MODELLING THE DRIVER FOR TRAFFIC FLOW SIMULATION

by

NORMAN ERIK STANG

B.Sc (Civil Engineering), The University of Saskatchewan, 1991

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

Department of Civil Engineering

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

April 1993

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

(Signature)

Department of Civil Engineering

The University of British Columbia
Vancouver, Canada

Date 12 October 1993

DE-6 (2/88)

ABSTRACT

For 60 years, engineers have modelled traffic flow for use in roadway analysis and design. A continuing problem with such models, though, is their inability to adequately capture the human element in the system. The human driver does not think and act in precise ways, making him difficult to model using conventional mathematical means. This research explored an alternative method of formulating driving models. It started from a psychological basis and used fuzzy logic. Fuzzy logic provided a systematic way of handling imprecision, and its constructs allowed a more intuitive model of driving, one more closely resembling the thinking and acting patterns of humans.

A new fuzzy logic driver model was developed. Its structure was based on a general psychological model of human information processing. It was designed so it could be programmed with specific driving behaviour by an end user. To demonstrate and validate it, the model was programmed for two lane rural highway driving and used in simulations of these facilities. Results were compared to field data and the Highway Capacity Manual and proved favourable.

The result of the work was a computer based road simulation toolbox containing the new fuzzy logic driver model. The toolbox provides a user with the ability to construct his own road networks, driver types, and vehicle types. With these, he can simulate traffic and examine both isolated incidents and overall performance measures.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENT

This research owes its beginning to Dr. Gordon Sparks of the University of Saskatchewan. He originated its concept, and supported and supervised its early stages. I am very grateful to him. For the contents and completion of the research, I also owe a big debt of gratitude to my thesis supervisor, Dr. Frank Navin of the University of British Columbia. His guidance, patience and advice were much appreciated. He focused the research on improvements to the driver model, ensuring that advances were made where they were most needed. He also suggested the use of fuzzy logic. Both the research and I benefitted greatly from his direction.

Along the way, the ideas of other noteworthy people helped shape the final product. I would like to thank Doug Kaweski, Mark Nickeson and Paul deLeur for their reviews of the software package and helpful suggestions for its improvement. I would like to thank Roy Klymchuk, Lynne McInally and Mike MacNabb for their encouragement and insight into the psychology of driving. Finally and most especially, I would like to thank Rob Thomson for his many helpful comments, for reading and evaluating this report, for fielding a thousand questions and for being my cohort in the world of object oriented programming.

I would like to dedicate this work to my parents Ike and Laila, and my brothers Howard and Gerald, whose love has sustained me in my endeavours over the years and continues to do so.

1 INTRODUCTION

It has been almost sixty years since B.D. Greenshields (1934) proposed his seminal mathematical model relating traffic speed, flow and density. Since that time, despite the efforts of many, a definitive model of traffic flow has remained elusive. The difficulty lies in the most variable and unpredictable system component: the automobile driver.

The driver is the only component of the system which is not a machine and does not operate like one. Traditionally, however, engineers deal with systems in which the overriding factors are mechanical and physical. So, in modelling traffic flow, they make the (often implicit) assumption that drivers can be modelled as simple mechanistic controllers. This assumption makes the system amenable to conventional analysis but sacrifices its human element. Rather than adopt this assumption for such convenience, the present research built a driver model starting from a psychological basis. This was a departure from history.

The history of traffic flow modelling has three distinct, though overlapping phases: macroscopic mathematical modelling, microscopic mathematical modelling, and computer simulation. In the first phase, analysts developed models which related aggregate traffic variables such as average speed, density and flow rates. Examples cited by Algea (1964) include the following:

> (1) Traffic movement is analogous to the motion of molecules in a gas as described by the kinetic theory of gases (Newell, 1955), (2) vehicles are elements in a queue (Haight, 1958), (3) the traffic problem is a mathematical problem (Pipes, 1953), (4) traffic flow is treated as the flow of a continuous fluid (Greenburg, 1959), and (5)

traffic dynamics are studied in terms of servomechanisms (Chandler, Herman and Montroll, 1958).

The problem with these models, as Algea noted, was that they ignored the importance of the driver. Consequently, there was always considerable scatter in the data, and the results were poor.

In the second phase, analysts made efforts to model traffic flow using its components: drivers and vehicles. They developed equations describing driver behaviour in terms of inter-vehicle interactions. From these microscopic models, they integrated to determine system performance. Efforts at this level, though, only demonstrated the complexity and stochastic nature of traffic flow. This suggested the problem was not amenable to conventional mathematical analysis.

The third phase of traffic flow modelling covered the last twenty years. With the development of digital computers, operations researchers created simulation techniques. Traffic flow analysts adopted these techniques at first to provide a means of more easily integrating microscopic models to the system level. Simulations also allowed them to include many more variables in the analysis. To date, however, traffic flow models have remained bound to conventional mathematical functions. This has not addressed the fundamental problem existing in the earlier models. *Human drivers do not conform to conventional mathematical description because they do not think or act in precise ways.*

People think and act in fuzzy ways. They base their actions on relative terms like slower, faster, higher, lower, comfortable and uncomfortable. For example, in choosing to pass a vehicle, a driver must think that the impeding vehicle is travelling *too slowly*. The oncoming vehicle must be *far enough away*. And the sight distance should be *adequate*. In passing the impeding vehicle, the driver attempts to travel *somewhat faster* than that vehicle. After passing, he resumes a *comfortable* speed. The driver does not incorporate exact values (ie. numbers) in any of these decisions. Given this, his actions are best described using fuzzy logic.

Fuzzy logic is imprecise mathematics. It allows one to model behaviours in a more intuitive manner than conventional mathematics. This is because fuzzy if-then relations have the appearance of reasons rather than causative relationships, and people think with reasons. Reasons and causes are distinguished as follows:

> Cause in the scientific sense is closely bound up with contingency, and a causal explanation consists of a specification of a set of antecedent events sufficient in the prevailing conditions for the occurrence of the caused event. The person whose behaviour is caused is, so to speak, at the mercy of antecedent events which cause it... Reasons ...are often stated as purposes: they refer to the ends to which the actions being explained are the means... It is manifest that most of what one does relates to rules and objectives. A Martian observing human behaviour would understand it much more quickly by discovering the rules which govern it rather than the causes which make it happen... The main reason why driving task analysis has been so singularly unsuccessful is that no one can say scientifically what the objectives are nor how to judge whether or not they are achieved (Taylor 1976).

Fuzzy logic is a mathematical description of human reasoning. The use of fuzzy logic requires one to look at driving from the point of view of a driver, rather than as an analyst of a bigger system. This is appropriate for two reasons. First, driver actions are for the most part voluntary and are the result more of internal reasoning than external causes. Drivers are more proactive than reactive.

Second, drivers make decisions based not on the state of the objective environment but on how they perceive it. They do not have perfect knowledge of the system they are operating in and consequently can make mistakes. Perceptual, attentional and experience limitations restrict their ability to learn about events and objects around them. Traffic flow models must account for this.

Driver models have four components which, in simulation, act sequentially. The components are environmental perception, anticipation, decision and action. Most other models effectively skip or oversimplify the first two components. The result is simulated drivers who do not make mistakes because they perceive the environment perfectly, and who are only reactive and not proactive in their interactions with other vehicles.

The focus of this research was the development of a more correct representation of the driver. Specifically, it improved the environmental perception and anticipation components of driver models. Environmental perception abilities have three factors: perception limits, attentional limits and driver experience. Perception limitations are physical. People do not have 360 degree vision and cannot see an infinite distance.

Attention, though, is the more important factor. Much of what people sense in their environment is missed because they have limited and divided attention. Driving is not a difficult task provided a person devotes enough attention to it. However, distractions exist, and it is when drivers divert too much attention elsewhere that mistakes are made. The current work used a stress index to model the broad level of attention paid to driving. Drivers in stressful driving situations pay more attention to driving than those in non-stressful situations. In the latter case, when stress drops below a certain distraction level, drivers pay no conscious attention to driving at all. Then, unconscious experience-driven processes take over. Provision is also made for the distribution of driving attention paid to specific components of the driving environment. This distribution is a function of the driver and the situation he is in.

The third factor in environmental perception is driver experience. Drivers who have been in more traffic situations can better interpret new ones. They not only know where and when to focus their attention but are better judges of distances, speeds, roadway curvatures, etcetera.

The second component of driver models slated for improvement was anticipation. Much of the reason that traffic works as well as it does is drivers' capacity to act on events before they occur. They can have negative reaction times. Models which do not have an anticipation component are often unstable and therefore unrealistic. As with environmental perception, drivers can err in their anticipation. And anticipation is even more a function of experience than perception.

The last two driver model components, decisions and actions, have been modelled in other work using conventional mathematics. By contrast, the present model uses fuzzy logic throughout, and this includes these two components. The advantage of this approach, as mentioned previously, is the more intuitive model structure. Instead of operating on a mathematical curve, simulated drivers carry out activities step by step, as would a real driver. Like the first two components, decisions and actions provide opportunities for driver error and provision is made for that.

After the driver model was developed, it was combined with vehicle and the road environment models to form an integrated traffic flow system. Think of this system as a toolkit for developing road ways for study. The system was programmed in a PC computer based application. The application is a shell with which a user can build his own road networks for simulation. He can also create his own drivers and vehicles to suit local populations. The driver prototype shell for creating drivers provides only a psychological basis for development of driver types. It does not assume any behaviourial structure beyond the sequential application of the four components: environmental perception, anticipation, decision and action. Within each of these, a user is free to not only calibrate but structure behaviour in any way. He can include and concentrate on whatever aspects of driving he wants. In this way, the shell does not pre-bias driver prototypes, and the user has complete freedom to build his model as he pleases.

This report is structured in the following way. Sections 2 through 4 describe the three components of the traffic flow system: the driver, the vehicle and the environment, respectively. Section 5 discusses the integration of these three into the overall system. To demonstrate the model, driver and vehicle prototypes were created for use on two lane rural highways. They were calibrated and validated using simulation runs on such facilities. Sections 6 and 7 address this. Section 8 considers application of the model, and section 9 provides a conclusion.

2 THE AUTOMOBILE DRIVER

The function of the automobile driver in the traffic flow system is to control his vehicle. Specific tasks in driving can be classified according to levels of performance in a hierarchy as follows:

> The vehicle-control subtasks low in the hierarchy are those observed in looking at the fine details of driving. They are referred to as micro-performance, or control. Macro-performance, or navigation, refers to the large behaviourial subtasks at the high end of the hierarchy. The remaining subtasks in between consist mainly of responding to roadway and traffic situations and are referred to as situational performance or guidance (Allen, Lunenfeld and Alexander 1966).

The present model dealt exclusively with situational performance tasks. The execution of each sub-task in the hierarchy affects tasks at lower performance levels. However, the present model divorced situational performance tasks from navigation and control ones. First, it assumed correct and invariable execution of lower tasks. These lower tasks mainly involve obtaining desired responses from drivers' vehicles. Second, the model requires users to specify traffic volumes, including turning volumes, thereby obviating the need for navigation.

The more critical of these assumptions was that of correct and invariable execution of micro-performance subtasks. The justification lies in that most drivers operate vehicles with which they are familiar. Even those who do not operate familiar vehicles quickly adapt to the ones they drive. In a study comparing driver performance as a function of different and unfamiliar vehicles, Koppa and Hayes (1976) determined the following:

> In general, drivers' steer (front wheel, not steering wheel) inputs were not significantly different between vehicles even though the vehicles had different steer ratios ...drivers normalized brakeline pressure to produce the desired deceleration which showed little difference between vehicles ...This suggests that drivers tend to normalize their inputs to achieve an acceptable level of lateral acceleration which is relatively independent of vehicle capability.

The basis of the driver model was a generalised psychological model of human behaviour. Section 2.1 describes the latter as it relates to driving. Section 2.2 details the driver model formulation, as implemented.

2.1 Psychological Foundations

2.1.1 Human Perception. As shown in Figure 2.1, a person driving a motor vehicle is modeled as a controller within a larger system (Algea 1964). The controller has four basic components: an input receiver, an information processor, a memory faculty and an output generator (Figure 2.2).

The senses - seeing, hearing, touching, tasting and smelling - receive information about the environment. The first stage of processing is data driven. It is the interpretation of these inputs, classifying and organising them for use in higher processing.

At the same time, conceptually driven processing uses contextual information to speed up recognition of the environment (Figure 2.3). A person forms expectations about new inputs

