Semi-autonomous Map-making and Navigation

Karen L. Myers **Kurt Konolige**
Artificial Intelligence Center
SRI International
333 Ravenswood Ave.
Menlo Park, CA 94025
myers@ai.sri.com konolige@ai.sri.com

Abstract

Map construction and task-oriented navigation constitute two important capabilities for mobile robots. Despite their value, progress to date in the development of stand-alone robots endowed with these capabilities has been limited. We describe an architecture for mobile robots based on the concept of *semi-autonomy* that employs communication with a human advisor in order to simplify the tasks of map construction and navigation. Communication can be used both by the human to convey information to the robot that is beyond its own perceptual capabilities and by the robot to extract information from the advisor that may be required in the execution of an assigned task. The semi-autonomous philosophy has two main benefits, namely a de-emphasis on the notoriously difficult subtask of perception and an overall increase in flexibility.

1 Introduction

The construction of mobile robots that function autonomously in the real world has been a major objective within the AI community since the first days of work on intelligent systems. It has been envisioned that such robots would navigate through an environment, managing to both learn the lay-out and perform assigned tasks. Perceptual apparatus would provide information about the surroundings and somehow be assimilated into both the reasoning process of the robot and representation structures that correspond to learned maps of the environment. Thus, a robot placed in an office building might be expected to generate a representation of the floor plan of the building complete with markings for significant objects such as corridors and doorways, and to utilize that map in navigating to points of interest.

Maps are essential for purposeful navigation in unengineered environments, primarily because they enable strategic planning of routes to locations of interest. Navigation with respect to plans is certainly more efficient than the blind exploration that must occur when maps are unavailable. Although the information that maps provide can be engineered directly into the environment to a certain extent, such engineering greatly reduces the scope of possible navigational tasks. For example, preferred paths for travelling between various destinations in a building could be embedded directly into the environment in the form of stripes on the floor. This approach limits the range of the robot's travels to the predesignated set of paths, leading to a highly inflexible system. As well, environmental modifications of this nature can be expensive. While engineering the environment may be suitable for certain limited cases, maps are required for more general settings.

The question arises as to how and when maps should be constructed. For certain environments, maps can be handcrafted for the robot by a human prior to runtime. Handcrafting requires the efforts of an individual who is well-versed in the representational requirements of the robot; as well, the time and effort involved impose limitations on the scope of the domain that can be mapped. When robots reach the point where their commercial distribution is feasible, it will be untenable to impose the responsibility of map construction on non-expert purchasers of such systems. For many domains, the use of human-designed *a priori* maps is impossible rather than simply impractical: applications for outdoor mobile robots often require navigation through uncharted areas, such as exploration of ocean floors or other planets. Certainly, automatic map-learning is an important skill for many situations.

To date, work on map-learning for mobile robots has focused on the low-level problems of feature recognition and extraction of topological structure and metric information from sensor data [2, 3, 5, 4, 7]. Maps produced by these methods provide geometric descriptions of the environment. Two fundamental problems restrict the utility of these maps: *uncertainty* and the lack of *aperceptual information*.

Uncertainty arises in map-making due to the inherent complexity of sensing and perception. Sensors are subject to noise and interference, leading to substantial difficulties in interpreting sensed data. Even with completely accurate sensing, interpretation can be problematic due to a lack of sufficient perceptual cues. In general, representations of an environment constructed

solely from sensor information will contain gaps corresponding to regions where perception was inadequate.

Aperceptual information consists of those properties of the environment that cannot be determined from sensed physical characteristics. The classification of a property as aperceptual or perceptual depends upon the sensing mechanisms in use. One important type of aperceptual information for range-finding sensors is an *identifier* (*i.e.*, a unique name) for a perceived object, such as the room number of an office. A second type is a *label*, which provides descriptive information about individual perceived objects that cannot be derived from sensed information. For example, doorways to different types of rooms in a building (*e.g.*, offices, closets, supply rooms) can have very similar or equivalent dimensions despite vastly differing functionalities for the corresponding rooms. This type of situation arises with the openings in our laboratory; in our case, it is impossible to distinguish office doors from supply room doors based on physical characteristics only. For our laboratory, type labels for openings are aperceptual.

