white mark on one side. Pippi started from one position in the room, and tried to find gb-b by visually scanning the room. As the mark was initially turned towards Pippi, the correct gas bottle was easily found. While the robot performed other tasks the gas bottles where rearranged so that the position of gb-b was no longer accurate. The next time Pippi needed to reacquire gb-b, it could not determine which gas bottle was the correct one by using position information or by observing the, now occluded, mark.

Pippi then generated a plan to go to different positions in the room observing the gas bottles and looking for the mark. A number of variation of the scenario were tried, varying the number of gas bottles, their rotation, initial and final positions. For the result in Table 1 we count here only the reacquire operations since the initial find operations where all trivial to perform, counting also the find operations would instead have yielded success rates in the range 90-95%.

When several gasbottles were involved the recovery was non-trivial because multiple perceptual faults often led to additional recursive recovery plans being executed.

(c) Anchoring with relations

Two series of experiments involving relations to other objects have also been run. The first scenario uses a description "the green garbage can near the red ball", where "near" is defined as a fuzzy relation on the estimated positions of

percepts. There are two green cans visible, but the red ball is hidden behind one of them. Here, Pippi generates a plan to tour the room looking for the red ball, and when it is detected, to check which can it is near. The second scenario uses the description "the green garbage can near the red ball with a mark", and involves a single can and a red ball with the mark turned in different directions. In this scenario, Pippi generates a plan where it searches for a mark on the secondary object, i.e. the ball.

(d) Multiple recoveries

The final set of experiments was meant to test recovery from multiple anchoring failures while executing a longer top-level plan. We gave Pippi the goal that three known garbage cans should be inspected and that Pippi should return home. From the initial information the planner generated a plan consisting of 15 steps, including 3 anchoring operations and which was executed successfully in only some of the setups. In some other setups perceptual ambiguities were encountered during the execution: a mark was not visible and/or a nearby ball was occluded. In different cases different top-level anchoring operations required additional actions resulting in the total execution of up to 27 actions. Pippi successfully handled most of these cases in ways similar to those described in the previous experiments and achieved the top-level goals. Most of the failures were caused by faults in the low-level perception, and a few by bad self localization. The cases where subsequent anchoring operations where not executed due to earlier failures where also counted as failures here. Otherwise we would have a success rate of 86%.

Conclusions

To a mobile robot that is sensing, reasoning and acting in a complex environment, the ability to anchor symbols to perceptual data is essential. In this paper we have extended the anchoring framework to deal with ambiguous cases where there is more than one percept that matches the symbolic description of an object. We have proposed an approach for actively solving such cases by automatically analyzing and encoding the ambiguous situation as a planning problem and generating a conditional recovery plan. Our approach has been implemented and tested in a variety of experiments involving a mobile robot with sensing capabilities such as vision and olfaction. The algorithms and domain knowledge used in the recovery planning do not depend on the specific failure type and the larger examples could be handled automatically using the same mechanisms as in the single-failure cases. An additional contribution is the use of relations in the description of objects. Relations allow for richer descriptions and more complex ambiguity cases.

Acknowledgements

References

Beetz, M., and McDermott, D. 1997. Expressing transformations of structured reactive plans. In *Proc. of the European Conf. on Planning,*, 64–76. Springer.

Beetz, M.; Arbuckle, T.; Cremers, A. B.; and Mann, M. 1998. Transparent, flexible, and resource-adaptive image processing for autonomous service robots. In *Proc. of the 13th ECAI Conf*, 158–170.

Broxvall, M.; Karlsson, L.; and Saffiotti, A. 2004. Steps toward detecting and recovering from perceptual failures. In *Proceedings of the 8th Conference on Intelligent Autonomous Systems*, 793–800. IOS Press.

Coradeschi, S., and Saffiotti, A. 2000. Anchoring symbols to sensor data: preliminary report. In *Proc. of the 17th AAAI Conf.*, 129–135.

Coradeschi, S., and Saffiotti, A. 2001. Perceptual anchoring of symbols for action. In *Proc. of the 17th IJCAI Conf.*, 407–412.

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

Cyranose Sciences Inc. 2000. The cyranose 320 electronic nose. User's Manual Revision D.

Fikes, R.; Hart, P.; and Nilsson, N. 1972. Learning and executing generalized robot plans. *Artificial Intelligence* 3(4):251–288.

Giacomo, G. D.; Iocchi, L.; Nardi, D.; and Rosati, R. 1997. Planning with sensing for a mobile robot. In *Proc. of the 4th European Conf. on Planning*, 158–170.

Karlsson, L. 2001. Conditional progressive planning under uncertainty. In *Proc. of the 17th IJCAI Conf.*, 431–438.

Kovacic, S.; Leonardis, A.; and Pernus, F. 1998. Planning sequences of views for 3-D object recognition and pose determination. *Pattern Recognition* 31:1407–1417.

Lazanas, A., and Latombe, J. 1995. Motion planning with uncertainty: A landmark approach. *Artificial Intelligence* 76(1-2):285–317.

Loutfi, A.; S.Coradeschi; Karlsson, L.; and Broxvall, M. 2004. Putting olfaction into action: Using an electronic nose on a multi-sensing mobile robot. In *Proc. of the IEEE Int. Conf. on Intelligent Robots and Systems*.

Pell, B.; Bernard, D.; Chien, S.; Gat, E.; Muscettola, N.; Nayak, P.; Wagner, M.; and Williams, B. 1998. An autonomous spacecraft agent prototype. *Autonomous Robots* 5(1):1–27.

Saffiotti, A.; Konolige, K.; and Ruspini, E. 1995. A multivalued-logic approach to integrating planning and control. *Artificial Intelligence* 76(1–2):481–526.

Seabra-Lopes, L. 1999. Failure recovery planning in assembly based on acquired experience: learning by analogy. In *Proc. IEEE Intl. Symp. on Assembly and Task Planning*.