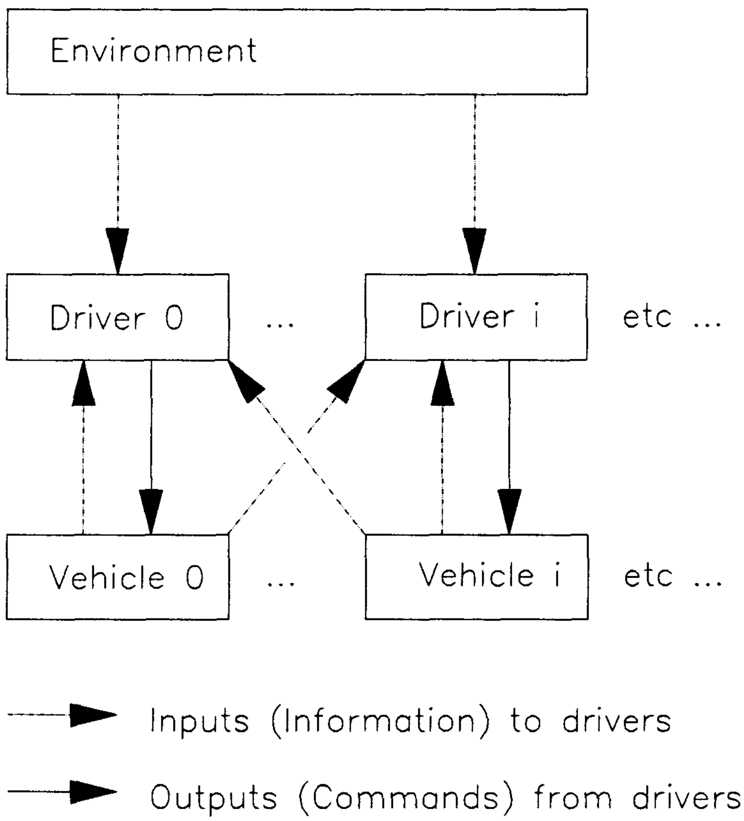


Figure 2.1 Traffic Flow System

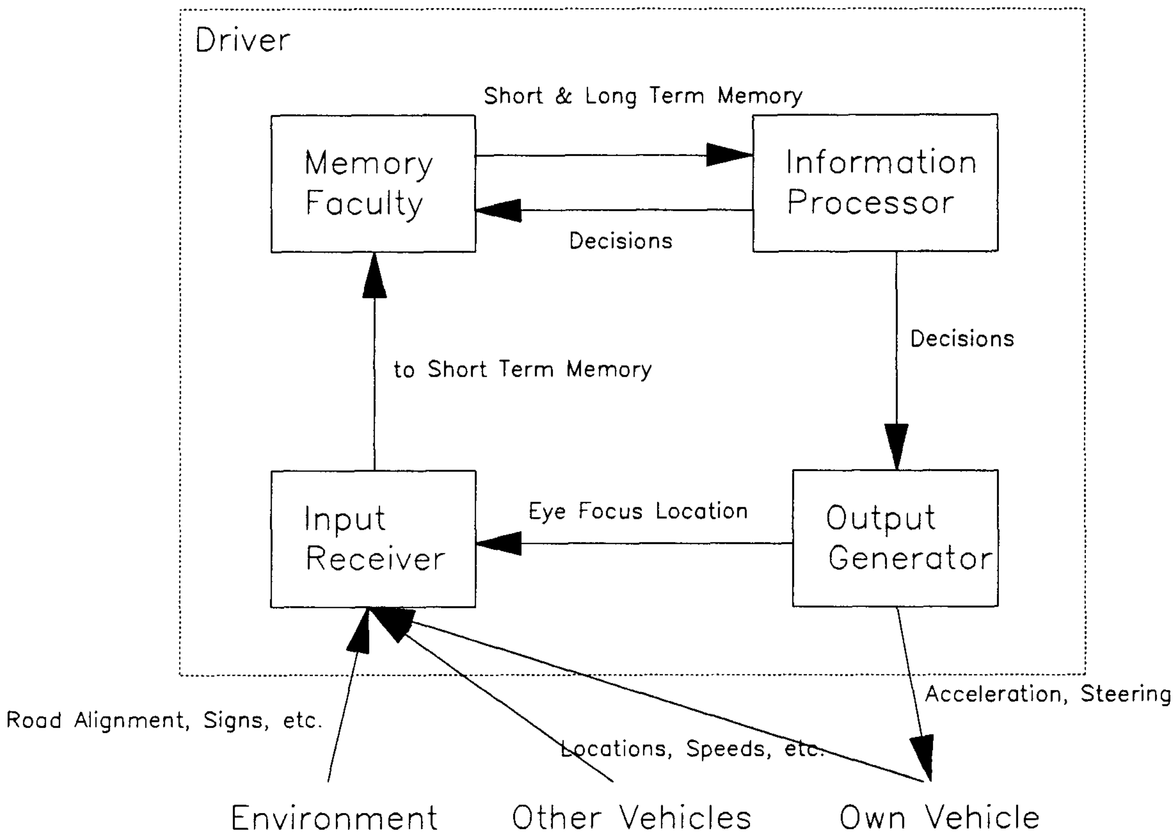


Figure 2.2 Driver Components

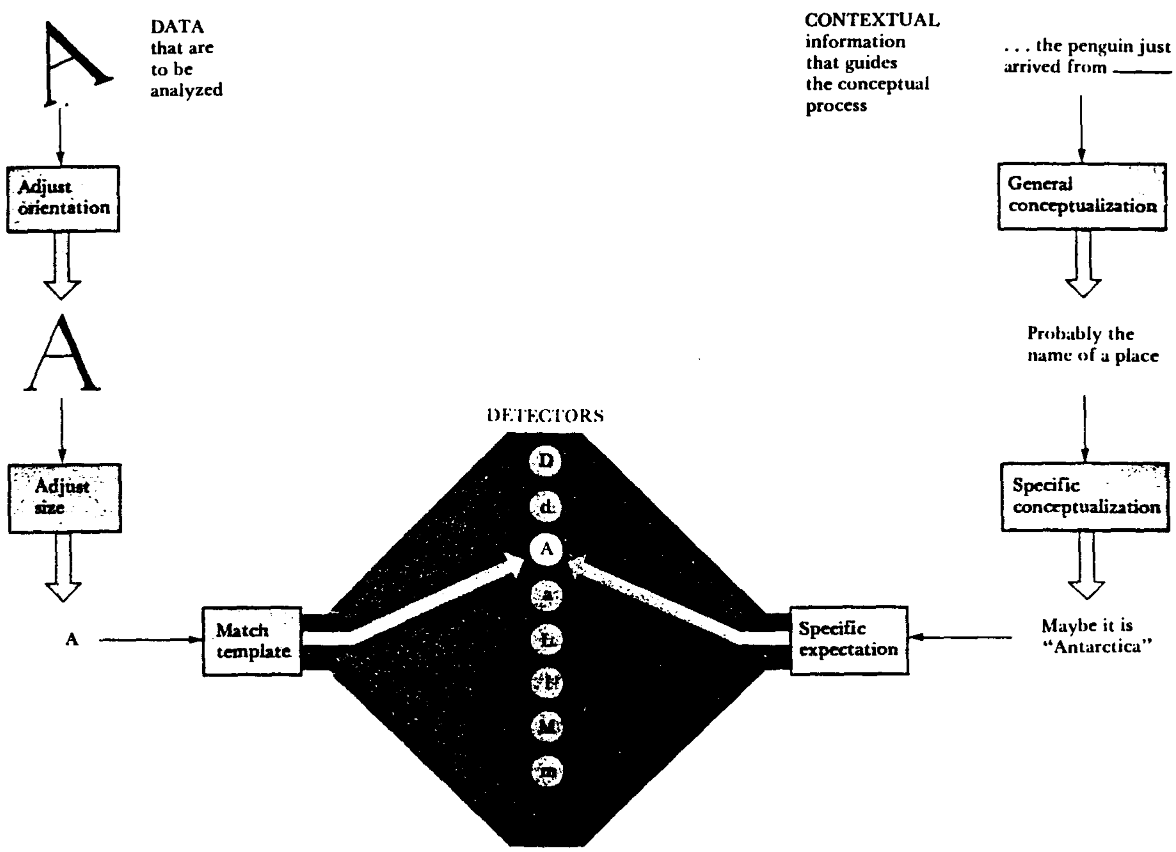


Figure 2.3 Data-Driven and Conceptually Driven Pattern Recognition

(Source: Lindsay and Norman 1977)

based on what he has received so far, and then seeks confirming evidence. It allows him to, for example, quickly read a word by taking in only the first several letters. Based on these letters and their context, he uses expectations to complete the word (sometimes incorrectly). Conceptually driven and data driven processing almost always occur together. Each contributes something to the interpretation of the environment (Lindsay and Norman 1977).

The literature on driving uses the term expectancy for conceptually driven processing.

> An important concept in publications on perception in driving is expectancy... The term indicates that driver behaviour is not only dependent on what the driver actually sees but also what he expects to see. Consequently the driver will scan his environment by body-, head- and eye-movements to 'test' his expectations. Therefore, in building a model that accounts for the visual strategies of drivers, the central cognitive processes that are responsible for 'expectations' must be incorporated (Wierda 1991).

Before considering interpretation processes further, however, one must look at the senses. For driving, the important senses are sight and hearing.

2.1.2 Sight. First, consider sight. The eyes receive over 90 percent of driving information (Hills 1980). The information received by the eyes is dependent on both light levels and where a person focuses. Changes in overall light intensity cause the eyes to alter their sensitivities. As light intensities decrease, sensitivity to low intensity light increases (Algea 1964). This does not happen instantaneously, however. Depending on the initial light intensity, a person requires up to thirty minutes to reach close to maximum dark adaptation. The process in the other direction also takes some time. Moreover, initial light adaptation begins slowly. For this reason, the bright lights of oncoming cars at night do not severely affect dark adaptation.

The importance of focus location stems from there being two types of light detection cells on the retina: rods and cones. Rods are more prevalent in the periphery of the eye and are very sensitive to light intensities, less so to colour. Cones concentrate in the fovea, the central focused section of eye. They are more responsive to colour and less to light intensities. The most detail-discriminating part of the fovea extends to only 1.5 degrees from the centre of focus. The fovea can extend, though, to about 15 degrees from centre at the sacrifice of some detail. Peripheral vision, which detects motion, extends 210 degrees horizontally (Wierda 1991).

During driving, a person focuses on successive objects in the environment, taking in details that he misses for objects in the periphery. Peripheral vision serves to attract attention (ie. move the focus) to objects which suddenly appear or change their state. An example is vehicles turning or slowing down.

2.1.3Hearing. Though yielding less information than sight, hearing is important to a driver. This is primarily in drawing attention to things outside his present field of view. For example, hearing can alert a person to a passing vehicle or an ambulance trying to get through an intersection. The attention-getting capability of sounds in the environment is a function of perceived loudness and the human ability to localise them.

The physical signals received by the ear do not exactly correspond to the sensation of sound one experiences. The pitch and frequency of physical sound does not have a one-to-one correspondence with experienced loudness and intensity. Loudness is a function of both the pitch and the frequency of physical signals received by the ear. Mid-frequency signals sound louder than both low and high frequency signals of the same pitch. Sounds can mask each other depending on relative frequency and intensity. In order for one sound to mask another, it must have a higher intensity, especially if the frequencies are different. Lower frequency sounds are more likely to mask higher frequency sounds than vice versa.

2.1.4Pattern Recognition. After the senses obtain information from the environment, it must be synthesized into meaningful patterns. One way to describe pattern recognition is using a metaphor developed by Lindsay and Norman (1977) called the Pandemonium system. This system consists of a succession of imaginary 'demons', which are analogous to different neural patterns in the brain. Demons works on sensory patterns, each performing a different job in analyzing them (Figure 2.4). The first set of demons, the image demons, record the initial image of the external signal. Next, feature demons extract particular characteristics in the patterns. These include, for example, certain types of lines, angles, and curves.

Cognitive demons watch the responses of the feature demons, each responsible for recognizing a more complicated and meaningful pattern. When a cognitive demon sees an appropriate feature, it begins yelling. The more features it finds, the louder it yells. Finally,

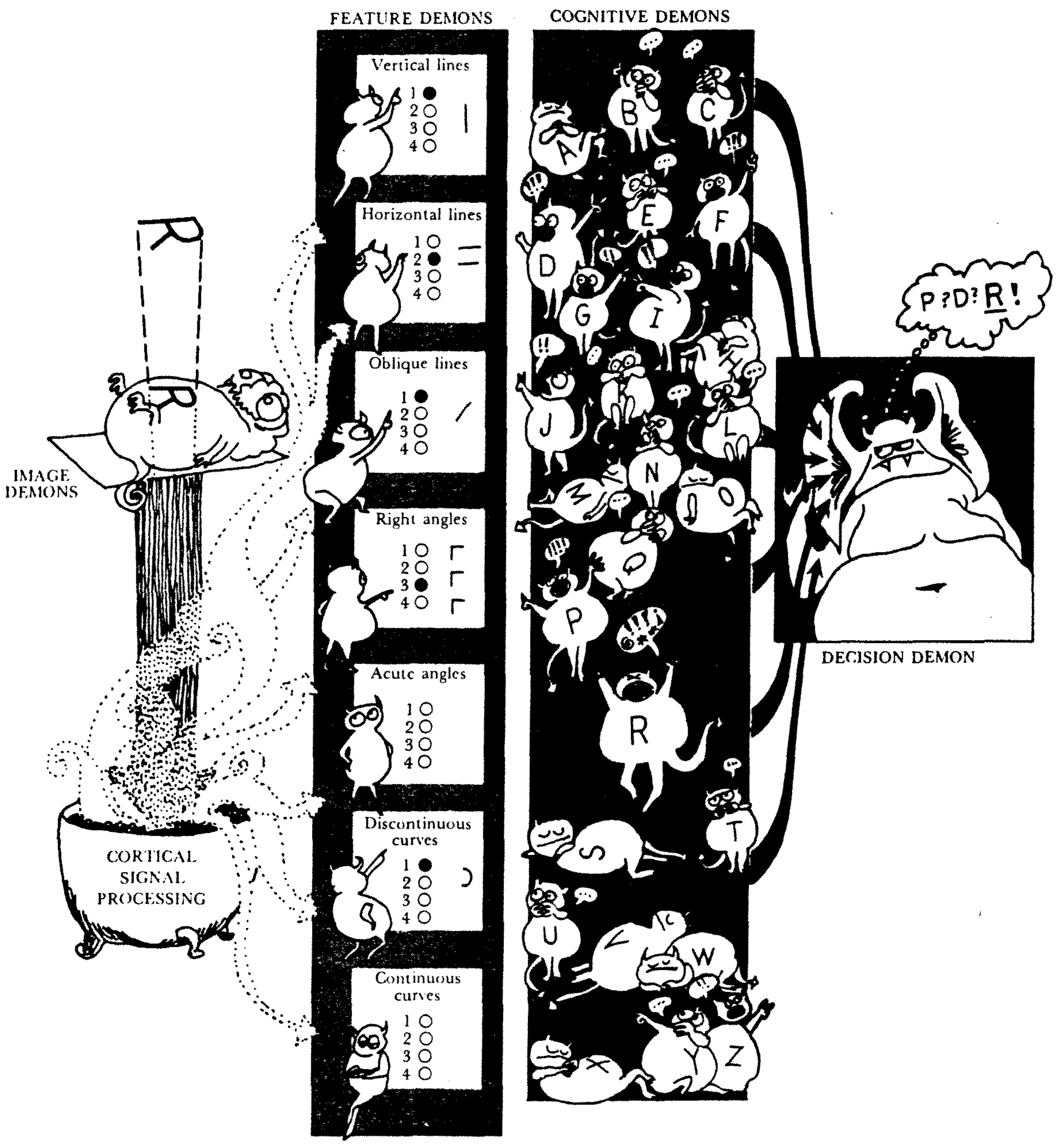


Figure 2.4 Pandemonium Data-Driven Processing

(Source: Lindsay and Norman 1977)

a decision demon listens to the pandemonium produced by the cognitive demons. It selects the pattern most likely occurring in the environment according to the cognitive demon yelling loudest. Experiments with reading and listening to speech have shown this to be a good model. People most often confuse letters that look like each other when reading, and syllables that sound like each other when listening.

As described so far, the Pandemonium model accounts for data driven processing. Adapting the model for conceptually driven processing requires the addition of specialist demons for context and expectations (Figure 2.5). The processing of specialist demons is not sequential but still interrelated. There is a need for a means of communication among them.

The central communication process is symbolized by a blackboard to which each demon has access. Cognitive demons involved in environment interpretation write their results to the board. Specialist demons watch the board for information they can analyze. As soon as information applicable to a particular demon's speciality is put onto the blackboard, it gets to work. When each specialist demon finishes its own specialized task, it writes the result on the blackboard. Some other demon can then use the result. In this way, individual demons (corresponding to specific neural patterns) need know nothing about the activities of the other demons. The total job still gets done.

People have limited capacity to take in information. When driving a vehicle, they are bombarded with a lot of information from the environment, most of which they do not

TAKE THIS **SENTENCE** AS AN EXAMPLE.

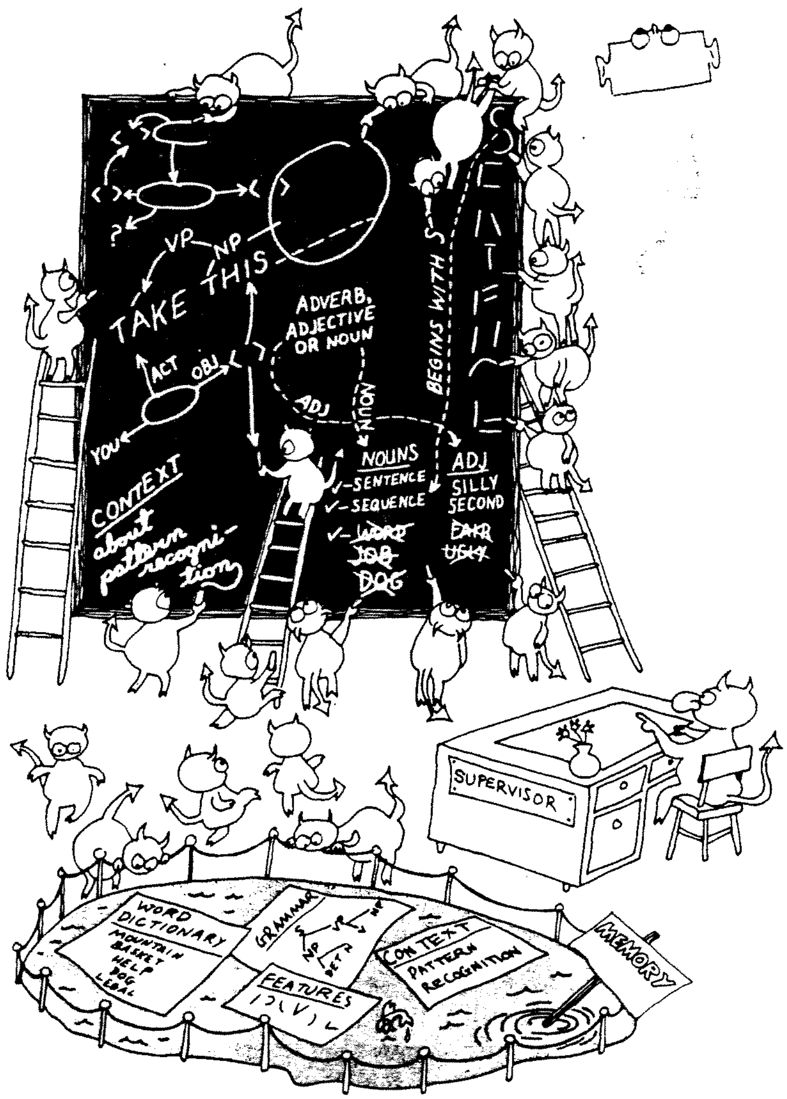


Figure 2.5 Pandemonium
Conceptually-Driven Processing

(Source: Lindsay and Norman 1977)

notice. The model considers this limitation. In the pandemonium model, there is a finite number of demons and fixed size blackboard. Any demon working on one analysis cannot be simultaneously working on another. And the blackboard will only allow room for a limited number of analyses. As described later, the size of the blackboard represents the limits of short term memory and sensory information storage.

In order to avoid conflicts and carry out analysis in promising directions, another specialist demon called the supervisor coordinates and directs activity generated around the blackboard. This demon's task is to direct the other demons based on what is on the blackboard. Because the blackboard limits it, the supervisor demon is as fallible as the rest.

2.1.5Attention. People can do more than one task at one time. They are not single channel processors (Lindsay and Norman 1977). This is because the mind operates at several levels of consciousness simultaneously. In short, people are able to walk and chew gum at the same time. This is even if they are not always consciously aware of one or the other. A single channel processor would have to alternate between chewing and taking a step. It could not do them at the same time.

The ability to perform multiple tasks at one time improves with the differences among tasks. For example, it is easy (quick) to read aloud a passage of text tapping one's finger after each noun. It is more difficult to read the text and say 'yes' after each noun because there is some conflict in processing. The first task of reading the text aloud is too similar to the that of saying 'yes'. Both require speech. In the pandemonium model, each is trying to use the same specialist demons. This explains why people can listen to the radio or a passenger while driving. The two are quite different and separable tasks so people divide their attention without much difficulty.

Overall processing is limited by the processing capacity noted previously. The number of individual tasks is not. In the pandemonium model, the supervisor is able to divide the blackboard among tasks. It assigns some specialist demons to each one (Figure 2.6). This models the division of attention one pays to different tasks. Driving is like other activities in that it can overload people with tasks to perform. When listening to two different conversations, people remember only the sound of voices and notable words from the conversation given least attention. Such words are notable for their loudness or meaning (eg. one's name). This suggests that people can combine well learned driving activities with another novel one. Attempting more than one novel activity at a time, however, causes difficulties.