Although identifiers and certain labels constitute aperceptual properties for range-finders, other sensing modalities could be employed to extract some of this information from the environment. For example, cameras could be used to read identification plates on doors or corridor walls. With any sensor system there will be properties that cannot be perceived; we focus on sonars in our work, for which identifiers and opening labels are aperceptual.[1]

Maps that lack aperceptual information are of limited utility to a robot charged with navigational assignments. Without type labels, it would be impossible for a robot to locate a supply room on a map. Similarly, maps without naming information are of little value to a robot assigned the task of delivering mail or of coordinating locations with other robots. To appreciate the power that aperceptual information provides, consider the analogue for the case of maps defined for human use: what is the value of a map of California if it does not distinguish rivers from roads or railway lines? Or worse, if it omits the names of streets, cities and highways?

We adopt the pragmatic position that perception alone is insufficient for many desired functions in environments where *a priori* maps are unavailable. As a corollary, we believe that domain-dependent information provided by non-perceptual sources is essential to successful execution of tasks such as map-making and task-oriented navigation. In this spirit, we have explored the design of *semi-autonomous* architectures for map-making and navigation. In our semi-autonomous system, an adviser interacts with the robot to provide information about the environment that is beyond the perceptual capabilities of the robot. This information, which can encompass both aperceptual and geometric properties, is expressed in some high-level symbolic language. We focus on the use of a sentential language (first-order logic) for the sake of generality, although one could easily envision the use of alternative languages (such as diagrams). In order to improve the quality of its internal maps, the robot will perform sentential reasoning to integrate communicated information into its internal representation structures. Furthermore, execution of navigation tasks may involve reasoning about the current map in order to determine how best to proceed.

Our notion of semi-autonomy can be viewed as a lessening of the perceptual responsibilities of the robot. In essence, communication of domain information constitutes a form of assistance for the problem of modeling the environment. Note that the robot remains fully autonomous with respect to decision-making in our framework; in particular, no guidance is provided for executing tasks. The dimension of semi-autonomy explored in this paper differs greatly with previous semi-autonomous architectures in which humans play a direct role in robot *activity*, either by preprogramming the robot's behaviour, by providing run-time advice for achieving goals or by tele-operating the robot.

In this document, we present a brief overview of the capabilities of a semi-autonomous robot and our progress to date in building a semi-autonomous architecture for SRI International's custom-built mobile robot, Flakey.

2 Semi-autonomous Robots

The semi-autonomous framework provides two basic capabilities not found in traditional architectures.

The first is the ability to assimilate information about the environment provided by the human advisor into the current internal map, resulting in a more complete representation than can be attained by perception alone. Integration of communicated information in this manner allows the robot to both eliminate areas of uncertainty from its internal maps and to augment the maps with aperceptual information. Information provided by the human could be formulated as a partial domain theory provided to the robot prior to its exploration or could be communicated to the robot during run-time.

The second extension is the robot's ability to reason about its assigned tasks in order to determine whether or not they can be executed given the current knowledge of the environment. In cases where the task is determined to be unexecutable, the robot would initiate a dialogue with the human advisor in which it

[1] We note that the problem of aperceptual information can be eliminated through appropriate instrumentation of the domain, such as strategic placement of barcodes that embed all relevant information about objects of interest. Although such engineering of the environment is possible in certain restricted situations (fixed factory floors, robotics laboratories), many applications of interest do not permit such environmental aids.

would pose questions designed to elicit the information required for execution of the task. The responses obtained by the robot may affect both the strategy for executing the current task and the map currently held by the robot.

We consider two simple scenarios that demonstrate the above capabilities.

Scenario 1: Map Improvement

Consider the following representation of a hallway. The diagram contains the kind of information that might be present in a map generated solely from a robot's perceptual input. Solid lines represent corridor walls while openings indicate doorways or corridors. The dashed line indicates a region where sensor information could not be accurately interpreted (due to noise, interference, sensor error or a lack of perceptual cues). The shaded circle denotes the current approximate location of the robot.

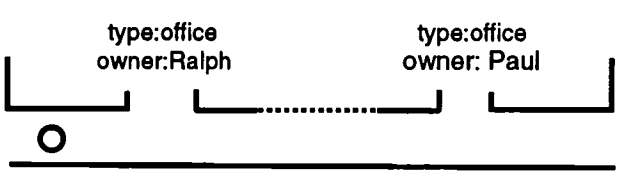
(1)

The diagram consists solely of perceptual information, namely the identification of openings and walls, along with spatial relations among them.

Suppose that after the robot has generated this initial map from its perceptual input, an advisor communicates the following facts to the robot:

The opening closest to you is Ralph's office.
The opening furthest from you is Paul's office.

