CHAPTER I

INTRODUCTION

Mobile robots, in order to be successful in a world as unconstrained as a human's, must be capable of responding intelligently to changes in their environment. Safe and successful goal-oriented navigation can only occur if the robot is able to emulate intelligent behavior. It should be recognized that in the real world:

1. Things are not always as they appear.
2. The world changes over time.
3. The spatial limitations of sensing when combined with mobility lead to new perceptions as previously unknown parts of the world are encountered.

Intelligent behavior can be defined as the ability to respond to changes in the perceived world in an advantageous manner. It should be made clear that a sharp distinction exists between reality and the perceived reality of the senses. A view of the world is just that, a view, and is not the world itself. If this fact is ignored and all perceptions are viewed as valid unambiguous interpretations of reality, any cybernetic system (animal or machine) is doomed to a short-life span. Consequently, it is of fundamental importance for such a system to be able to extract from the wealth of data provided by its senses a coherent interpretation that is subject to later revision and possible revocation based on freshly acquired information. Simply put, a robot must be able to alter its beliefs.

An AI system with the ability of handling this difficulty is usually vulnerable to the frame problem [46] i.e. the system is not fully aware of all the consequences of any action that is undertaken. Fortunately, action-oriented sensing can serve as a means to cope with this ill.

In order to design a mobile robot system with the ability to behave intelligently, the real-world characteristics enumerated above must be addressed. Handling errorful perception requires a mobile robot system capable of frequent sensor sampling and uncertainty management. Multiple sensors utilizing appropriate sensor fusion techniques can also ameliorate problems arising from the differences between the real and perceived worlds. Spatio-temporal changes in the real world require frequent updating of the robot's internal world model. It is unsafe to assume, except in highly structured workplaces, that a static environment exists. World modeling and navigation should additionally extend to regions where the robot has never been before or which it has not encountered recently. Ideally, learning can also be used to adapt to slowly varying changes in the world and to add new locales to the robot's memory.

Most mobile robot systems built to date have had a narrow focus of interaction with their environment. There are road-following vehicles, hall-following robots, guide-dog robots, etc. Little effort has been placed, however, on the development of a more general purpose robot capable of functioning well in both indoor and outdoor environments. If robots are expected to be other than special-purpose (although programmable) machines, a bridge must be built to allow the transition from constrained environments to more open ones. The key to accomplishing this transition is the construction of an architecture and accompanying representations that are less restrictive than those previously employed.

The thrust of this dissertation is the development of a mobile robot navigation system (AuRA - Autonomous Robot Architecture) that can operate in environments to which civilized man is accustomed. This system is developed and experimentally tested in what are termed extended man-made environments. These environments include the interiors of buildings, streets, sidewalks, city and campus settings. It is not intended to deal with navigation in jungles, swamps, forests and other unstructured environments, in which indeed a human - without a compass, clear skies, or additional information - can become hopelessly lost. The principal reason for the choice of these environments for the mobile robot's domain is the wealth of visual cues they afford the vision system. The assumption is made that man-made objects are more readily discernible via machine vision than natural ones, largely due to a significant population of regularly shaped edges or characteristic colors of such objects. These cues can be used by the robot for localization, guiding it to greater confidence in its position relative to the world, as well as goal recognition and obstacle avoidance.

§1. Navigation

Navigation can be simply defined as moving from one place to another in an effective manner. One may navigate amongst that which is known ahead of time, or in unknown territory. In a perfectly modeled world, a path can be computed in advance that is completely acceptable. The robot's task becomes simply to maintain its bearings relative to the precomputed global path so that it does not deviate from this predetermined route. At the other extreme, where nothing is modeled ahead of time, the robot reacts to its environment, seeking an unknown goal amidst unforeseen barriers. The first case, moving in a known world, will be referred to as map-navigation, and the second case, reacting to unknown events, will be referred to as piloting.

Most real world situations involve both navigational forms. When given a task, any *a priori* knowledge of the world (if available) is mustered prior to the initiation of movement. In a perfectly modeled world, simple path adherence is all that is required. In a less than completely known world, (i.e. the real world), the robot starts moving along a predetermined path. If no unusual events occur, the pilot maintains the map-navigator's path. If an unexpected event occurs, the pilot initiates dynamic sensor-based replanning. As long as the deviations are minor, the pilot handles the replanning. If the path deviations are too severe, the pilot informs the map-navigator, who computes an alternate path.

An important distinction is that map-navigation is model-driven and sensor-independent whereas piloting is largely sensor-data driven. It is entirely possible for the pilot to inform the map-navigator of newly sensed information so the world model can be updated to incorporate this new data. This is in essence a learning mechanism. It is important to note however that not everything the pilot detects should be entered into the map-navigator's model. Moving obstacles (cars, people, etc.) may be present only for a very short time and should not be added to the navigator's map. This implies that semantic interpretation of the sensed data is essential for navigational updating of the world map.