People performing more than one activity at a time, divide their attention in an apparently rational way. In some situations they do this according to the expected value of information received from each of various sources. They pay more attention to the most promising source. In other situations, people pay more attention to more difficult tasks than to easier ones. Sometimes, a person seems to favour a particular sense such as vision over hearing. There is also evidence that to some degree, people control the division of attention

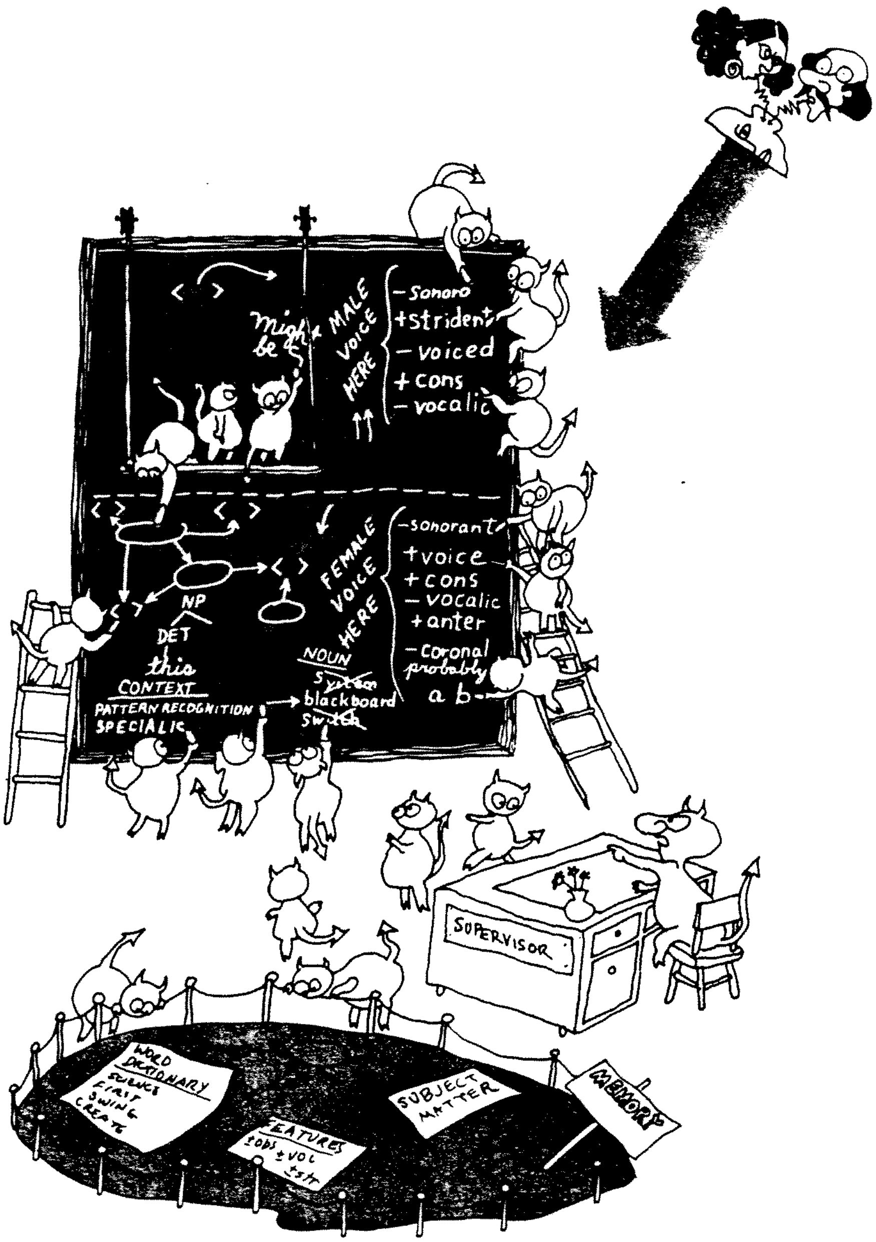


Figure 2.6 Pandemonium Division of Attention

(Source: Lindsay and Norman 1977)

unconsciously. There are some stimuli, such as hearing one's name, which people assign permanently high weights. These stimuli 'attract attention' regardless of the intentions of the person involved. Another example, pertinent to driving, is relative movement, especially if seen peripherally. People instinctively turn toward it (Shulman and Fisher 1972).

People are able to perform multiple tasks at the same time. There exists, however, a mental processing capacity limit which encompasses both perceptual and higher order processing. As demands on higher order processing increase and people reach capacity, attentional and perceptual functions are adversely affected. Everyone periodically immerses themselves deep in thought, oblivious to the world. Suddenly, they 'snap out of it' or 'wake up'. Such daydreaming can take away from one's ability to perceive the immediate environment. This occurs only when capacity limits are being pushed, however. When there is excess capacity, perception and other mental activities operate independently (Shulman and Greenberg 1971).

Attention is not only a conscious phenomenon. People take in and analyze more information than that which they consciously deal with. They see many objects in the environment and analyze them at a low level. They either discard the objects as unimportant or transfer them to short term memory. In the first case, people never consciously acknowledge the objects. In the second, they deem them important enough for conscious consideration.

> The factors that determine whether an element of information in iconic memory will be transferred to short-term memory, and therefore attended to, are the sensory

> conspicuity of the element, its information content and the informational needs of the observer. The central processor making this choice of elements in iconic memory can be considered to be driven by both the incoming sensory data and the cognitive processes of the observer. The central processor has access to long-term memory and to cognitive processes related to the observer's task and information from these sources will bear on the strategy used to scan the transient contents of iconic memory and on the criteria for selection of particular elements of information contained in it for transfer to short-term memory (Hughes and Cole 1986).

Evidence of this behaviour lies in drivers' use of road signs. They consciously acknowledge only about one in ten road signs. Those that the driver sees he considers to be most important (Hughes and Cole 1986). This says there is a subconscious screening mechanism taking place. Though a driver is only consciously aware of the important signs, he must at some level see them all. This is so he can determine which of them are important. Section 6.3 considers this further.

Other evidence comes from peripheral vision. Drivers see objects of potential danger with greater frequency than any other objects. The other object may even be of brighter contrast or bigger size. Also, considerable visual clutter, especially in an urban environment, does not significantly affect driving abilities. Roadside advertising, for example, provides little distraction in the driving environment. Drivers examine advertising only using spare attentional capacity (Hills 1980).

Expectancy plays a key role in attention. Those objects in the environment which are commonplace and expected are least likely to attract conscious attention. Novelty, of any kind, attracts attention (Hughes and Cole 1986). A driver is as likely to pay attention to a new but irrelevant part of the scenery as to new driving danger. The driver requires conscious attention to resolve the importance of the new object. Only the next time will the subconscious be able to screen for its importance. It is for this reason that novice drivers notice more road signs than experienced drivers. Expectancies are a function of experience residing in long term memory.

At sites requiring greater than usual caution, signs and signals are sometimes overdesigned. This takes advantage of a driver's reaction to novelty. Unusual signs attract his attention where a normal sign would not (Hills 1980). Obviously, highway engineers must use such overdesigned signs and signals sparingly to have effect.

2.1.6 Short Term Memory. Psychologists divide memory into three distinct types: sensory information storage (or iconic memory), short term memory, and long term memory. The sensory information storage system maintains a current accurate and complete picture of the world for a short duration - perhaps 0.1 to 0.5 seconds. This system creates the shadow images one sees when waving a hand quickly in front of the face. The subconscious importance screening described in the last section uses sensory information storage.

Short term memory is the working memory of conscious and some unconscious mental activity. It holds a less complete picture of the world than iconic memory. Rather than actual sensory perceptions, it retains their immediate interpretation. For example, after seeing a vehicle a driver may remember its approximate location, speed and heading. He may also forget its (unimportant) colour, model and size. The capacity of short term memory is the 5 or 6 most recently interpreted items. It is not very sensitive to item length or size, though, as long as these are coherent wholes. For example, people are able to keep an almost equal number of random letters and random words in short term memory.

Unlike information stored briefly in sensory information storage, a person can maintain short term memory for an indefinite period by maintenance rehearsal. Maintenance rehearsal is the conscious repetition or use of items. There are two possible processes by which a person forgets information (ie. drops it from short term memory). The first is interference. Forgetting by interference assumes that short term memory can contain a limited number of items. To add new items, one must discard old ones. More realistically, items in memory may have a strength which is a function of the number of items added after them. Thus memory does not limit its items by number. A person can bring back fading items in short term memory before they are entirely forgotten. This is modeled as follows:

$$A = A_0 r^n \tag{2.1}$$

where A = memory item strength
A_0 = original memory item strength
r = interference factor (between 0 and 1)
n = number of new items added

The second process of forgetting is time decay. It assumes that for each instant of time, the strength of items in short term memory decay in strength by a certain factor. A model similar to the one above can be used:

$$A = A_0 s^t \tag{2.2}$$

where A $\quad$ = memory item strength
A_0 $\quad$ = original memory item strength
s $\quad$ = time decay factor (between 0 and 1)
t $\quad$ = elapsed time since memory item added

The experimentation required to decide which is the correct short term memory model is difficult and has proved inconclusive. The correct model is probably a combination of the two. Though items in memory fade such that they can eventually be only partially remembered, one often can reconstruct the memories by filling in missing parts. This is much like receiving a partial sensory input and piecing together the rest using conceptually driven processing.

2.1.7Long Term Memory. Long term memory is the most complex and important type of memory. Anything that is remembered for more than a few minutes resides in it. All learned activities, including driving, are part of long term memory. For practical purposes, its capacity is unlimited. In fact, its capacity would allow the retention of every detail of sensory experience throughout a lifetime. The problem would be to efficiently retrieve that information.

Entry and retrieval with short term memory is immediate and relatively effortless. Entry and retrieval with long term memory requires time and effort. The only reason it is possible at all is because of the structure of long term memory. Human memory contains an enormous variety of concepts, each with a supporting structure of relationships to other concepts.

It is primarily through this supporting structure that concepts take on meaning. Compare long term memory to a dictionary. A dictionary defines each word using other words. The difference is that definitions in a dictionary are ultimately circular. People ground long term memories using experienced events and sensory perceptions. The concept of 'dog', for example, references experiences one has had with dogs, images of them, memories of their actions. The ability to drive exists in memory as a collection of driving experiences not as a collection of rules.

Lindsay and Norman (1977) modeled long term memory as a semantic network of records (distinct items in memory). Each record has an array of pointers and references to other records. Figure 2.7 gives an example of this. The structure shows a network of ideas related to a tavern. Note the various types of interconnections (properties, classes and examples) and the pointer directions. Some memory records are prototypical. Thus, we can conceive of a prototypical dog with default properties: four legs, a tail, fur covered, about so big. At the same time, we can recognise examples of dogs which do not fit the prototype exactly. Objects in the driving environment are recognised using similar prototypes (Wierda 1991).

Driving experiences in long term memory are mostly events, however. This model of long term memory assumes that people remember events as centring around actions, which they may join with conditional relationships. Episodic and semantic memories exist in the same

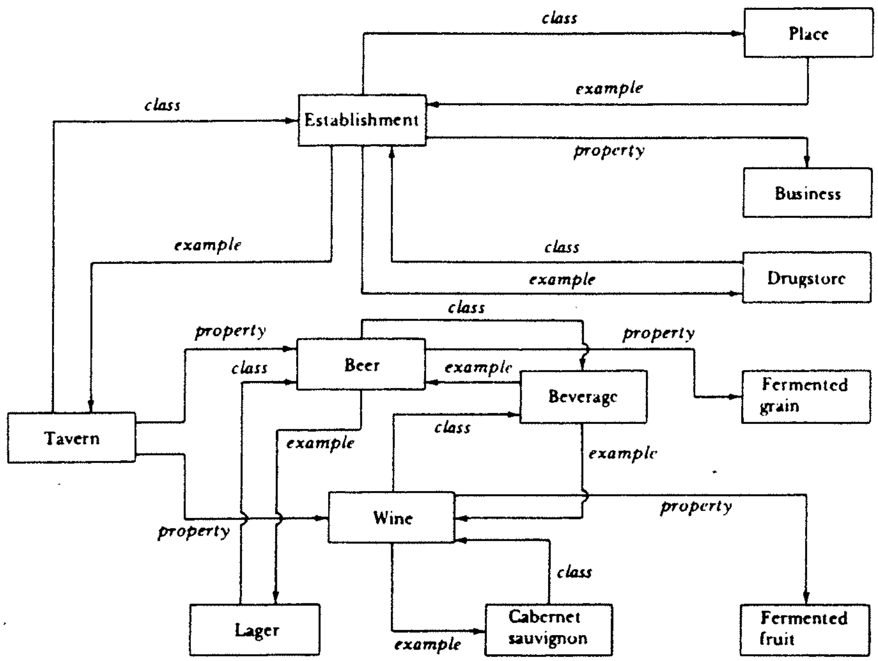


Figure 2.7 Long Term Memory Semantic Structure

(Source: Lindsay and Norman 1977)

structure, with pointers between records in both (Figure 2.8). As with semantic concepts, long term memory keeps prototypical events. People use them in planning an activity or imagining hypothetical situations. They also use them as the basis for acting (driving) in new situations which are similar to ones previously met.

The importance of long term memory at the situational performance level of driving is described by Allen, Lunenfeld and Alexander (1966):

> Performance at this level is a function of the driver's perception of a situation and his ability to respond in an appropriate manner. Therefore, the driver must have a store of a priori knowledge on which to base his control actions as will as an understanding of what the situation demands.

Using long term memory involves moving about its structure. One follows pointers and references and loads records into short term memory. Referring to the Pandemonium model, the supervisory demon directs searches of long term memory. He writes a window on the long term memory to the blackboard. Specialist demons then continuously update it until they retrieve the desired records.

Integration is the addition of records to long term memory in adults (ie. learning). In order for a person to permanently remember items in short term memory, he must integrate them into the structure of long term memory. This may involve adding a new part to the structure with many pointer and reference connections to the old part. It may also involve altering existing records and structures. The key to long term retention is the relationships established between new and existing knowledge. It is through these connections that a

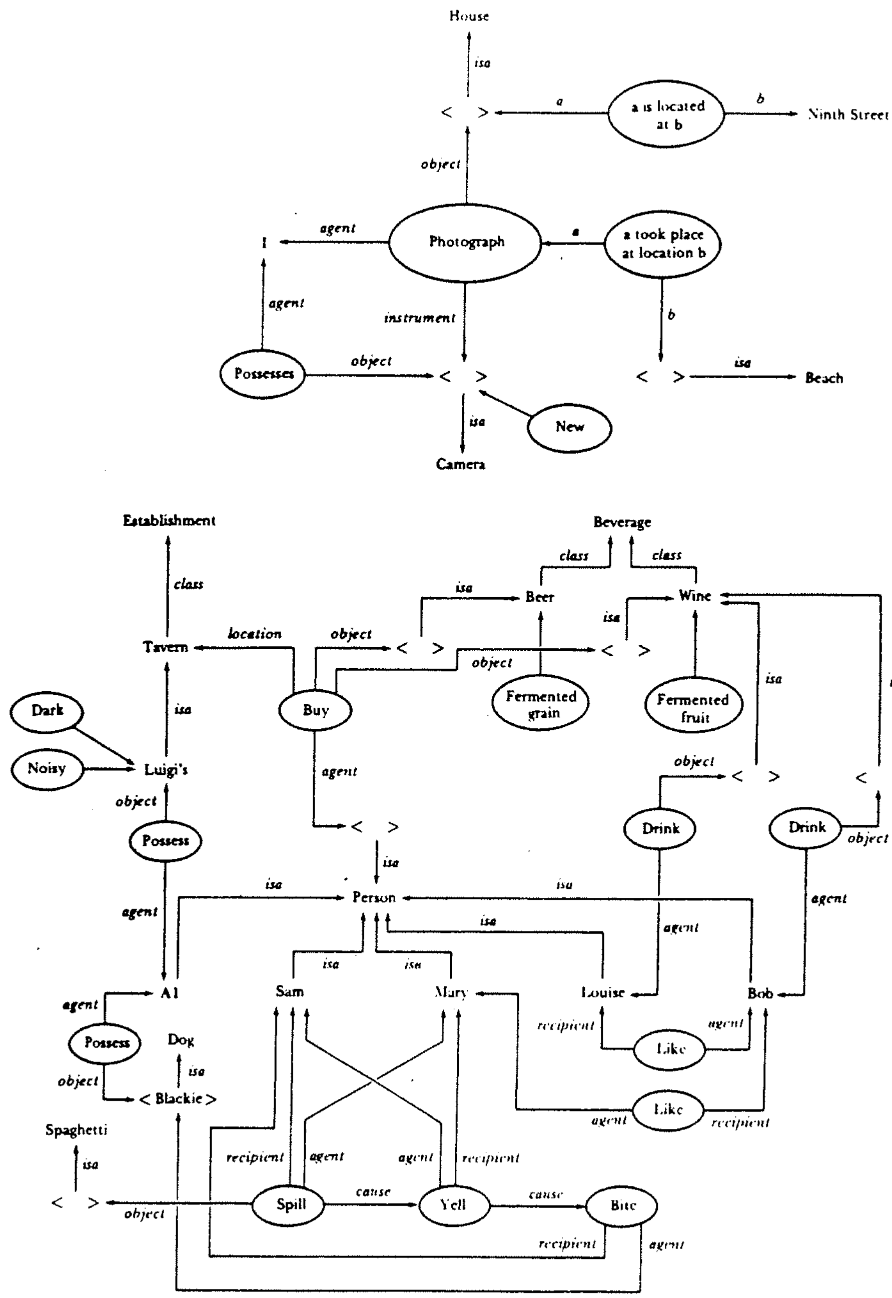


Figure 2.8 Long Term Memory Event Structure

(Source: Lindsay and Norman 1977)

person accesses the new memory once he forgets it from short term memory. The current work does not account for the learning process in driving, but considers it in the recommendations.

