Chapter 1

Introduction

1.1 Motivation

One of today's most serious social, economical, and environmental problems in the world is traffic congestion. In the United States alone, the problem is costing over 100 billion dollars each year. In addition, more than 40,000 persons are killed and another three million are injured (an additional cost of $137 billion per year). This is compounded by the fact that it is becoming increasingly difficult, for both financial and environmental reasons, to continue to build additional highway capacity [Hayes94]. Therefore, the solution to the problem must lie in other approaches, one of which is to make better use of the existing highway infrastructure. "The current infrastructure has begun to plateau in its ability to meet the operational requirements of highway transportation." Consequently, congestion and safety risks are increasing. Previous solutions of building or widening roads are not feasible anymore; the average cost of a mile of urban freeway is $38 million. Furthermore, the "tens of thousands of lives per year is a steep price to pay for mobility" [Hayes94].

New approaches discussed in the next section have the potential of increasing the safety, reducing the congestion, and improving the driving conditions. Studies conducted at the beginning of the 'intelligent transportation' era show that it is possible to cut accidents by 18%, gas emissions by 15%, and fuel consumption by 12% by employing new approaches. It is claimed that these improvements will result in $10 return to the economy for each dollar invested in the new transportation systems [Card93].

Considering the potential of the application of intelligent systems to surface transportation problem, US Department of Transportation deems the investigation of new technologies as crucial. Most of the "smart highway" technologies were first suggested in the '60's, but the idea did not gain momentum until Intermodal Surface Transportation Efficiency Act (ISTEA) passed by Congress in 1991. The technology needed to create an intelligent transportation system is already available, although still expensive for full implementation. Today's vehicles have at least four or five microprocessors monitoring and controlling such things as ignition spark, fuel and emission controls, automatic transmission, cruise control, air bags, and anti-lock brakes. Further development is underway on computerized controls for active suspensions, traction control, all-wheel drive, and four-wheel steering. Few people realize, in fact, that today's car has four times the computing power of the first Apollo moon rocket [Card93].

1.1.1 ITS and AHS

The approach taken by the US Department of Transportation (DOT) is called Intelligent Transportation Systems (ITS; previously, Intelligent Vehicle Highway System, IVHS). ITS is actually a broad range of diverse technologies, including information processing, communications, control, and electronics. DOT's IVHS Strategic Plan [Plan93] summarizes the approach, mainly defining the use of intelligent vehicles and extensive communications between the vehicle and the roadside infrastructure. This initial plan set the goals, milestones, and objectives of the ITS program through 1997. The federal government is one of the many players in the implementation of a nationally compatible system. Others are found among the private sector, professional societies, consumer and industry groups, academia, and State and local governments.

Some of the direct goals and objectives of the ITS program are:

- Significant reduction in the number of annual fatalities and injuries due to accidents,
- Improvement in the safety of commercial vehicles and hazardous material movement,
- Significant reduction in the cost associated with congestion,
- Increase in the volume of people and goods that can be moved on existing highways,
- Improvement of travel time predictability, and reduction in the level of stress associated with travel,
- Reduction in harmful vehicle emission, and in surface transport energy consumption per vehicle-mile and per passenger-mile traveled.

While congestion is the primary problem in urban areas, rural areas suffer from a higher traffic fatality rate due to road conditions and higher rates of speed. Increased margins of safety provided by ITS technologies might reduce the number of accidents in both urban and rural areas. Most importantly, communication capabilities could improve emergency response times, and improve levels of rural public transportation services.

The established process of delivering the ITS program has four major components covering the following issues:

- Research and Development: The categories of research applicable to the development of the program include (a) providing basic research tools and a knowledge base, (b) creating a favorable environment and (c) defining potential opportunities for applications, and (e) developing ITS applications.
- Operational Tests are designed to evaluate applications of new technologies and system concepts, facilitating the transition into operational use.
- The Automated Highway System (AHS) is established to pursue the more technically challenging, longer term goals of having a fully operational vehicle-highway system that automates the driving process.
- Deployment Support: Unlike major national systems in aviation, space, defense, and other sectors, the DOT will not own and operate the national ITS.

The DOT has established a set of discrete milestones to be accomplished until the end of the year 1997. Some of these milestones related to the work presented here are system architecture, communication frequencies, collision avoidance, and automated highway systems.

A major long-term element of DOT's research and development is AHS. The deployment of early phases of AHS will begin in the next ten years, but will continue beyond that period. Current AHS program activities address the imminent user services involving safety advisory and driver assistance products, *e.g.*, in-vehicle warnings of impending danger, and emergency control intervention. The ITS architecture must be structured to provide a sensible evolution to accommodate these types of user services, as well as the eventual realization of total automation.