In our framework, this information would be assimilated into the map (1) to produce the new map

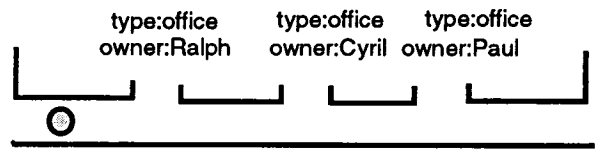


(2)

The new map contains a mixture of both geometric information and aperceptual information (specifically, type and ownership labels). Given the further domain knowledge

Ralph and Cyril are neighbors.
Cyril and Paul are neighbors.

we would expect the robot to extend its map to the following:



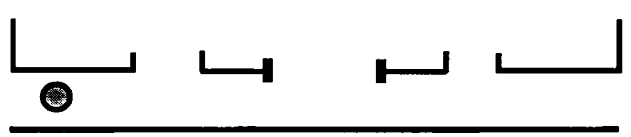
(3)

This scenario illustrates how the problems of aperceptual information and uncertainty inherent to the traditional approaches to map-learning can be overcome through a combination of communication and sentential reasoning. Here, reasoning with the communicated information resulted in both the addition of aperceptual information (*i.e.*, type and ownership labels) to the map and the elimination of the gap of uncertainty between the two perceived openings. With the last map (3), the robot would now be capable of executing the task *Go to Cyril's office*; this task could not be performed with the map (1) generated solely from perceptual input.

The reasoning required to progress from map (1) to map (3), although simple for humans, is nontrivial to mechanize. Details of the issues involved can be found in [9].

Scenario 2: Taskability

Suppose the robot has built the following map of a particular hallway using only its perceptual capabilities. In this representation, the middle opening corresponds to an elevator.



(4)

Consider tasking the robot with the command: *Go to the supply room in your hallway.* The robot cannot execute the command given the knowledge of the hallway displayed in map (4) since the supply room could be either the first or third opening in the corridor. An appropriate behavior on the part of the robot at this point would be to pose the question:

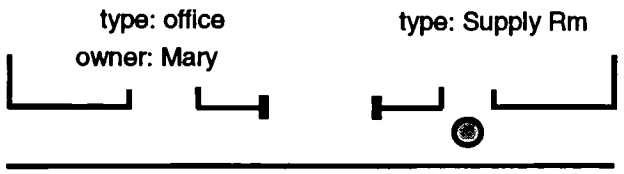
Is the supply room the first or last room on the left-hand side?

The human advisor may respond by informing the robot that the third opening leads to the supply room, or may supply the indirect but more informative answer:

The first doorway is Mary's office.

In the latter case, we would both expect the robot to proceed to the third doorway and to update its map

to the following:



3 Progress to Date

Realization of the capabilities described in the previous section constitutes an ambitious research project. Our efforts to date have focused on the task of semi-autonomous map-learning. Here, we describe what has been accomplished so far and discuss outstanding issues that are currently under investigation.

Semi-Autonomous Map-learning

The fundamental technical challenges for semi-autonomous map-learning are (A) bridging the gap between the sentential communication language and the robot's internal representations of perceptual information, and (B) designing appropriate perceptual interpretation routines for generating initial map structures.

Consider first the integration problem (A). Successful perception involves many layers of interpretation and hence a corresponding hierarchy of representation structures [1]. Using the terminology of [6], we can subdivide these layers into three types: *sensor data* is refined into an *image-level* representation, which is further interpreted to provide a *scene-level* description. Although sentential reasoning can be put to good use throughout, we focus primarily on applying sentential reasoning to improve a scene-level interpretation of sensed information. We assume that special-purpose analogical representations serve as the primary repository for geometrical interpretations of sensor data.

We have developed a formal framework for unifying sentential and scene-level map representations, along with an inferential calculus for integrating the contents of these two representational media. A formal account of the integration framework can be found in [8, 9]. The framework has been proven correct, meaning that it generates only maps and new sentences that are indeed consequences of the initial map and the sentential theory.

We have implemented the integration framework on top of the KLAUS automated deduction system [10]. The framework has been successfully applied to extend scene-level maps that are *conservative*; conservative maps can be partial but do not register spurious objects. To date, these maps have been produced by hand rather than *via* perception. The implemented system can successfully duplicate the behavior described in Scenario 1.

The second major challenge is to generate preliminary maps from perceptual input. The perceptual apparatus employed by our robot Flakey for the map-generation task consists of four Polaroid sonar sensors, two on each side of the robot.[2] The sonars have a range of two meters and a dispersion angle of thirty degrees. Preliminary interpretation routines for the perceptual input have been implemented and are producing reasonable results. The routines recognize different classes of openings, such as doorways (either open, half-open or closed), corridors and elevators. In addition, the routines identify areas where perception is unreliable; this information is used to register explicit regions of uncertainty in the constructed maps. As the basis of the interpretation routines, sonar readings from the previous five seconds are registered and a segmentation algorithm is used to construct piecewise linear artifacts. A combination of model information for features in the domain (for example, the widths of doorways and elevators and the offset of closed doors from walls), linear segments and sonar clusters that have interesting grouping properties are used by the interpretation routines in the feature extraction process.