2.1.8 Decision Making. Psychological experiments with human decision making show that people follow a utility maximisation model. They measure alternatives in a decision combining a number of yardsticks specific to the decision. They then total and compare the results (Lindsay and Norman 1977).

People exhibit two different comparison techniques. Sometimes, they rank alternatives in each category of comparison and then average the rankings (perhaps using weights). At other times, they grade alternatives on cardinal scales and then compare average results. The latter technique requires more processing effort. Without the use of external aids (such as writing on paper), people use it only for simple decisions.

It is clear, however, that people use averaging in comparisons of alternatives. For example, suppose a person considers sincerity and cleanliness both good traits but the former somewhat better than the latter. He will have a higher opinion of a person described as sincere than one described as sincere and clean. This is because he averages sincerity and cleanliness giving a value less than sincerity alone.

What are the factors that make up utility? They are anything to which a person attaches importance in a particular decision making situation. They could include everything from monetary and emotional values to the desire to make a quick decision because of fatigue. These factors can change over time. A person can make different decisions about outwardly similar situations depending on his physical and emotional state. As a consequence, he may no longer understand the reasoning behind some past decisions. This suggests there is an internal variability in drivers in addition to the variability between drivers.

Something else to consider is people make decisions on the basis of whatever knowledge is available at one time. The available knowledge is whatever is present in short term memory. In analyzing a complicated problem, one can reach partial decisions guiding further analysis and then forget the reasons for the earlier decisions. External aids, of course, help overcome the limitations of short term memory. Also, people can use long term memory to reconstruct faded memories of earlier parts of a decision.

An important consideration in decision making is uncertainty. How do people handle it? Basically, they choose the alternative in a decision problem yielding the highest expected value of utility. However, they modify this value by an assessment of risk. A person who is risk averse in a given situation will choose an alternative with lower expected value and less risk (probability of non-occurrence). People are more risk averse when the stakes are high, both positively and negatively.

Risk in decision making is a subjective phenomenon. Utilities are functions of perceived or subjective probabilities. A person's perceptions and experiences bias probabilities used in decision making. People generalise from their limited experiences. If a person has had a good experience with the make of car he owns, for example, he probably considers it a good make. If he has had bad experiences with it, he likely considers it a bad make. Because of these two tendencies, the following generalities are common in human decision making:

1. People tend to overestimate the occurrence of events with low probability and underestimate the occurrence of events with high probability.

2. People tend to exhibit the gambler's fallacy, predicting that an event that has not occurred for a while is more likely to occur in the near future.

3. People tend to overestimate the true probability of events that are favourable to them and underestimate those that are unfavourable (Lindsay and Norman 1977).

In driving, these tendencies result in a particular pattern of risk taking. Drivers underestimate the risks in commonplace activities such as following too closely behind another car. They overestimate the risks in passing manoeuvres by choosing gaps in the opposing traffic stream larger than necessary. On the other hand, individual risk behaviour can be very inconsistent from one manoeuvre to another. This is because each driver's individual driving history is different.

2.1.9Thought Processes. Psychologists know little about thought processes and consciousness. Much of what follows is therefore their speculation. People think both consciously and unconsciously. Psychologists suspect that conscious processes are fundamental to intelligent choices, to learning and to guidance. Many subconscious processes can operate without conscious guidance for a while but must seek supervision and direction periodically.

Subconscious processing comes at a price. It uses up some of a person's processing capacity. Even though he is not consciously aware of it, his effectiveness at other tasks may be hampered.

The difference between conscious and subconscious thought in Lindsay and Norman's (1977) Pandemonium model is as follows. Conscious thought is the actions of demons directed by the supervisor. Subconscious thought occurs when demons continue their activities after the supervisor has diverted its attention to other demons doing other tasks. How does the processor (ie. the demons) know how to do its job? Here one must differentiate between tasks which are innate and those people learn. It is difficult to determine which tasks belong to which category. It is clear though that learning itself is an innate ability.

Both learned and innate tasks are programs or sequences of instructions in long term memory. Whenever the processor is in operation, it follows some particular program of instructions. Programs can search long term memory, alter memory structures and run or alter other programs. They can search short term memory for information or seek information in the environment using the senses. They can issue motor control commands to different parts of the body. Also, in performing tasks, programs often must deal with a trade-off between short term memory and other processing. This is because of the capacity limitations discussed before (Lindsay and Norman 1977).

This model of thought processes is analogous to another information processing device: a digital computer with parallel processing capabilities. Obviously, the human brain is much more complicated, flexible and powerful, but it also differs in another important aspect - it has direction. A digital computer takes its direction from its programmers and users. Human processors take direction from themselves.

The human processor, by this model, has three different control modes:

(a) conceptually guided control. Whenever the process supervisor attempts to satisfy its goals, it exerts conceptual guidance on the flow of processing. It directs other processors to perform their operations, probably by specifying the programs to follow. As the results of the operations become available, the supervisor evaluates them along with interpretations of newly arriving sensory inputs. It decides on further actions.

(b) program control. Program control exists at a level below conceptual control. Processes at this level simply follow the instructions of programs in long term memory.

(c) data-driven control. New sensory inputs exercise data-driven control of feature analysis and cognitive interpretation processes. However, a full analysis of incoming data requires consideration by the supervisory processor. There is usually some system which interrupts the activity of the supervisory processor and brings newly arriving data to its attention. Then it is up to the supervisory processor to decide what actions to take.

Why are there some tasks that a person is able to do best when he is not thinking about them? How come he can ride a bicycle but could never fully explain how he does it? People learn tasks like this to such a degree that they can perform them subconsciously. They firmly established the programs in long term memory through much repetition and practice. As drivers gain experience, more and more of the driving task becomes such an established program (Lindsay and Norman 1977).

There are some activities in which the guidance of conscious thought makes it more useful than subconscious thought. An example is the initial stages of problem solving. However, it is just this lack of guidance which makes the latter better for well learned activities. Subconscious thought is much quicker and more efficient than conscious thought because it requires no assessment to determine succeeding stages (instructions). Continuing with the digital computer analogy, conscious thought is like running a program in interpretive mode. The interpreter must check each instruction for errors and code it into machine language before execution. It does this each time someone runs the program. Subconscious thought, on the other hand, is like running a compiled program. A compiler has already checked the entire program and coded it into machine language. Someone only needs to execute it. Compiled programs run much faster than interpreted ones. Consider, for example, a person trying to consciously balance himself on a bicycle. By the time he figured it out, he would have fallen.

2.1.10 Social Interactions. In section 2.1.7, it was noted that people keep prototypes of objects and situations in long term memory. They also maintain prototypes of people and groups, commonly known as stereotypes. Until people get to know one another well, such prototypes can form the basis of social interaction.

People use prototypes for a number of reasons. Prototypes make for efficient memory use, processing and deductions. They allow one to respond quickly and efficiently to external events. Prototypes fill in missing essential details which would otherwise require time consuming effort to find out.

When meeting others for the first time, people fit them with prototypes based on their appearances and what they say and do. This is why first impressions are so important. It is usually an uncertain business, though. People can interpret a single action in a wide variety of ways in the absence of other supporting evidence. They do attempt interpretations, however.

Stereotypes are very important in driving because usually drivers do not know each other. They form initial expectations usually based on the type of vehicle being driven. For example, one may expect aggressive behaviour from the driver of a sports car. Such behaviour may be surprising from the driver of a station wagon loaded with children.

2.1.11 Stress and Emotion. Webster (1980) defines stress as "a physical, chemical or emotional factor that causes bodily or mental tension ..." Stress affects the cognitive processes discussed thus far. It has two basic causes: situations for which a person does not know how to respond, and situations which differ from what he predicted (Lindsay and Norman 1977).

The effects of stress on cognitive processes are as follows. People under stress narrow their perception of problems, concentrating on limited aspects of them. They also narrow the range of responses they consider. They select one and apply it over and over even if it is not at first successful. In a less stressful situation, they would likely consider more options if the first one failed. For example, people have died in building fires because they tried unsuccessfully to push open unlocked doors. The doors opened inward. Instead of rethinking the situation they just pushed harder when at first not successful. Parenthetically, this is why all fire doors are now required to open outward.

In driving, this introduces the idea of being pushed to panicking. As the stress level rises, drivers have less time to think of a proper response to a situation and react almost instinctively. For more experienced drivers, the instinctive response is more likely to be the correct one.

Stress arising from differences in predicted and actual situations comes from significant differences in sensory perceptions and long term memory. There is a conflict between data driven and conceptually driven processing. The model employs a cognitive comparator. The cognitive comparator continuously compares the two processing flows. It generates both neural activating responses and biochemical ones (hormones) as necessary (Figure 2.9).

2.2 Model Formulation

The psychological theory presented in section 2.1 provided the basis of the driver model. This section details the model formulation. In order to model a variety of 'driver types', many traffic flow simulations index driver behaviour to particular variables. These variables include desired speed, aggressiveness and gap acceptance thresholds. Psychological evidence suggests, however, that people are not so easily classified. The choice of such variables can bias a model. To avoid this problem, this model used the concept of defining a driver prototype by learned abilities and experiences. A driver prototype is a collection of perception abilities, predictive capabilities, and proactive and reactive behaviours and responses. It is no more than a collection of past perceptions and experiences which govern how it responds in new situations. These perceptions and experiences are long term memory prototypes as described in section 2.1.7.

Though indexed variables can be implicit in the driver prototype, this is entirely dependent

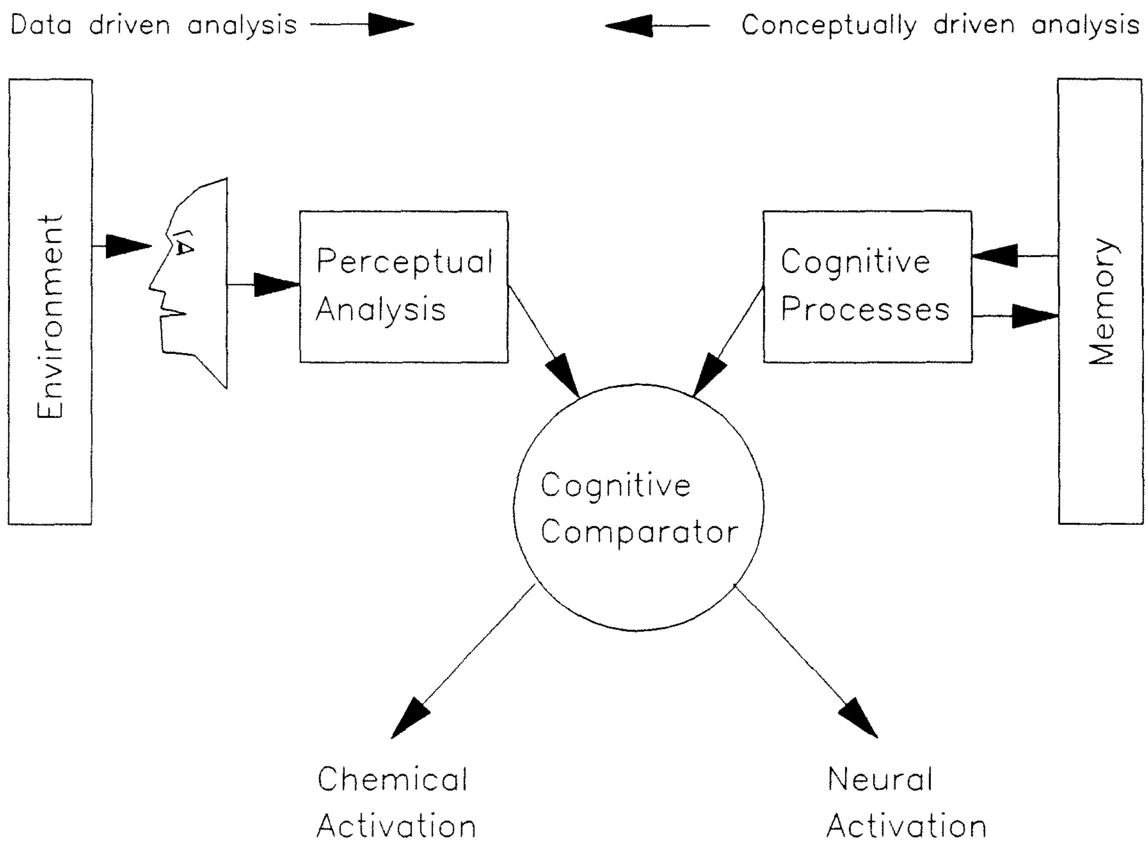


Figure 2.9 Stress and the Cognitive Comparator

(Source: Lindsay and Norman 1977)

M09*

program instructions in order from first to last each time step. The started instructions contribute to the acceleration, steering and focus location outputs. The processor combines the control sets from each of these instructions using union operators. This is the same as the predictive processor's combination of predictive control sets (Section 2.2.2).

In most programs, it is desirable that some instructions precede others in application. The driving processor must not apply later instructions until earlier ones are either in the started or finished state. The processor achieves such order using two progression flags at conditional instructions. The first of these flags requires that its instruction be in a started state before the processor executes any subsequent instructions. The second requires that its instruction be in a finished state before the processor executes any subsequent instructions.

Referring to Figure 2.16, which depicts a typical program instruction set, this works in the following way. The graphs for each instruction show the control set membership grade on the ordinate and time on the abscissa. The origin is the time that the program began execution. At time 0, all the instructions were in the not-started state. The driving processor evaluated the membership grade for instruction 1 and found it was less than the starting grade. This meant that the processor left the instruction in the not-started state and moved to instruction 2. It then evaluated the grade for this instruction and found it to be less than its starting value. It therefore left instruction 2 in the not-started state. The processor did not move to instruction 3, however, because instruction 2 had a start-required

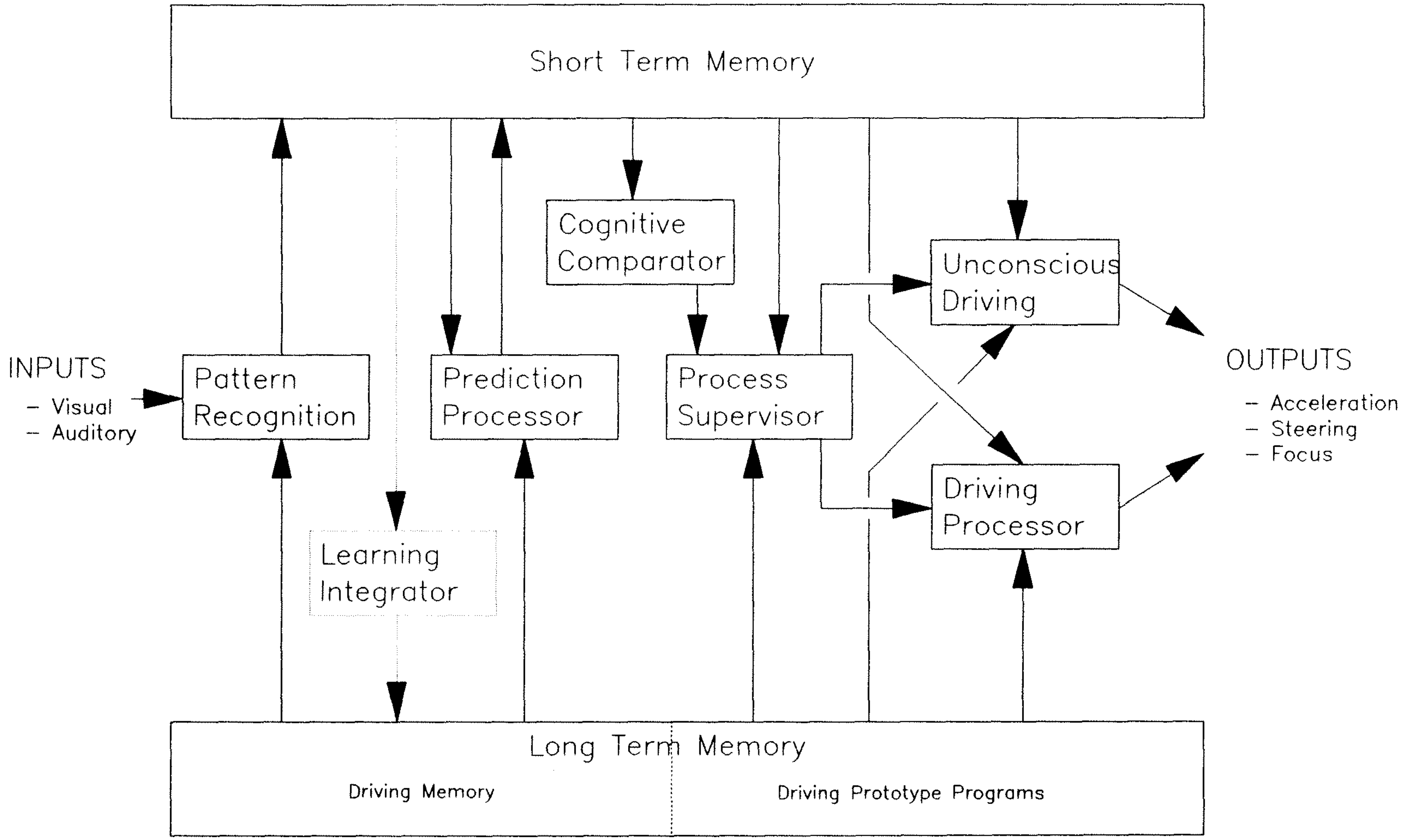


Figure 2.10 Skeleton Driver Model

on how a user defined and calibrated its experiences. Because the model is a generalised shell, the user has the ability to 'program' driver prototypes of his own. He does this by specifying experiences using fuzzy logic if-then rules[1]. The rules consist of relationships among a limited but fairly exhaustive set of variables. The variables describe the state of each element in the driving environment including the driver himself.