The 1997 demonstration requested by Congress is the AHS program's first major milestone, and will assess proof of concept feasibility for the AHS. Preparation for demonstration will include the identification and evaluation of feasible alternative conceptual approaches. The demonstration will feature many feasible concepts. Subsystems and control, including lateral and longitudinal control, transition between manual and automated control, maintenance of position in the roadway traffic flow, lane changing, and various malfunction management capabilities will be demonstrated all in a collision-free automated driving environment.

The demonstration results, if successful, will lead to selection of the AHS conceptual approach, documentation of the approach in performance specifications and standards, and establishment of a partnership with industry to design, develop, and operationally test an AHS with public participation. Underpinning the development of the first AHS demonstration is an understanding of the human factor issues related to AHS driver operations. The DOT has begun a research project to develop comprehensive driver performance guidelines for AHS. Beyond 1997, deployment of a nationally compatible system will become the primary concern of DOT. It is expected that the work will shift towards supporting deployment of ITS user services, and continuing progress on long-term ITS goals, especially the AHS. Details of the AHS concept, and major important factors are described in detail in Chapter 2.

1.1.2 Intelligent Vehicle Control

Vehicle control is probably the most important part of the advanced AHS applications. Implementation of AHS requires automatically controlled vehicles for the reasons listed in Chapter 2. The aim is to increase safety by removing the driver from control of the vehicle, and the throughput by automation of vehicle following, lane changing, and other maneuvers.

The past and present research on vehicle control emphasizes the importance of new methodologies in order to obtain stable longitudinal and lateral control. Combined lateral and longitudinal control needs to be implemented for full AHS applications. As indicated in Section 2.2, there is a large group of investigators working on vehicle control issues. However, being able to control the dynamics of a vehicle does not necessarily mean that we have an AHS. What we mean by 'intelligent vehicle control' is the ability of an automated vehicle to make decisions on its steering, and velocity commands, *i.e.*, its ability to plan its path. In an environment where there are multiple fast moving vehicles, making the right decision to avoid collisions and optimize vehicle path is difficult. There are several different approaches currently considered in the current AHS research. These and our position relative to current research are discussed next.

1.2 Context

1.2.1 Different Approaches to Intelligent Vehicle Control

There are two main approaches to the implementation of a fully automated highway system: hierarchical structure and autonomous vehicle approaches. The first approach, defined in [Varaiya93], centers around the notion of "platooning" (See Section 2.3). In this approach, there are different layers of control hierarchy, each responsible for a different task required to implement an automated highway system. The layers of the architecture, starting at the top, are: *network*, *link*, *planning*, and *regulation*. The network layer assigns a route to each vehicle as it enters the system. The link layer assigns each vehicle a path which balances traffic for all lanes and assigns a target speed for each section of highway. The planning layer creates a plan which attains the desired path. The regulation layer controls the vehicle trajectory so that it confirms to this plan.

The tasks assigned to each layer differ from each other in three "dimensions:" time scale, information span, and spatial impact of decisions. As we go up in the hierarchy, the frequency of decisions decrease, more aggregated information is used, and the decisions affect vehicles over larger distances. For example, the planning layer of the hierarchy makes decisions to create a plan that evolves every minute by coordinating with neighboring vehicles. Path decisions affects only the neighboring vehicles.

This approach assumes an information flow from the higher hierarchy and to/from other vehicles for making decisions about the optimal path. The information available to an automated vehicle has global characteristics since it is relayed from a higher layer of the hierarchy.

A simpler approach for reasons of implementation is the autonomous vehicle approach, where each vehicle is treated as autonomous agents using local sensors (and maybe limited communications with neighboring vehicles). It does not use the concept of closely separated vehicles and aims to solve the problem in a shorter time interval, and to reach a fully automated highway systems by an evolutionary implementation, *i.e.*, by slowly increasing the capabilities of autonomous vehicles over time.

In the first case, the control commands are relayed by higher layers in the hierarchy, and the vehicles do not need to make intelligent decisions for path planning; the hierarchical structure 'controls' the vehicles. In the latter approach however, the vehicles need to make decisions based on local information. Our learning planner can be visualized as part of the second approach, or a backup system for the first one.

The first attempt to solve the problem of real-time decision making in an automated highway environment dates back to 1992. Mourou and Fade described a "planning method applicable to agents with perception and decision-making capabilities and the ability to communicate with other agents" [Mourou92]. The agents, in this example, are assumed to transmit their actions to their neighbors. This work emphasized the possibility of achieving a more flexible, fast, and reactive system by multi-agent planning methods and execution monitoring. However, this approach requires constant and complete data transfer between agents.