Another important distinction between map-navigation and piloting lies in scope.

Map-navigation is more global in nature, ignoring the small detail to arrive at a satisfactory global solution (based on available knowledge). The short-sighted pilot, on the other hand, is concerned with the immediate environs and deals with sensed but unmodeled obstacles that are not represented in the navigator's map. Meystel describes the role of scope in navigational planning in [83].

Sidestepping the issue of learning for the moment, the maintenance of two distinct representations for the different levels of navigation facilitates planning: long-term memory (LTM) for the static world representation used by the map-navigator, and short-term memory (STM) which is used by the pilot for building up a perceived model of the world. Localization (determining the robot's position relative to the global LTM map) becomes a matter of correlating STM with LTM, and learning involves moving relevant features from STM to LTM. The pilot also draws on LTM but in a more limited way, LTM providing specific cues for the pilot to look for (e.g. landmarks) which can be used for localization. The pilot monitors only those landmarks in its vicinity and thus can more effectively use its available computational resources.

For the purposes of piloting in AuRA there exist two strategies: a low-level reactive or reflexive approach using schema-based control structures which does not draw directly on memory (either STM or LTM) but instead uses sensor data as it is received, and another method using both the local context of LTM and the accumulated sensor data in STM when the reflexive approach fails. Unmodeled and changing world conditions are treated through relevant schema and STM representations. The distinctions between these two methods will be discussed in Chapters 3 and 5.

A partial *a priori* model (LTM) of the domain of interaction is provided for the map-navigator's use. It includes static objects, (lampposts, walls, buildings, etc.), but omits dynamic ones (people, cars, chairs, etc.). The question of the robot learning and adapting its representations to meet the demands of a changing world is considered in the theoretical development of the architecture presented in Chapter 3.

§2. Characteristics of mobile robotics

Mobile robotics in many respects is decidedly different from conventional robotics. It is worth describing some of the characteristics that distinguish it from the more conventional robot arms and manipulators. Some of the material in this section is loosely adapted from Thorpe [126] and Andresen et al [3]. These characteristics include:

- Inherent inaccuracy
- Limited degrees of freedom
- Cumulative error
- Incomplete model
- Environmental uncertainty
- Non-repetitive paths
- On-line, continuous path planning

Inherent inaccuracy

The sensors relied upon by mobile robots can easily give rise to imprecise and inaccurate data. Even if the devices themselves are highly accurate, (e.g. shaft encoders), the correspondence of the changes in the sensors to the changes in the robot's environment may be poor (e.g. due to wheel slippage). Feedback in the traditional sense of control theory is not immediately applicable and can only be used as a guide to establish expectations for higher level processing.

Limited degrees of freedom

The number of degrees of freedom for the mobile robot are significantly less than those of a robotic manipulator. Assuming no translational motion is allowed in the vertical (up and down) direction (generally a necessity due to available locomotion systems - this would change if applied to legged, flying and submarine robots), there are 2 DOFs of translation and one DOF of rotation. This is half of the 6 DOFs commonly found in a robot arm and wrist, a decided decrease in complexity.

Cumulative error

Errors if left uncorrected will tend to increase. This is typical of any type of dead-reckoning system. A path cannot be computed and the robot sent off to execute it without frequently verifying and correcting the robot's internal model of its position. Merely avoiding obstacles along the way is insufficient to guarantee that a robot will reach its goal or even recognize when it reaches it. Consequently, information must be maintained in a representation that enables this type of updating to be performed.

Incomplete model

Any model by definition is incomplete, otherwise it would not be a model. Internal world representations for mobile robots are perhaps more incomplete than most, due to the larger and more unstructured world in which it operates when compared to industrial robots. Space-versus-time computational tradeoffs must be made in order to meet the real-time constraints of path planning, replanning, and obstacle avoidance. Excess representational baggage is a luxury that generally cannot be afforded. It is difficult to see how any representation can be maintained, updated and accessed by algorithms that must process the data in *real-time* with *existing* hardware and yet is complete enough for accurate positioning of the robot, semantic interpretation of high level commands and objective statements, supporting multi-modal sensors, coping with uncertainty, handling goal recognition and choosing alternate path-planning strategies dependent upon external factors.

Representational incompleteness can also be encountered when the robot is required to traverse areas in which it has never been before. The representation may have to be built dynamically from only partially correct and possibly contradictory sensor data.