From the theory in section 2.1, a skeleton driver prototype structure provides the basis on which to define the fuzzy logic rules. Figure 2.10 shows this structure. It consists of a series of activity centres (small blocks) which perform different functions in the thought processes. These centres make use of both short and long term memory. Short term memory keeps the results of the processing done by the activity centres. The results are then used by other activity centres. Long term memory is a store of past driving experiences.

This structure is analogous to Lindsay and Norman's (1977) Pandemonium model described in section 2.1. Activity centres are like the demons performing limited and specialized tasks in isolation from other demons. Information passed between the centres via short term memory is like these demons writing their results to the Pandemonium blackboard. Like the fixed blackboard, the present model limits short term memory. The Pandemonium demons dipped into a permanent pool of long term memory. Similarly, activity centres make use of long term memory in their processing. The following sections describe the operations of

[1] Appendix A contains a brief introduction to fuzzy logic.

each activity centre, leading from sensory inputs to motor command outputs.

2.2.1 Pattern Recognition. Sensory inputs are transformed and screened before reaching the process supervisor. They pass first through the pattern recognition activity centre which transforms them from objective data to subjective impressions. The pattern recognition centre first sorts out which objects are part of the driver's 'active environment'. Objects are things external to the driver such as the road and other vehicles. The active environment is the set of elements the driver would see if he had 360° vision. It encompasses all objects which may affect the driver's actions at a particular time. Objects in the active environment include:

(a) the roadway and any intersections ahead extending to the limit of forward sight distance.

(b) vehicles ahead, behind, oncoming, or arriving on side roads at intersections ahead.

(c) Road signs and traffic lights ahead

The pattern recognition centre then reads the characteristics of each object and puts fuzzy representations of them into short term memory. Characteristics include relative positions, speeds and sign messages. Table 2.1 lists the readable characteristics of various types of objects. The fuzziness represents a person's varying ability to judge them. This depends on the object type, the particular characteristic, where the object is in the active environment and the degree of attention being paid to it.

Table 2.1 Object Characteristics

Object Type		Characteristic
Roads	(a)	distance to the start of a single alignment section*
	(b)	distance to the end of a single alignment section*
	(c)	length of a single alignment section*
	(d)	type of horizontal alignment (eg. right curve)
	(e)	horizontal radii
	(f)	spiral parameters
	(g)	vertical grades
	(h)	vertical radii
	(i)	superelevations
	(j)	median widths
	(k)	surface types (eg. asphalt)
	(l)	shoulder types (eg. paved)
	(m)	median types (eg. divided)
	(n)	passing allowances
Lanes	(a)	distance to the start of a single alignment section*
	(b)	distance to the end of a single alignment section*
	(c)	type of lane (eg. merging)
	(d)	movements allowed if at intersection (eg. right turn lane)
	(e)	width
Other Vehicles	(a)	distance to
	(b)	length
	(c)	frontal area
	(d)	speed
	(e)	relative speed
	(f)	lane position (eg. curb lane)
	(g)	turn signals
	(h)	brake lights
	(i)	headlights
	(j)	flashers
	(k)	position within lane
Traffic Lights	(a)	distance to
	(b)	colour of through light
	(c)	colour of left turn light

* A single alignment section is a sub-section of a road which has a constant horizontal and vertical alignment.

The conspicuity of an object is its ability to attract attention in a given situation. It is a function both of the physical qualities of the object itself and the priority given to it by the driver in question. Thus, conspicuity is variable. One can define two types of conspicuity:

> One is *attention conspicuity*, which is the capacity of an object to attract attention, and which might be measured by the probability of the object being noticed when the observer has not had his or her attention directed to its likely occurrence. The other is *search conspicuity*, which is the property of an object that enables it to be quickly and reliably located by search (Cole and Hughes 1984).

Attention conspicuity manifests itself in peripheral vision. Its measure is the maximum angle of eccentricity from a person's line of sight at which he notices an object at a single glance. The model therefore used this angle as a variable in its objective-to subjective transformation functions.

Search conspicuity, on the other hand, manifests itself in foveal vision. One must make the assumption that drivers always look directly at objects they are paying attention to. Thus, conspicuity functions are based on a driver's selected eye focus location. In this way the conspicuity of a particular object in the environment varies depending on where a driver is looking. In addition, given a particular driver and object, both attention and search conspicuities are functions of brightness relative to the background and angular size on the retina. They are not functions of absolute object brightness or size (Cole and Hughes 1984 and Hills 1980).

Each objective to subjective transformation, then, is a function of (a) the type of object in the environment; (b) the individual characteristic; (c) the solid angle subtended by the object at the driver's eye (visible angle); and (d) the angle between the line of sight to the object and the fovea (object angle) (Figure 2.11). The model did not include relative brightness as a variable on the assumption that for most objects average values suffice. Visual clutter is an important variable in conspicuity. This is difficult to measure and the model assumes that driver prototypes developed for particular types of driving, such as highway driving versus city driving, have the effects of average visual clutter inherent in them.

The visible angle is calculated as follows:

$$A_v = \frac{hw(v \cdot x) + hl(v \cdot y) + lw(v \cdot z)}{d} \tag{2.3}$$

where A_v = visible (solid) angle
l, w, h = object length, width and height
v, x, y, z = unit vectors (see Figure 2.11)

Some characteristics of objects in the fovea (ie. objects in focus) may not be fuzzy at all. Object characteristics become more fuzzy as objects become more peripheral in the field of view. Any recognition at all of some characteristics may require that an object be in the fovea or close to it. Reading a sign is an example. The model otherwise leaves such characteristics undefined (ie. unknown to the driver).

Auditory inputs, such as engine noise, tire noise and horns come from objects out of the visual field. The model treats them similar to visual inputs. It gives those characteristics

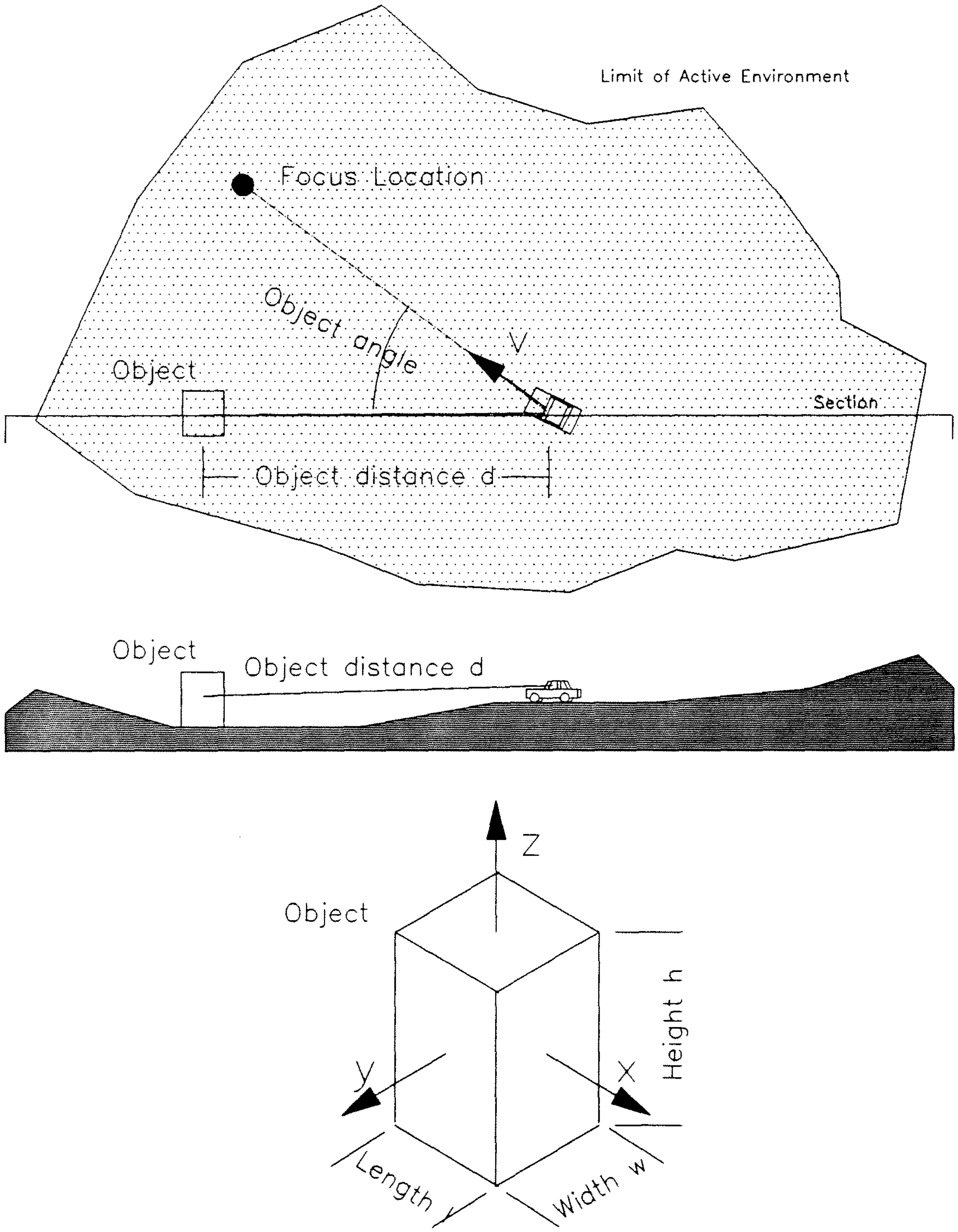


Figure 2 . 11 Active Environment

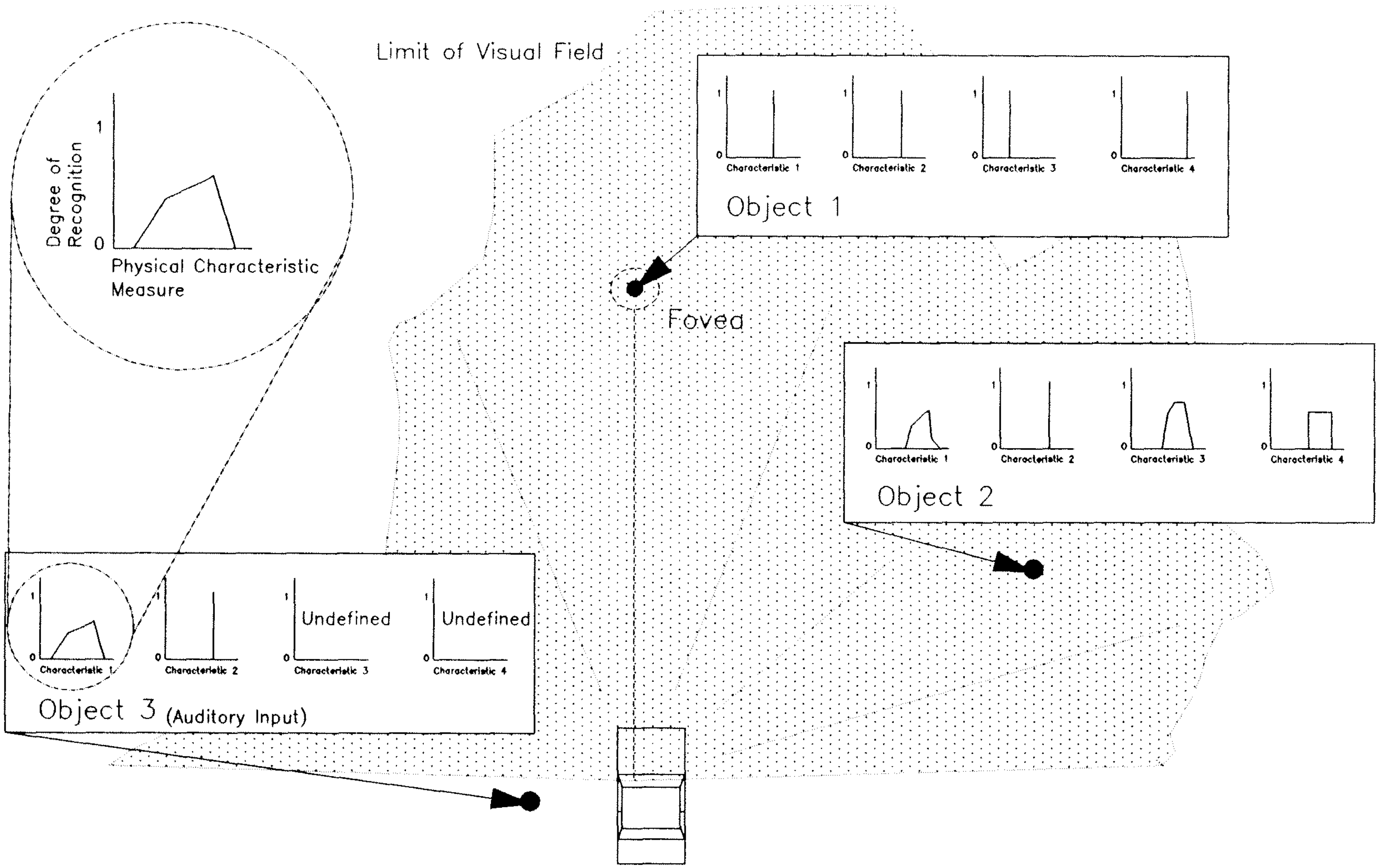


Figure 2 . 12 Pattern Recognition

which a person can deduce based on sound fuzzy values. Others it leaves undefined (See Figure 2.12).

The pattern recognition centre effectively generates a subjective view or 'snapshot' of the environment. The model keeps a limited record of the most recent of these snapshots in the short term memory. Stringing them together produces a motion picture of the environment consisting of time connected series of characteristics for each object (Figure 2.13). The model uses these 'object histories' in conceptual processing, decision making and program execution, all described subsequently.

Like for other objects in the active environment, a driver maintains the characteristics of his own vehicle in short term memory. These fuzzy characteristics are simply functions of the variables themselves. Table 2.2 lists them.