We are currently in the process of increasing the robustness of the perceptual routines and organizing their output to produce maps. We note that since we are not relying exclusively on perception to build the robot's representations of the environment, the routines need not meet the high standards of accuracy sought by traditional work in perception. Rather, they need simply be conservative. Certainly, useful conservative perception and the recognition of unreliability are nontrivial tasks. It remains to be seen to what extent we can develop sufficiently robust routines. We hope soon to extend our integration framework to handle *almost conservative* maps, in which a small number of spurious objects may appear.

Semi-autonomous Navigation

Semi-autonomous navigation will rely extensively on the hybrid reasoning subsystem described in the previous section. In addition, strategies for introspectively reasoning about knowledge will be required to both evaluate the feasibility of assigned tasks and, for tasks that are infeasible with respect to the current knowledge, to identify the specific information that is required to successfully complete the tasks. Our efforts in this area are still preliminary.

4 Summary

The proposed semi-autonomous approach to map-learning and navigation has two main benefits not found in traditional architectures.

First of all, the ability to communicate domain information to a robot decreases the reliability require-

[2] Input from a structured light system may be used in the future.

ments imposed on the robot's perceptual processes. Rather than demanding high standards for the notoriously troublesome task of perception, simpler perceptual modules can be employed and communication used to improve upon the weaker results. In this regard, semi-autonomous systems for map-making and navigation are more realistic than traditional architectures that assume accurate perceptual routines.

Semi-autonomous operation has the added benefit of increased flexibility. Consider building robots for commercial distribution. It would be impractical to expect the purchaser of such a system to hire a programmer to customize the system to the situation at hand. Communication of relevant environmental information by some sort of adviser is a natural initial learning process experienced by any thinking organism placed in a new environment. In realistic domains, the learning process is never-ending: people move offices, buildings are modified, new tasks arise, *etc.* Thus, having established communication protocols for interacting with humans is of continuing importance.

We note that the independence of a robot built in accordance with our semi-autonomous philosophy will increase over time. As the robot uses communication to extend the information content of its maps, it will eventually acquire all of the domain knowledge that is relevant to its objectives. The robot will be fully autonomous at that point.

Acknowledgements

This research was supported by the Office of Naval Research under Contract N00014-89-C-0095.

References

[1] A. Bobick and R. Bolles. An evolutionary approach to constructing object descriptions. In *Proceedings of the Fifth International Symposium on Robotics Research*, Tokyo, Japan, 1989.

[2] R. Brooks. Visual map making for a mobile robot. In *IEEE Proceedings International Conference on Robotics*, pages 824–829, 1985.

[3] R. Chatila and J. Laumond. Position referencing and consistent world modeling for mobile robots. In *IEEE Proceedings International Conference on Robotics*, pages 138–170, 1985.

[4] T. Dean, K. Basye, R. Chekaluk, S. Hyun, M. Lejter, and M. Randazza. Coping with uncertainty in a control system for navigation and exploration. In *Proceedings of the Eighth National Conference on Artificial Intelligence*, pages 1010–1015, 1990.

[5] B. J. Kuipers and Y.-T. Byun. A robust, qualitative method for robot spatial reasoning. In *Proceedings of the Seventh National Conference on Artificial Intelligence*, pages 774–779, 1988.

[6] A. Mackworth and R. Reiter. A logical framework for depiction and image interpretation. *Artificial Intelligence*, 41, 1989.

[7] P. K. Malkin and S. Addanki. LOGnets: A hybrid graph spatial representation for robot navigation. In *Proceedings of the Eighth National Conference on Artificial Intelligence*, pages 1045–1050, 1990.

[8] K. L. Myers and K. Konolige. Integrating analogical and sentential reasoning for perception. In *Proceedings of the AAAI Spring Symposium on Reasoning with Diagrammatic Representations*, 1992.

[9] K. L. Myers and K. Konolige. Reasoning with analogical representations. In B. Nebel, C. Rich, and W. Swartout, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the Third International Conference (KR92)*. Morgan Kaufmann, 1992.

[10] M. E. Stickel. The KLAUS automated deduction system. In *Proceedings of the Ninth International Conference on Automated Deduction*, 1988.