Environmental uncertainty

Things are not always where they are expected to be, even when they are modeled. Not only is the robot's position uncertain, but the location of objects will also have to be treated with skepticism. In addition, objects may have moved since their last observation or even be in motion relative to the world or the robot. This is in marked contrast to the robot manipulator's highly structured workplace.

Non-repetitive paths

The path an autonomous mobile robot executes is unlikely to be the same twice. Although the general route may be the same, changing conditions combined with positional errors usually will require the actual path taken to differ from the high level specifications each time it is traversed. If this were not so, as in some manufacturing situations, a stripe or wire following automatic guided vehicle (AGV) would be the robot of choice instead of an autonomous vehicle.

On-line, continuous, path planning

The robot must not close its "eyes" for long while moving. Constant monitoring for collision avoidance is essential. To obtain enhanced performance, path planning should be maintained during robot motion, just in case an unexpected event arises that would necessitate a path change. These might include such things as an unanticipated barrier (necessitating a detour) or the absence of a modeled obstacle (opening up a better path). This dynamic replanning must be conducted in real-time.

§3. Overview of the dissertation

The ultimate goal of this dissertation is to provide a broad, relatively unrestricted approach to the problem of mobile robot navigation. To accomplish this, several specific issues will be addressed. These will include:

1. The development of an architecture to support intelligent navigation of a multi-sensory (predominantly visual) robot in a "civilized world".

   Approach: The Autonomous robot architecture (AuRA) is forwarded as the structure to accomplish this goal. The "civilized world" includes both indoor and restricted outdoor travel. The outdoor case will assume a significant population of man-made objects (roads, paths, sidewalks, buildings, etc.) but also will allow for substantial natural surroundings (grass, trees, sky, etc.). Extensibility and generalizability are considered fundamental design goals.

2. Effective fusion of the visual data with other sensory input (e.g. shaft encoders and ultrasonics). The development of techniques that are appropriate for resolving conflict between contradictory sensory data while enhancing cooperative input.

   Approach: The use of motor schemas and associated perceptual schemas as control mechanisms to funnel relevant sensory data to the appropriate motor task.

3. Representation of spatial uncertainty and its use to guide expectations for perception. The use of sensing to restrict the limits of uncertainty through feedback.

   Approach: The use of specific modules in the AuRA architecture dedicated to the management of uncertainty, including the spatial uncertainty map and its manager, and the Expecter used to provide expectations to the perception subsystem.

4. Selection of appropriate vision algorithms for specific tasks.

   Approach: Associating specific strategies through perceptual schemas to provide action-oriented perception. These include:

   - A depth-from-motion algorithm for obstacle avoidance.
   - A line extraction algorithm for path following.
   - A region segmentation for path following and localization.
   - Scene interpretation and interest operators for landmark recognition.

5. The use of knowledge representations that can effectively deal with navigation in "civilized" environments and are general enough to be used both indoors and outdoors.

   Approach: A multi-level representation managed by a cartographic process that maintains both an *a priori* model of the environment in addition to a dynamic model of newly encountered obstacles and features.

Experimental testing of the resulting robot system is in two arenas: navigation within the Graduate Research Center at the University of Massachusetts - Amherst (UMASS) and in the outdoor area surrounding the same building. The results are presented in Chapter 8.

The dissertation is structured as follows:

- Chapter 1 is this introduction.
- Chapter 2 describes relevant prior work in the field of mobile robot navigation, including an analysis of the types of representations and control strategies used. An overview of work in the use of vision as a sensor for robot navigation is also presented.

- Chapter 3 presents the structure of the Autonomous Robot Architecture (AuRA) that is used as the framework for the experimental work of this dissertation. Motivation for its structure is also included.

- Chapter 4 describes in detail the role of the navigator and long-term memory representations used to develop a path for the robot based on *a priori* knowledge. The inclusion of both indoor and outdoor terrain types are important extensions to previous work. The roles of the mission planner, pilot and their accompanying representations are also presented.

- Chapter 5 puts forth the motivation for motor schema based mobile robot piloting. Action-oriented perception is a fundamental tenet of this approach. The use of schemas in AuRA is described. The inter-relationship between the motor schema manager and pilot is also discussed.

- Chapter 6 describes the specific sensor algorithms, both visual and ultrasonic, that are used to provide environmental data to AuRA.

- Uncertainty management is discussed in chapter 7. An exposition of the role of the spatial uncertainty map and its manager and the use of represented uncertainty to guide perceptual expectations is provided.

- Chapter 8 presents the actual experiments performed to validate the concepts presented in this dissertation. These include both indoor and outdoor runs with differing levels of *a priori* knowledge available.

- Chapter 9 concludes the dissertation with a summary of accomplishments as well as a discussion of future work.