Table 2.2 Characteristics of a Driver's Own Vehicle

(a)	length	(g)	headlights
(b)	wheelbase	(h)	flashers
(c)	speed	(i)	position within lane
(d)	lane position (eg. curb lane)	(j)	heading relative to road
(e)	turn signals	(k)	sight distance
(f)	brake application		

2.2.2 Prediction Processor. Drivers take actions and decisions not only on the basis of the present and past active environment. They also use anticipated future conditions. The prediction processor activity centre allows for that. It takes the object histories created by

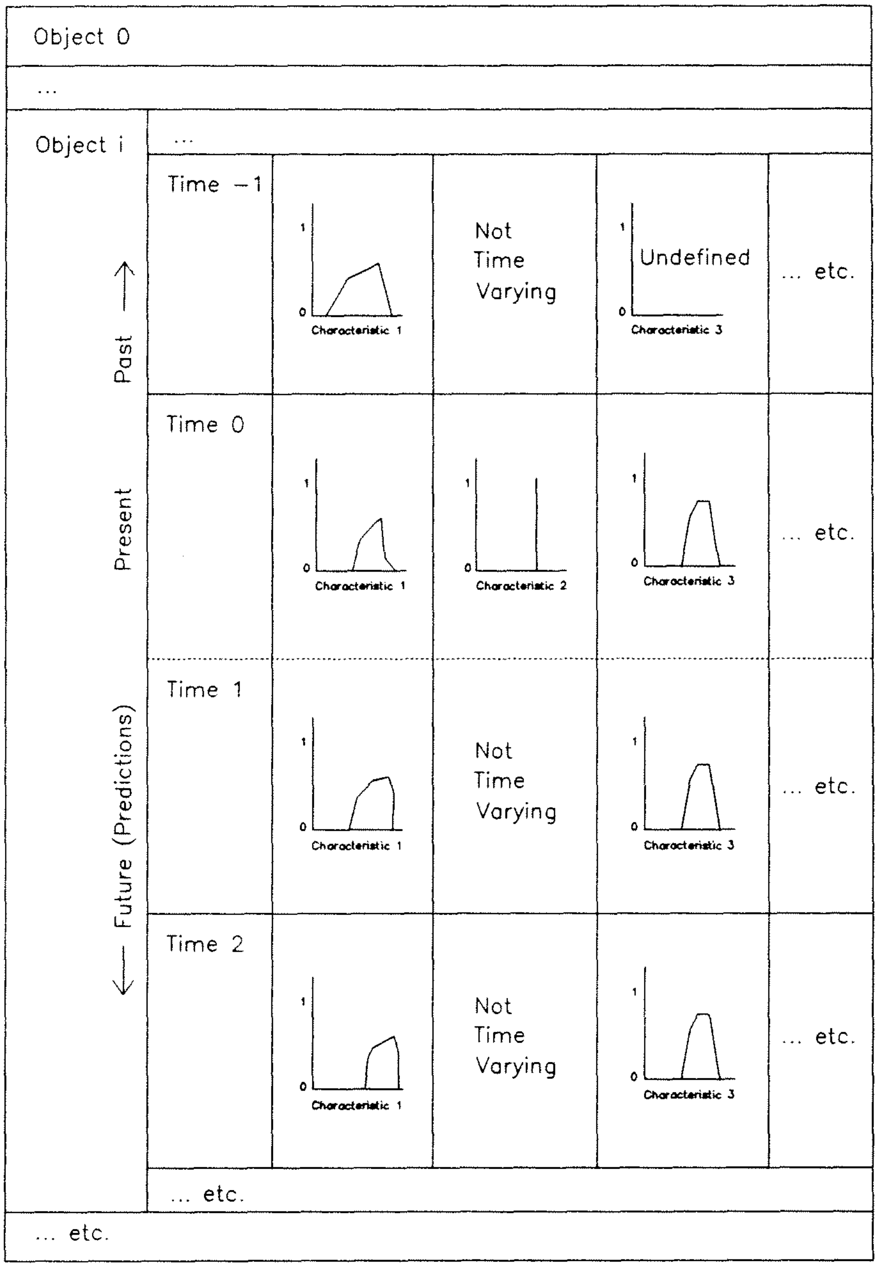


Figure 2 . 13 Structure of Short Term Memory

the pattern recognition activity centre and, using long term driving memory, forms expectations of future situations. The long term memory consists of a set of fuzzy relations. The model's allowing for decisions based both on predictions as well as past events makes the driver proactive as well as reactive. This is realistic.

The prediction processor makes predictions only for objects in the fovea, ie. objects under detailed consideration, and only for characteristics which are time varying. The driver's long term memory has, for each object type, a series of if-then fuzzy relations. The relations associate antecedent past and present object characteristics with consequent future ones (Figure 2.14).

Referencing Figure 2.15, prediction processing takes place in the following way. First, the prediction processor 'softens' each fuzzy predictive relation. It does this in accordance with how well the relation's antecedents match current conditions. It performs an intersection operation between each antecedent (relations can have multiple antecedents) and the corresponding characteristic in short term memory (Step 1 in Figure 2.15). The processor then performs a min-max operation across the resultant fuzzy sets yielding a single membership grade (h). It uses the membership grade to determine a control set. The grade represents the extent to which the antecedents matched existing conditions in the active environment as perceived by the driver.

In Step 2 in Figure 2.15, the prediction processor determines control sets for the relation by

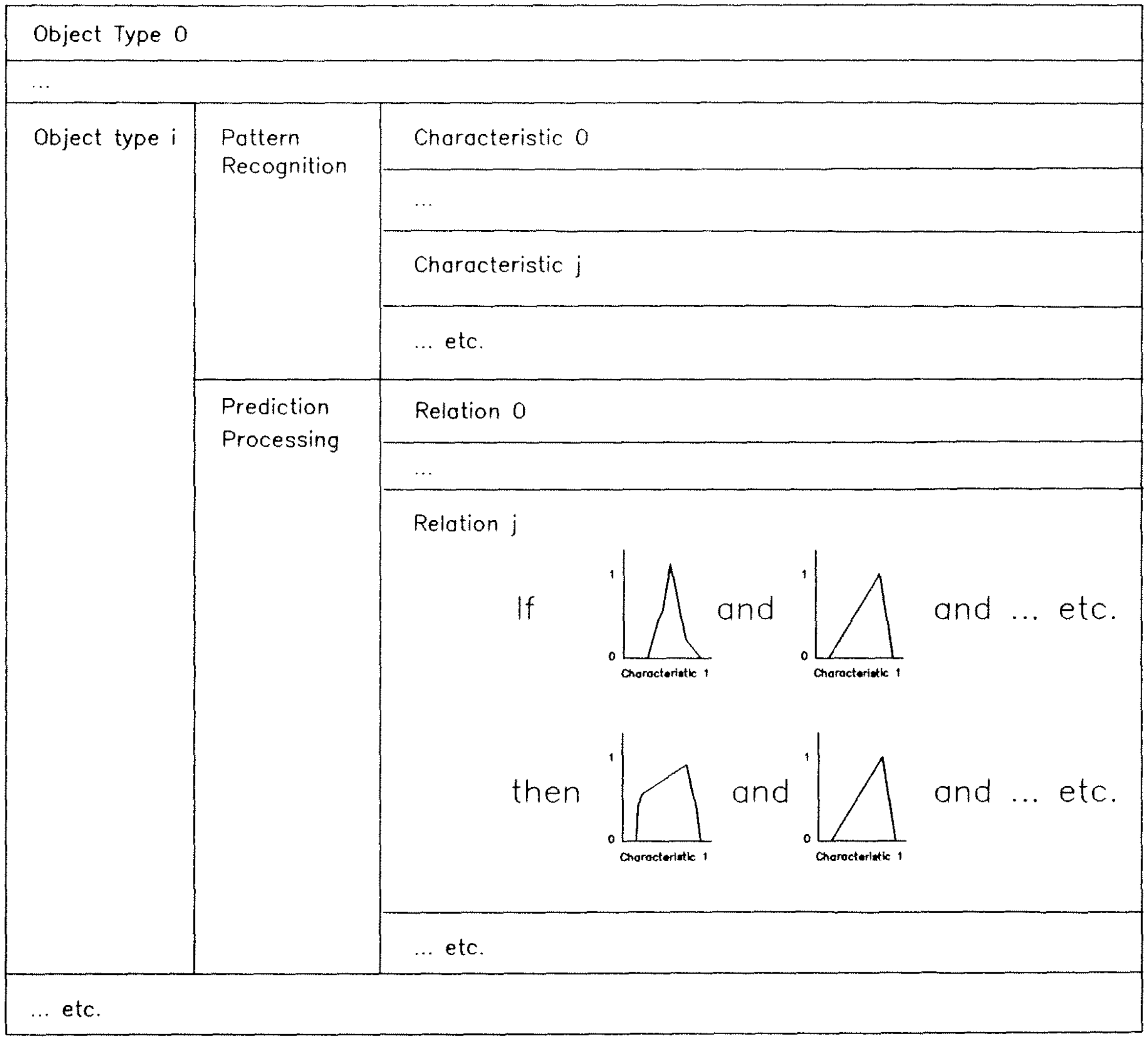


Figure 2. 14 Structure of Long Term Driving Memory

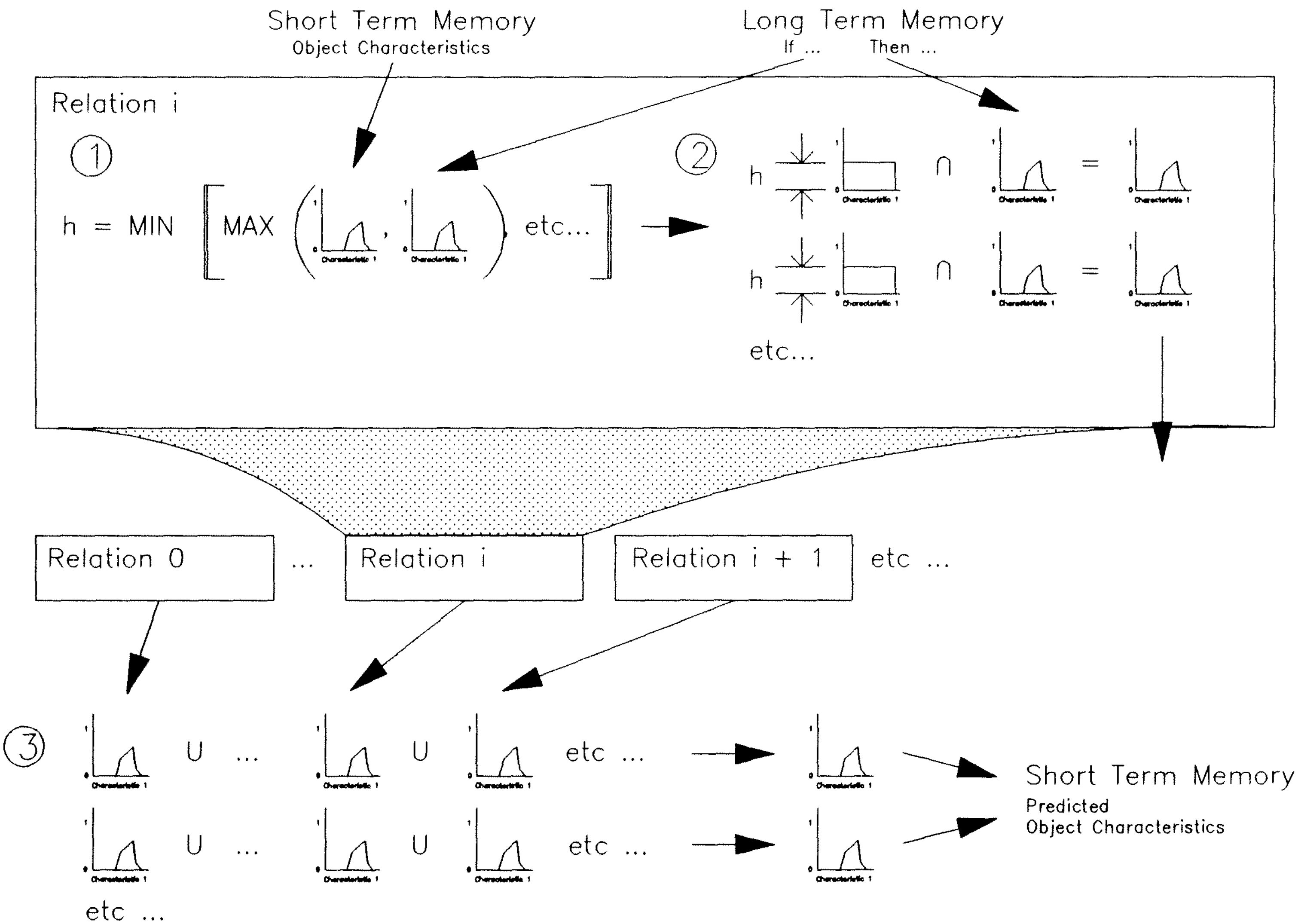


Figure 2.15 Prediction Processing

performing intersection operations with the consequents and a 'flat' set. The flat set has a uniform membership h.

The processor performs steps 1 and 2 for each predictive relation for the object type. This results in a number of control sets for the various predicted variables and future times. In Step 3 in Figure 2.15,the processor combines these sets for each predicted variable and time using union operators to yield single fuzzy predictions. Any characteristic of any object in the active environment can be used in the antecedents. This is regardless of the object for which predictions are being made. Table 2.3 lists the variables for which the processor can make predictions. Once the processor makes its predictions, it stores them in short term memory as shown in Figure 2.13. The predictions serve to extend the time histories discussed in section 2.2.2into the future and provide the basis for later decisions and actions.

Table 2.3 Time Varying Object Characteristics

Object Type		Characteristic
Other Vehicles	(a)	speed
	(b)	lane position (eg. curb lane)
	(c)	turn signals
	(d)	brake lights
	(e)	headlights
	(f)	flashers
	(g)	position within lane
Traffic Lights	(a)	colour of through light
	(b)	colour of left turn light

2.2.3 Cognitive Comparator. Section 2.1.11 introduced the idea of a cognitive comparator.

As noted there, one cause of stress is the prediction of situations in which a driver is not confident of his ability to respond. The second main cause of stress is situations which are unexpected (not predicted). To identify unexpected situations, the cognitive comparator compares past predictions with the present (subjective) active environment. This is the role of the cognitive comparator. In this model, it has a dual purpose. The first is to compare predicted and actual events to produce an overall stress level which affects processing elsewhere. The second is to bring peripheral objects to the attention of the process supervisor if they deviate dramatically from a steady state. In this way drivers notice sudden movements 'out of the corners of their eyes'. If an object does nothing unusual, it may not come to the conscious attention of a driver at all.

The cognitive comparator calculates stress levels using relative Hamming distances between predicted and actual object characteristics. It does this when time catches up to the former. These distances represent the degree to which the two sets of characteristics differ (ie. how wrong the predictions were) and are calculated as follows:

$$S_{oc} = \frac{1}{n}\sum_{i=1}^{n} |m_{Ci} - m_{Pi}| \qquad (2.3)$$

where
- S_{oc} = stress contribution of characteristic c of object o
- n = number of intervals across characteristic measure
- m_C = membership function of characteristic c of object o at time 0
- m_P = objects in focus: membership function of latest prediction of characteristic c of object o at time 0
  other objects: membership function of characteristic c of object o at time -1

The prediction processor does not make predictions for most objects in the active environment. For these, the cognitive comparator watches for marked fluctuations in characteristics using Hamming distances as just described. Distances greater than a certain threshold invite the driver to change his focus to an object. This invitation competes with those of other objects, though. It also competes with the object under focus. Because drivers can quickly grasp suddenly changing situations and bring predictions into line, stress caused by the cognitive comparator has a shorter life than that caused by unfamiliar or dangerous situations. Thus stress increases because of incorrect predictions or the sudden changes of peripheral objects count only in the time step in which they occur. Prolonged stress can occur, though, if the driver continues to make bad predictions. It can also occur if he faces an active environment changing too fast for him.

2.2.4 Process Supervisor. The collection of a driver's prototypical driving experiences is a series of programs. Each program is a list of instructions to effect a desired outcome. There are two types of programs: conscious and unconscious. Conscious programs are those considered for execution by the process supervisor, which represents a driver's conscious direction of his thoughts and actions. At each time step, the process supervisor checks a set of starting conditions for each program. It then determines the next course of action. The starting conditions consist of a set of fuzzy antecedents on object characteristics in the active environment. The process supervisor evaluates them using the same technique as described in section 2.2.2 for predictive relation antecedents (Step 1 in Figure 2.15). For each, it determines a membership grade with a value between 0 and 1. It compares the membership grades of the programs and puts the program with the highest one into effect.

For each program, model users can specify an active stress range. Outside of this range, the process supervisor does not consider the program for execution. This is so when stress is high, the process supervisor does not execute programs alleviating trivial problems. Also, in order for the process supervisor to consider a program for execution, the program's grade must be higher than the current stress level. This has the effect of narrowing options in high stress situations. Eventually the process supervisor executes a panic program (the one with highest maximum stress). This is to reflect the stress effects described in section 2.1.11.

Once a program is in progress and before its completion, the process supervisor can replace it with another. This requires that the starting grade of the latter exceed the starting grade which originally prompted the first program's execution. For example, if a passing program is being executed, it may be effectively aborted by a close oncoming vehicle. The process supervisor executes a higher stress response, ie. a new program to return to the right lane.

So far the process supervisor has concentrated solely on driving. However, it must also consider other matters which are a distraction from the driving task. Examples of distractions are thinking of other things or listening to the radio. Others include talking on a telephone, talking to passengers or daydreaming. The process supervisor makes use of a level of distraction between 0 and 1, the latter representing full distraction. Consider the level of distraction as a starting condition value for non-driving programs. The process supervisor compares it to starting values of driving programs and to stress levels as described previously. If the level of distraction is higher than the current stress level and higher than the starting condition values for all considered driving programs, the driver enters a distracted state.

Drivers in a distracted state operate their vehicles using an unconscious program. The program has a similar structure to the conscious ones but has no starting conditions. This division of conscious and unconscious driving reflects the ideas of section 2.1.9. While driving in a distracted state, prediction processing does not take place. Drivers are able only to react to the environment and not be proactive in it.

One of the causes of stress, as described in section 2.1.11,is a situation for which a driver is not confident in his ability to make the correct response. In the model, situations occur in which the stress level is higher than the distraction level. At the same time, however, none of the starting condition values of the conscious programs exceeds the stress. None of them provides a proper response in the situation. The driver, because he lacks a previous (prototypical) experience of the situation, is unsure of how to respond. This causes a rise in stress. The process supervisor narrows the range of programs it considers. If the starting grades do not rise sufficiently, it eventually executes the highest stress program (panic).

2.2.5Driving Processor. Once the process supervisor issues the order to execute a particular conscious program, the driving processor carries out the order.