Recent research on intelligent vehicles includes an adaptive intelligent vehicle module used in a simulation and design system for tactical driving algorithms [Sukthankar96a, 96b]. The approach divides the driving task into three levels: strategic (route planning, goal determination), tactical (maneuver selection for short term goals) and operational (translating maneuvers into control operations). Intelligent modules are designed to answer the need for real-time tactical level decisions. A reasoning system called SAPIENT combines the high-level goals with low-level sensor constraints to control vehicles. The system uses a number of modules whose outputs are combined using a voting scheme. The Population-Based Incremental Learning (PBIL) method is used to adjust the large number of parameters defining the behavior of the modules. This approach is basically a method for finding the suitable parameters for the module that are used to fire lateral and longitudinal actions. The learning algorithm for the parameters is a combination of evolutionary optimization and hill climbing, and it is very similar to learning automata reinforcement schemes except for the use of a mutation factor for the probabilities of actions at the end of each iteration.

The learning automata controller we describe here can be visualized as a combination of intelligent modules at the tactical level. However, our approach differs from the above mentioned research in the use of the learning algorithms. Instead of learning the parameters affecting the firing of actions on repeated runs, learning automata learn which action to fire based on the local sensor information. In other words, the learning is not at the design phase, but at the *run* phase. The parameters defined in [Sukthankar96a] correspond to learning (Chapters 3, 4, and 6) and sensor (Chapters 4 and 5) parameters of our controller, and can be used to adjust those. Learning and sensor parameters defines the capabilities of an automated vehicle, and can also be used to model different driver behaviors. For example, large learning parameters decrease the decision time, and consequently result in a more 'agile' vehicle path.

Another approach to intelligent vehicle controller for autonomous navigation uses a decision-theoretic architecture with probabilistic networks [Forbes95]. The problem is modeled as a partially observable Markov decision process, and the optimal action is a function of the current belief state described as the joint distribution over all possible actual states of the world. However, this work currently considers only combinations of dynamic decision networks with a decision tree, *i.e.*, if-then rules where each predicate is actually a complex set of probability thresholds on specific variables, and it is similar to the previous work in the sense of firing actions. Similarly, Niehaus and Stengel defined a rule-based navigation system that uses worst-case decision making (WCDM) approach. Again, a stochastic model of the traffic situation based on sensor measurements is assumed [Niehaus94].

1.2.2 Learning Automata as an Intelligent Controller

The concept of the learning automaton is defined as a result of the work on modeling observed behavior, on the choice of actions based on past observations, on implementation of optimal strategies in the context of the two-armed bandit problem, and on the need for rational decisions in random environments.

In classical control theory, the control of a process is based on complete knowledge of the system. Later developments considered the uncertainties present in the system. However, all those assumptions on uncertainties and/or input functions may be insufficient to successfully control the system. It is then necessary to acquire additional information online since *a priori* assumptions are not sufficient. One approach is to view these as problems in learning.

Learning can be defined as a change in behavior as a result of past experience. In a purely mathematical context, the goal of a learning system is the optimization of a functional not known explicitly [Narendra74]. A stochastic automaton attempts to find a solution of the problem without any information on the optimal action (the control input to the environment/system). One action is selected at random, the response from the environment is observed, action probabilities are updated based on that response, and the procedure is repeated. A stochastic automaton acting as described to improve its performance is called a *learning automaton*.

As seen clearly from the general description of a learning automaton, our approach differs from previous and current efforts to intelligently control autonomous vehicles in the "use" of the learning. Instead of learning the parameter/behaviors/firing rules for best actions to take for achieving a safe and optimal path, our controller *learns the action* in real-time. In that sense, the decision to fire an action is never taken at exactly the same time for similar conditions. Furthermore, there are no "prescribed conditions" for actions. The parameters that define the learning process of a stochastic automaton, as well as the sensor parameters defining decision ranges, can be *learned* too using methods similar to ones described in Section 1.2.1.

The idea of defining a "fixed" structure to be utilized to find the optimal action has its own appeal, since the performance of the system is deterministic in the sense that the best action for a specific situation is known. However, even a good driver does not follow rules deterministically. In this sense, the learning automata approach is able to capture the dynamics of the driver behavior. On the other hand, a rule-based system, although known to perform well on many situations, has the disadvantage of requiring extensive modifications, even for a minor adjustment in the rules. Furthermore, such systems cannot handle unanticipated situations [Cremer94].