Programs consist of a series of instructions, executed sequentially. The instructions produce acceleration, steering and changes in eye focus location for the driver. In the same way that programs can have an active stress range, each instruction can have an associated stress range within which it applies. This allows a user to model the effects of stress on the performance of different tasks. A program may have sets of duplicate instructions, each designed for a different stress range. The instruction sets perform the same task but with different dexterity on the part of the driver.

Instructions are of one of two types: conditional or unconditional. Conditional instructions have the form of fuzzy if-then relations, and unconditional ones are simply fuzzy output commands. During program execution, each conditional instruction is in one of three states: not started, started, or finished. Before program execution begins, all conditional instructions are in the not started state. Execution at each time step begins with the first non-finished instruction. If an instruction is conditional and in the not-started state, the driving processor evaluates its fuzzy relation antecedents. It uses the technique described in section 2.2.2. If the antecedents' control set membership grade exceeds a specified starting value for the instruction, the driving processor moves the instruction to the started state. In the started state, the processor solves the if-then relation below a specified stopping value. At this point, the processor moves the instruction to the finished state and no longer uses it.

The driving processor issues unconditional instructions each time step regardless. It applies program instructions in order from first to last each time step. The started instructions contribute to the acceleration, steering and focus location outputs. The processor combines the control sets from each of these instructions using union operators. This is the same as the predictive processor's combination of predictive control sets (Section 2.2.2).

In most programs, it is desirable that some instructions precede others in application. The driving processor must not apply later instructions until earlier ones are either in the started or finished state. The processor achieves such order using two progression flags at conditional instructions. The first of these flags requires that its instruction be in a started state before the processor executes any subsequent instructions. The second requires that its instruction be in a finished state before the processor executes any subsequent instructions.

Referring to Figure 2.16, which depicts a typical program instruction set, this works in the following way. The graphs for each instruction show the control set membership grade on the ordinate and time on the abscissa. The origin is the time that the program began execution. At time 0, all the instructions were in the not-started state. The driving processor evaluated the membership grade for instruction 1 and found it was less than the starting grade. This meant that the processor left the instruction in the not-started state and moved to instruction 2. It then evaluated the grade for this instruction and found it to be less than its starting value. It therefore left instruction 2 in the not-started state. The processor did not move to instruction 3, however, because instruction 2 had a start-required

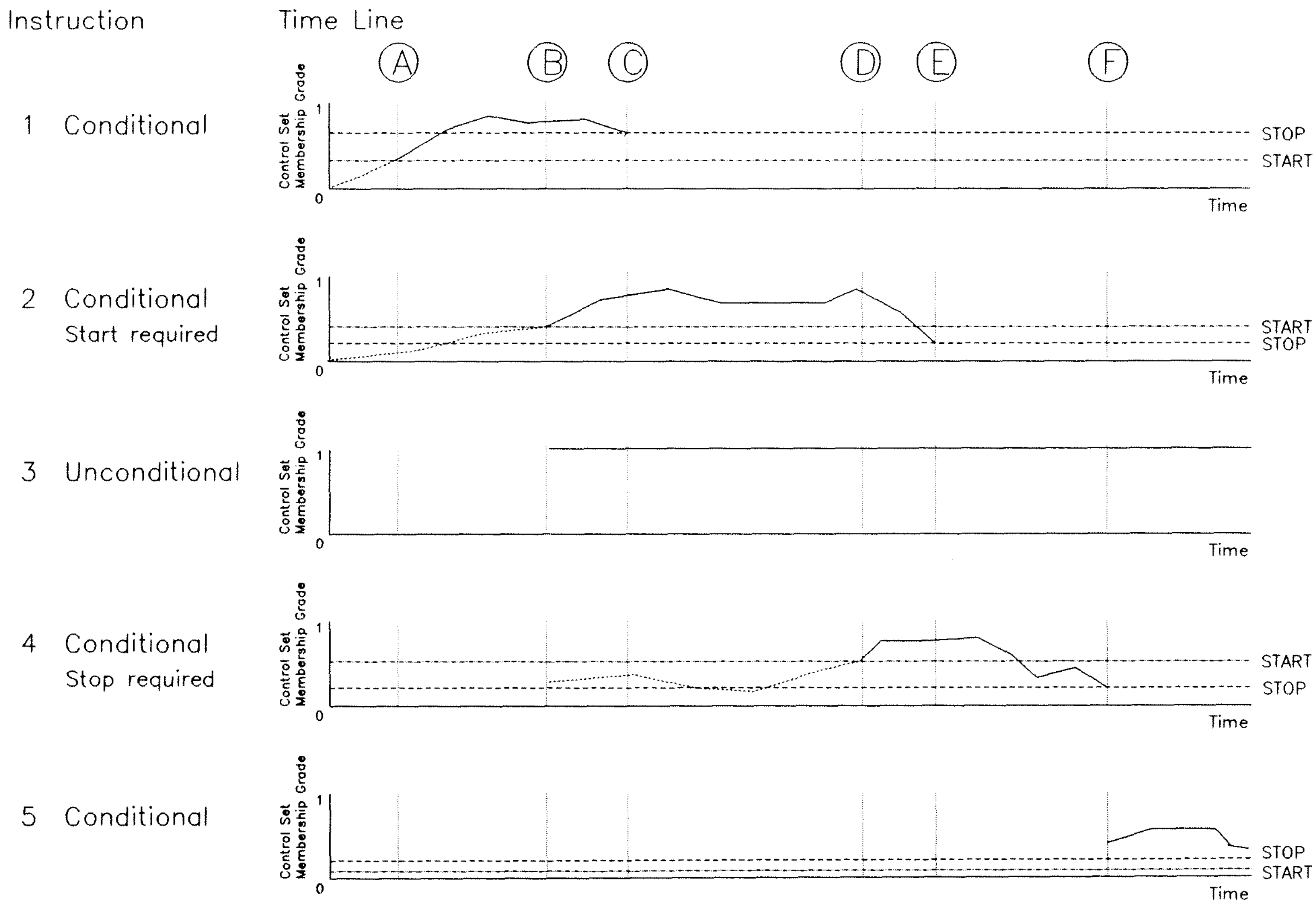


Figure 2 . 16 Conscious Program Progression

flag. Processing of the program stopped until the next time step when it began again with instruction 1.

This loop through the first two instructions continued until time A. At this point, the membership grade of the first instruction exceeded its starting value. The instruction entered the started state and began contributing to the acceleration, steering and focus location outputs. Instruction 2 did not exceed its starting value until time B. At this point, the first two instructions were in the started state and contributing to the outputs. Also, instruction 2 satisfied its start-required flag, and the processor moved to instruction 3. This instruction was unconditional and therefore began contributing immediately to the outputs. The membership grade of instruction 4 was less than its starting value. The driving processor therefore left the instruction in its initial not-started state. Processing for the time step stopped at instruction 4 because of its stop-required flag.

The driving processor looped through the first four instructions until time C. At this point, the membership grade of instruction 1 dropped below its stopping value. The driving processor moved the instruction to the finished state. It then looped through instructions 2 through 4 at each time step. Between time C and time D, instructions 2 and 3 contributed to the outputs. After time D, instruction 4 moved into the started state and also began contributing. The processor did not check instruction 5 until time F when instruction 4 entered the finished state satisfying its stop-required flag. After time F, instructions 3 and 5 determined the outputs.

Each instruction has a mental processing time associated with it. This delays its first application whether it be conditional or unconditional. The time delay for unconditional instructions is applied the first time execution reaches the instruction. The driving processor applies the time delay for conditional instructions when the instruction's state moves to started. Note that the processing time does not include the time required for muscle and vehicle response. The model assumes these to be functions only of stress for each of the three outputs regardless of the mental processing preceding them.

2.2.6 Unconscious Driving. The process supervisor directs conscious processing. However, depending on the level of a driver's experience, he may be performed much of driving unconsciously. In this driving model, unconscious driving controls those outputs which are not being consciously controlled. For example, during a passing manoeuvre, a conscious program may adjust acceleration to achieve a particular speed. At the same time, the unconscious program steers to stay within the lane. The unconscious program is only inactive when a driver is consciously controlling all the outputs.

The unconscious processor works much like the driving processor described in section 2.2.5 except that the processor does not use state and progression flags. It applies each instruction, depending on stress levels, at each time step.

3 VEHICLES

Figure 2.1 showed the automobile driver as a controller in a traffic flow system. The driver takes informative inputs and issues command outputs. The driver issues all his external command outputs to one other system component: his vehicle. The two main types of outputs to the vehicle are acceleration and steering angle. The model assumes the driver to be able to change these instantaneously, subject to reaction time. It assumes that drivers know how to make their vehicles respond to their wishes. There was no need to model systems internal to the vehicle. This means that they issue acceleration commands directly to the drive wheels and not via the accelerator, engine and transmission. Similarly, the model assumes drivers are able to set the angles of their vehicle's front tires directly. This cuts out the need to model the steering system.

The present work does not model systems internal to the vehicle in series with driver commands. It does, however, check the limitations of these systems to be sure that vehicles respond only within their capabilities. The following sections discuss the limitations on acceleration, deceleration and steering.

3.1 Acceleration

In most driving situations, one of the performance characteristics of most interest to drivers is vehicle speed. The model did not use speed as a control variable, though, because a driver cannot instantaneously change his vehicle's speed. To achieve a different speed, he must accelerate or decelerate which, given vehicle capabilities, takes time. One of these vehicle capabilities is the maximum acceleration it can achieve at a particular time. If a driver commands a higher acceleration than the maximum, his vehicle responds with the maximum value. Maximum acceleration for a particular vehicle type is a function of its present speed and the grade of the road.

In an application similar to the present one, St. John and Kobett (1978) incorporated simplified acceleration capability models for both passenger cars and commercial vehicles. The present work reviewed a more complicated (and accurate) model for passenger cars (Lucas 1986). Figure 3.1 shows a comparison of acceleration capability curves for a small passenger car using the two methods. The Lucas model was rejected because the parameters required for its calibration are not commonly available for vehicles in the population. Because most drivers operate their vehicles well within their limits most of the time, the St. John and Kobett models were deemed sufficiently accurate. Sections 3.1.1 and 3.1.2 describe the latter.

3.1.1 Passenger Cars. This model applies to passenger cars, pickups, recreational vehicles, light trucks up to 2-ton rating, and any of these vehicles pulling trailers. Maximum acceleration for these types of vehicles is approximately a linear function of speed according to the following equation:

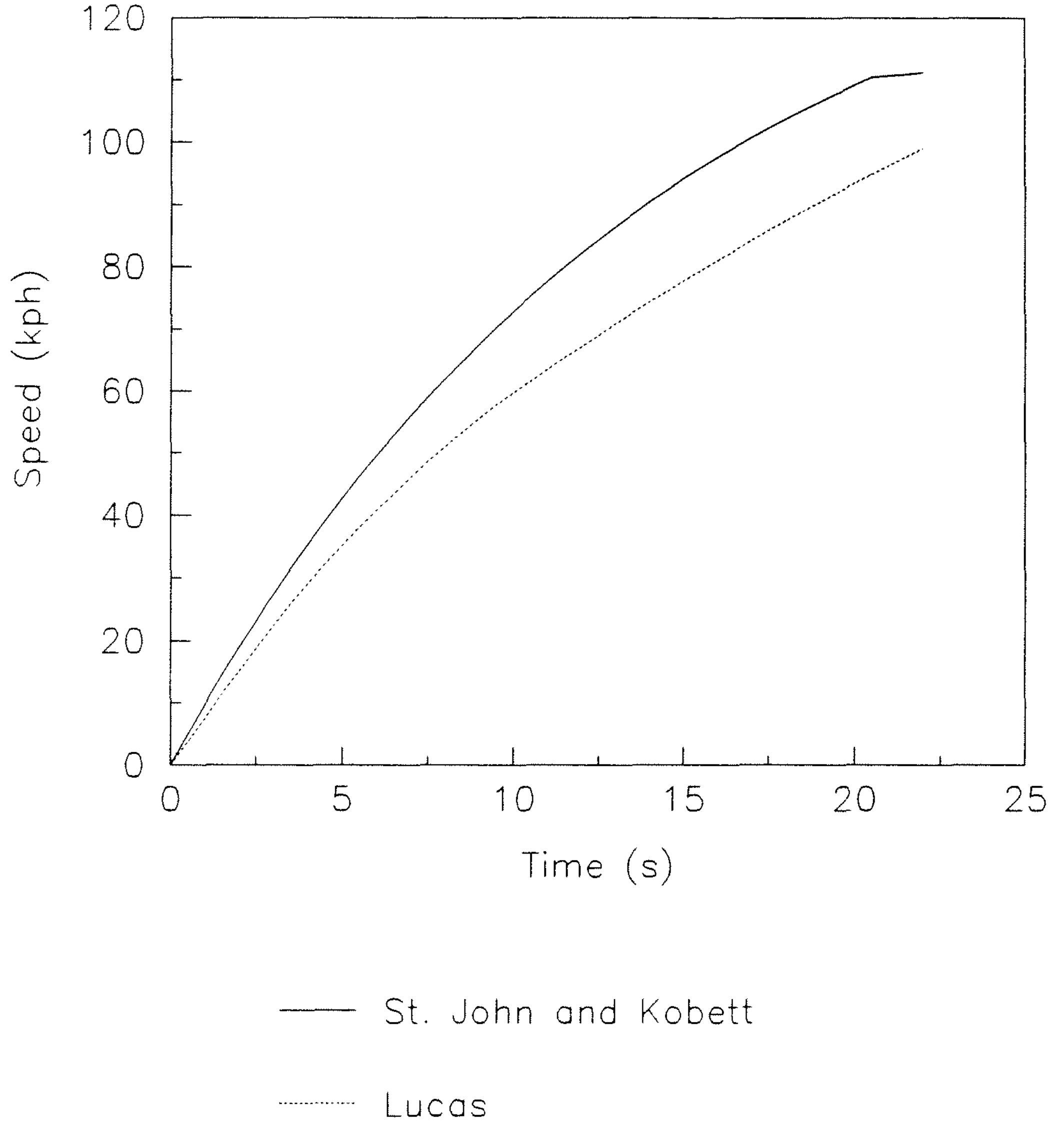


Figure 3 . 1 Time to Speed Curve Comparison

$$a = a_0(1-\frac{V}{V_m}) - g\sin G \tag{3.1}$$

where a = acceleration capability at speed V on a zero grade
a_0 = maximum acceleration at a speed near zero
V_m = a pseudo-maximum speed. This is the maximum speed indicated by the linear relation between acceleration and speed when data are fitted in the normal operating range. This maximum occurs at a = 0
g = gravitational acceleration
G = grade (upgrade positive)

The maximum acceleration a_0 and pseudo-maximum speed V_m are functions of vehicle characteristics. The maximum acceleration can be calculated as follows:

$$a_0 = -0.2055 + \frac{40.06 r_1 r_2 bhp(1 - 0.0013E)}{Wg}$$

for motor homes and large sport vans (3.2a)

$$a_0 = 0.26 + \frac{42.23 r_1 r_2 bhp(1 - 0.0013E)}{Wg}$$

for other vehicles (3.2b)

where a_0 = maximum acceleration (m/s^2)
r_1 = transmission first gear ratio
r_2 = rear axle ratio
W = vehicle or combination gross weight (kg)
g = gravitational acceleration
= 9.81 m/s^2
bhp = manufacturer's maximum rated brake horsepower
E = elevation (m)

An estimate of the pseudo-maximum speed can be found with this expression:

$$\bar{V}_m = 11.28 + 2.16 \sqrt[3]{\frac{746\,bhp(1-0.0013E)}{\rho(1-0.00002099E)^{4.255}C_D A + mg\,C_{RV}}} \qquad (3.3)$$

where V_m = pseudo-maximum speed of Equation 3.1 (kph)
m = gross mass (kg)
g = gravitational acceleration
= 9.81 m/s^2
bhp = manufacturer's maximum rated brake horsepower
E = elevation (m)
ρ = standard atmospheric mass density at sea level
= 0.03779 kg/m^3
C_D = aerodynamic drag coefficient
A = projected frontal area of vehicle or assembled combination (m^2)
C_{RV} = rolling resistance coefficient
= 0.000006997 s^2/m^2

3.1.2 Commercial Vehicles. For commercial vehicles, St. John and Kobett (1978) employed a different model of acceleration capacity. Their study specifically examined grade effects on traffic flow. Because truck acceleration capacities govern performance on upgrades, it was important that the model be as accurate as possible. This meant accounting for the following:

(a) the time required to shift gears
(b) the rolling losses
(c) the chassis losses
(d) the aerodynamic losses (St. John and Kobett 1978)

The importance of the time required to shift gears lies in the driver employing the engine a varying fraction of the time depending on the grade. The performance equation St. John and Kobett developed to account for this was as follows:

$$A_e = \frac{\eta V\bar{A}_p}{\eta V + S_p t_s(\bar{A}_p - \bar{A}_c)}$$

for V > V_1 (3.4)

where A_e = the effective acceleration (m/s^2)
n = a parameter dependent on the range of engine speeds normally employed; typical values range from 0.33 to 0.43
V = vehicle speed (kph)
A_p = power-limited acceleration (ie. with engine employed and vehicle at speed V); the bar indicates the use of average available net horsepower, assumed to be 0.94 or rated maximum (m/s^2)
S_p = one times the sign of Ap (which can be either + or -)
t_s = actual time required to shift gears (s)
A_c = acceleration in coasting at vehicle speed V (m/s^2); the bar indicates the use of an average gear ratio for the coasting chassis losses; Ac is normally negative except on downgrades
V_1 = maximum speed in lowest speed gear ratio (kph)

For trucks on upgrades and downgrades, St. John and Kobett determined that the simple addition of a gravity component to the zero grade performance equation, as for passenger cars (Equation 3.1), overestimated the effects of the grades. It was therefore not satisfactory. This was because of the varying fraction of time spent changing gears. On upgrades during the time used to change gears, the grade causes an additional deceleration during the coast. This is more than offset by the gear changes being made less frequently on upgrades than on level sections. On upgrades, acceleration in each gear ratio is lower so drivers spend more time in each gear. The opposite occurs on downgrades. The upshot of this is that the model must account for grade effects within the power-limited acceleration and coasting acceleration components of Equation 3.4. These components are:

$$\bar{A}_p = \frac{\dfrac{22360(1-0.00013E)}{(mg/NHP)V} - 0.07452 - 0.000122V - \dfrac{0.2742V^2(1-0.00002260E)^{4.255}}{(mg/A)} - g\sin}{1 + \dfrac{74831}{(mg/NHP)V^2}} \tag{3.5}$$

$$\bar{A}_c = \frac{-329.1}{(mg/NHP)V} - 0.07452 - 0.000122V - \frac{2.742V^2(1-0.00002260E)^{4.255}}{(mg/A)} \pm g\sin G \tag{3.6}$$

where V = vehicle speed (kph); use $V = V_1$ if $V < V_1$
m = gross mass (kg)
NHP = rated net horsepower at sea level conditions
E = elevation (m)
A = projected frontal area (m^2)
g = gravitational acceleration
= 9.81 m/s^2
G = grade

A number of vehicle prototypes were developed using vehicle specifications from Watanatada, Dhareshar and Rezende-Lima (1987). Table 3.1 lists the specifications, and Figure 3.2 graphs the vehicle acceleration capabilities, calculated using St. John and Kobett's method.