We visualize two learning automata employing a reinforcement algorithm as the heart of our path planner. Using local sensor data (and maybe limited communications with neighboring agents), longitudinal and lateral automata learn the optimal action to be taken for a given situation. Sensor information is *processed* in virtual teacher modules that evaluate each action in the light of the current sensor status. Given a large enough time and sufficiently large learning parameters, the automata indicate the best action to take and sends them to a lower layer which in turns 'fires' the action. The details of the concept are given in Chapters 4 and 5.

This application of learning automata is one of the very few applications to real-time learning problems [Narendra89, Najim91b]. The difficulty in modeling the environment (highway, the vehicles, and the drivers) makes learning automata a likely candidate for such an application, since a learning system is guaranteed to adapt to existing conditions provided that the "teachers" are designed carefully.

1.3 Scope and Structure of Dissertation

Chapter 1 described the overall problem of traffic congestion and safety, as well as ITS, currently considered as a solution by DOT as well as other agencies throughout the world. The direct goals and objectives of the ITS program, its advantages and the role of Automated Highway Systems (AHS) to reach these goals as well as the importance of the intelligent vehicle control for AHS are also emphasized in this chapter.

Chapter 2 introduces the AHS program in detail. The phases of the program and National AHS Consortium are introduced to give a perspective on the current developments. Background on the previous and current research on vehicle control, proposed control structures, and related issues such as sensors and communication difficulties, and safety analysis are also given in Chapter 2. The chapter concludes with the description of the evaluation and experimentation framework designed at Virginia Tech Center for Transportation Research. Current small scale and simulation studies for AHS are briefly described.

Chapter 3 is an introduction to learning automata and reinforcement learning. All the necessary descriptions and ideas to understand our application of learning automata to intelligent vehicle control are given. This chapter also lists past and present research on learning automata and previous applications of learning automata.

Chapter 4 and 5 describe the use of learning automata as an intelligent decision maker for vehicle path planning. The application of learning automata techniques to intelligent vehicle control is introduced in Chapter 4. The basic automata-teacher model and abstraction of the sensing modules are given; the idea of 'learning necessary actions' for a vehicle to survive in a dynamically changing environment is introduced. The idea differs from the previous and current research by its on-line learning approach. The method described in Chapter 4 is found to be capable of safely directing the autonomous vehicles. Simulations of learning model also indicated that there is a need for additional sensor information as well as faster learning algorithms.

Result of Chapter 4 leads to the extension of the decision capabilities of an automated vehicle by extending its capabilities sensing and communication capabilities and by defining a more complex decision structure, given in Chapter 5. Our investigation of the control of autonomous vehicles using local information determines the need for new and improved sensor structure. Chapter 5 details the necessary additions for more intelligent vehicles in order to optimize the overall "behavior" of the traffic, and provides solutions to problems generated by intelligent vehicle interactions. The 'flag structure' described in this work enables autonomous vehicles to reach their immediate goals while optimizing the traffic flow in close neighborhood.

The overall design structure, effects of the several parameters on the decision, and several other issues related to intelligent path planner conclude Chapter 5.

In Chapter 6, we introduce reinforcement schemes (learning algorithms) used for our application of learning automata. The schemes are compared according to their desired update characteristics. A new nonlinear scheme and a linear scheme previously not considered by previous research efforts are found to be more advantageous than others. For the linear scheme, proof of optimality for a specific case in our application to intelligent vehicle control is given. Instead of general linear differential equation approach [Narendra89], stability theorems for discrete-time systems are used for this purpose. The behavior of the new nonlinear reinforcement scheme is also discussed in this chapter. It is found that the algorithm is absolutely expedient. These two new schemes are the results of our attempt to find fast converging reinforcement schemes.

Chapter 7 introduces the concept of interacting automata and the treatment of multiple automata as *game playing* agents. The concept of the automata games is then extended to multiple automata in multiple vehicles. An approach similar to automata games [Narendra89] is taken for interacting vehicles as well. By defining the physical environment of vehicle as a switching environment, it is possible to treat a highway situation as state transitions in a Markov chain. Associating automata environments to each stationary physical environment, it is possible to analyze and design physical environment changes as state transitions based on the reward-penalty structure of the automata environment[1]. This approach led to the design of the flag structure explained in detail in Chapter 5.

A brief discussion of technical feasibility and user acceptance of AHS, the conclusions drawn from the work presented here, and our recommendations for future research as well as the discussion of the problems we encountered in this research are listed in Chapter 8.

Notations and definitions of the variables, list of abbreviations and acronyms, additional proof of convergence for a simple linear reinforcement scheme, a brief description of the simulation program and diagrams of the 'states' for highway scenarios mentioned in Chapter 7 are all given in the Appendices.

[1] The distinction between automata and physical environments is discussed in Chapter 4.