Table 3.1 Vehicle Acceleration Performance Specifications (Sources: Watanatada, Dhareshwar and Lima 1987; Lucas 1986)

Vehicle	First Gear Ratio	Rear Axle Ratio	Gross Weight (kg)	Rated Maximum Power* (hp)	Aerodynamic Drag Coefficient	Projected Frontal Area (m^2)
Cars:						
Small Motor Car	4.10	4.440	1000	34	0.45	1.86
Volkswagon 1300	3.80	4.375	1160	49	0.45	1.80
Chevrolet Opala	3.07	3.080	1466	148	0.50	2.08
Dodge Dart	2.71	3.150	1915	201	0.45	2.20
Utility:						
Volkswagon Kombi	3.80	4.375	2155	61	0.46	2.72
Large Bus:						
Mercedes Benz 0-362	8.02	4.875	11500	149	0.65	6.30
Light Truck:						
Ford F-4000	5.90	4.630	6060	103	0.70	3.25
Heavy Truck:						
Mercedes Benz 1113	-	-	18500	149	-	5.20
Articulated Truck:						
Scania 110/39	-	-	40000	289	-	5.75

* St. John and Kobett suggest that net rated horsepower for commercial vehicles is approximately 94 percent of rated maximum.

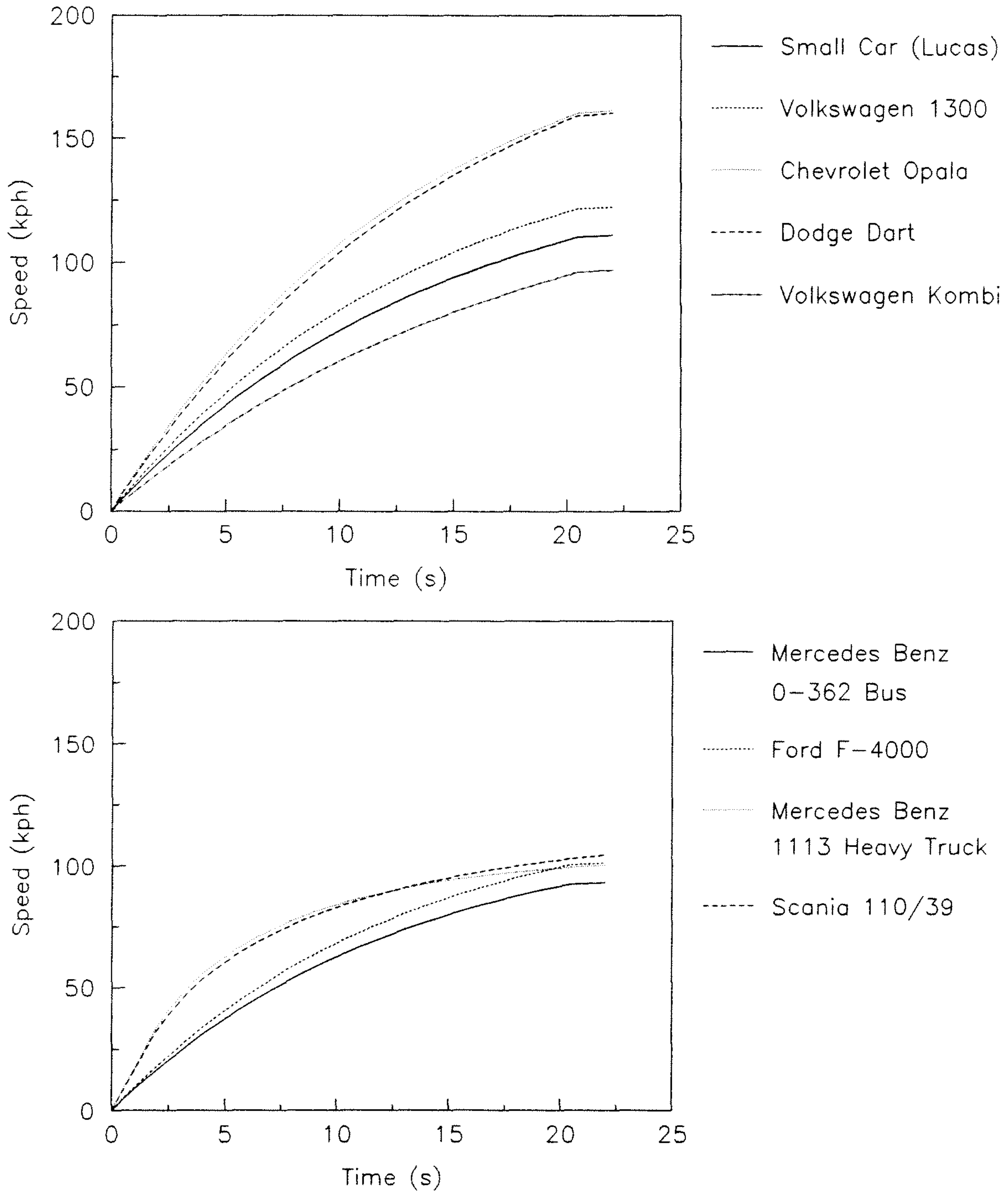


Figure 3.2 Acceleration Performance

3.2 Deceleration

Vehicle capabilities limit positive acceleration. The maximum deceleration response of a vehicle in a particular situation is also a function of the road the vehicle is travelling on. The grade and surface friction of the road provide an upper bound to maximum deceleration capability. Hutchinson and Parker (1989) call this ideal braking. They discuss it as follows:

> The maximum deceleration of a vehicle is obtained when the braking force developed at the tire-pavement interface at each wheel is just equal to the product of the normal force at each wheel and the tire-pavement coefficient of friction.

The tire-pavement coefficient of friction is a function both of the pavement and the tire. Figure 3.3 shows a range of friction coefficients for dry, wet and icy pavements. Figure 3.4 demonstrates that the coefficient is also a function of tire tread depth. Truck tires typically have lower values than car tires because of the rubber compounds used to produce better wear under heavy loads.

For a given pavement and tire, many vehicles, especially commercial vehicles, cannot achieve ideal braking. This is because of premature wheel lock-up or limited brake capabilities. Hutchinson and Parker defined the term braking efficiency to measure this lower performance. Wheel lock-up is caused when the braking force developed at one of the axles exceeds the friction force at that axle. For controlled braking, the braking force developed at the axle which locks up first governs maximum deceleration. Wheel lock-up occurs because in modern trucks the braking forces at each wheel are not proportional to the

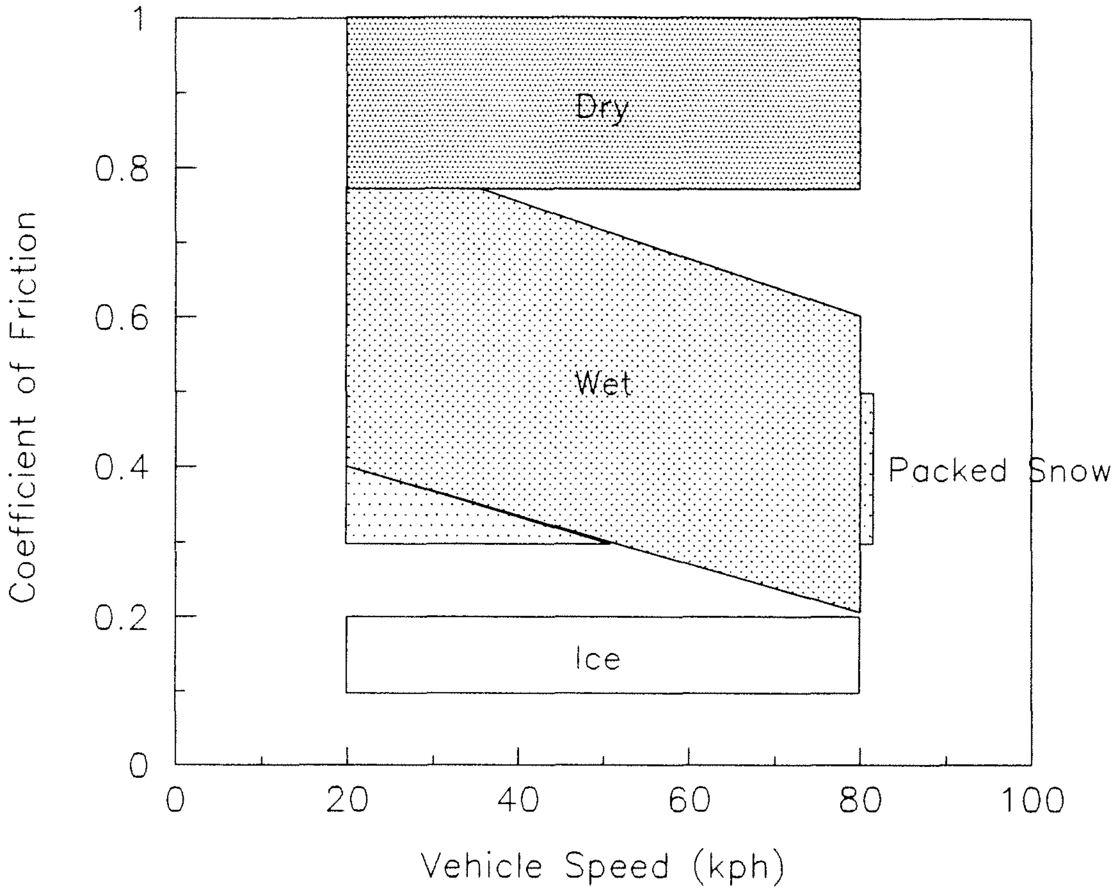


Figure 3.3 Tire–Pavement Friction Coefficients

(Source: Hutchinson and Parker 1989)

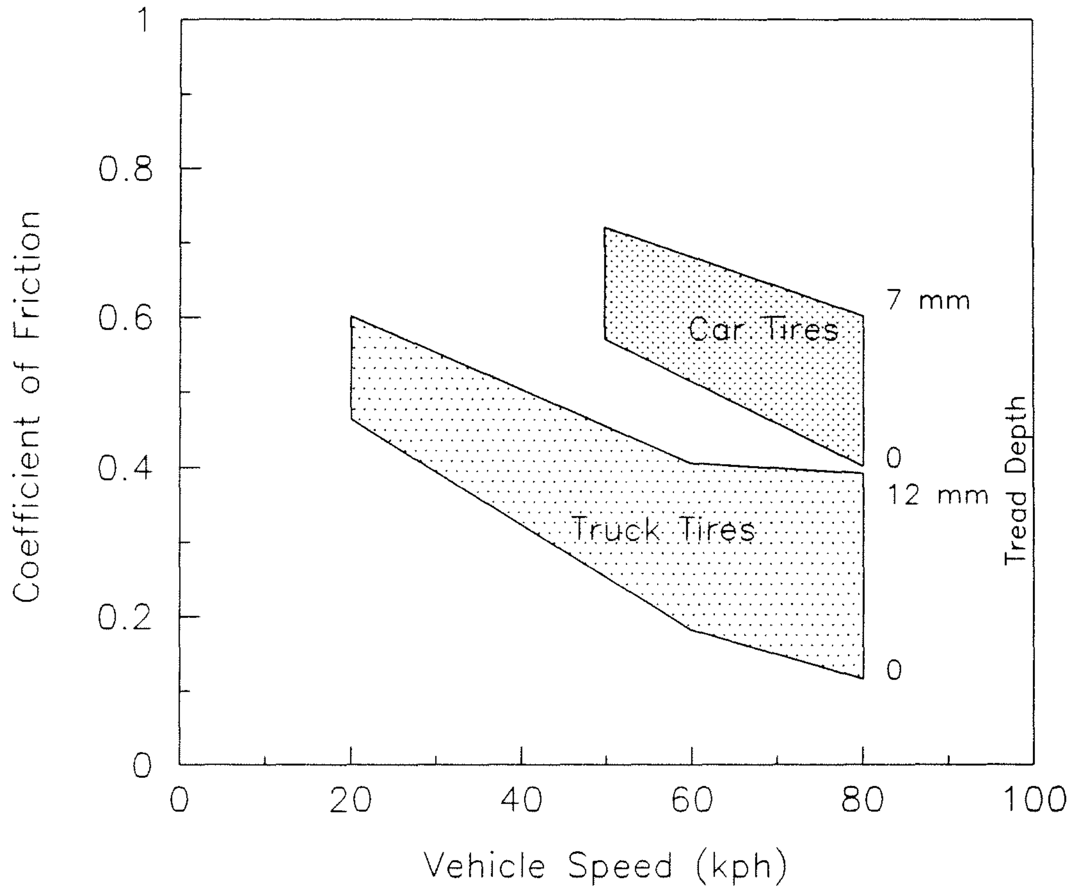


Figure 3.4 Car and Truck Friction Coefficients

(Source: Hutchinson and Parker 1989)

varying normal loads which they must carry. To maintain control once lock-up occurs at one axle, the available tire-pavement friction at the other axles is not fully mobilized. Therefore overall braking efficiency is less than ideal.

In trucks, then, braking efficiency is partially a function of load distribution. Less than ideal braking can also be caused by brake deficiencies. A truck may be underbraked for its load such that a it cannot develop a high enough braking force to use the full friction potential. Poor friction characteristics at the brake lining-brake drum interface may also limit braking force. They occur because of wetness, overheating or incorrect adjustment. Typically wheel lock up is a problem with lightly loaded or empty trucks. Underbraking is a problem with more fully loaded trucks (Navin 1986). In either case, braking may be less than ideal.

For use in the present model, a user must determine braking efficiencies externally for a particular vehicle prototype. Hutchinson and Parker (1989) describe a method of doing so. In the model, then, maximum decelerations are calculated as follows:

$$A_{min} = -g(nf\cos G + \sin G) \tag{3.7}$$

where
- g = gravitational acceleration
- f = tire-pavement average coefficient of friction
- n = braking efficiency (dimensionless)
- G = grade

3.3 Steering

Coordinates and a heading in two dimensional space give the trajectory of a vehicle. Speed is the rate of change of the coordinates. The driver cannot directly control that but he can control the rate of change of the speed (acceleration). Thus he can control the rate of change of the rate of change of the vehicle coordinates. In the same way, a driver can only control the rate of change of the rate of change of the vehicle heading. He determines the trajectory of his vehicle by controlling two acceleration quantities, one translational, the other angular.

This model assumed that the time over which a driver changes the steering angle from one value to another (ie. the time during which the angular acceleration is non-zero) is small compared to the model scanning time. Therefore it deemed him able to 'instantaneously' change the steering angle of his vehicle. The steering angle is the angle between the heading of a vehicle's front tires and the longitudinal axis of the vehicle (Figure 3.5). The rate of change of the heading is a direct function of the steering angle and the wheelbase of the vehicle (See Figure 3.5). The model calculates it as follows:

$$\Delta h = \frac{S(\Delta D)}{w} \tag{3.17}$$

where

h	=	heading
S	=	steering angle
D	=	distance travelled
w	=	vehicle wheelbase

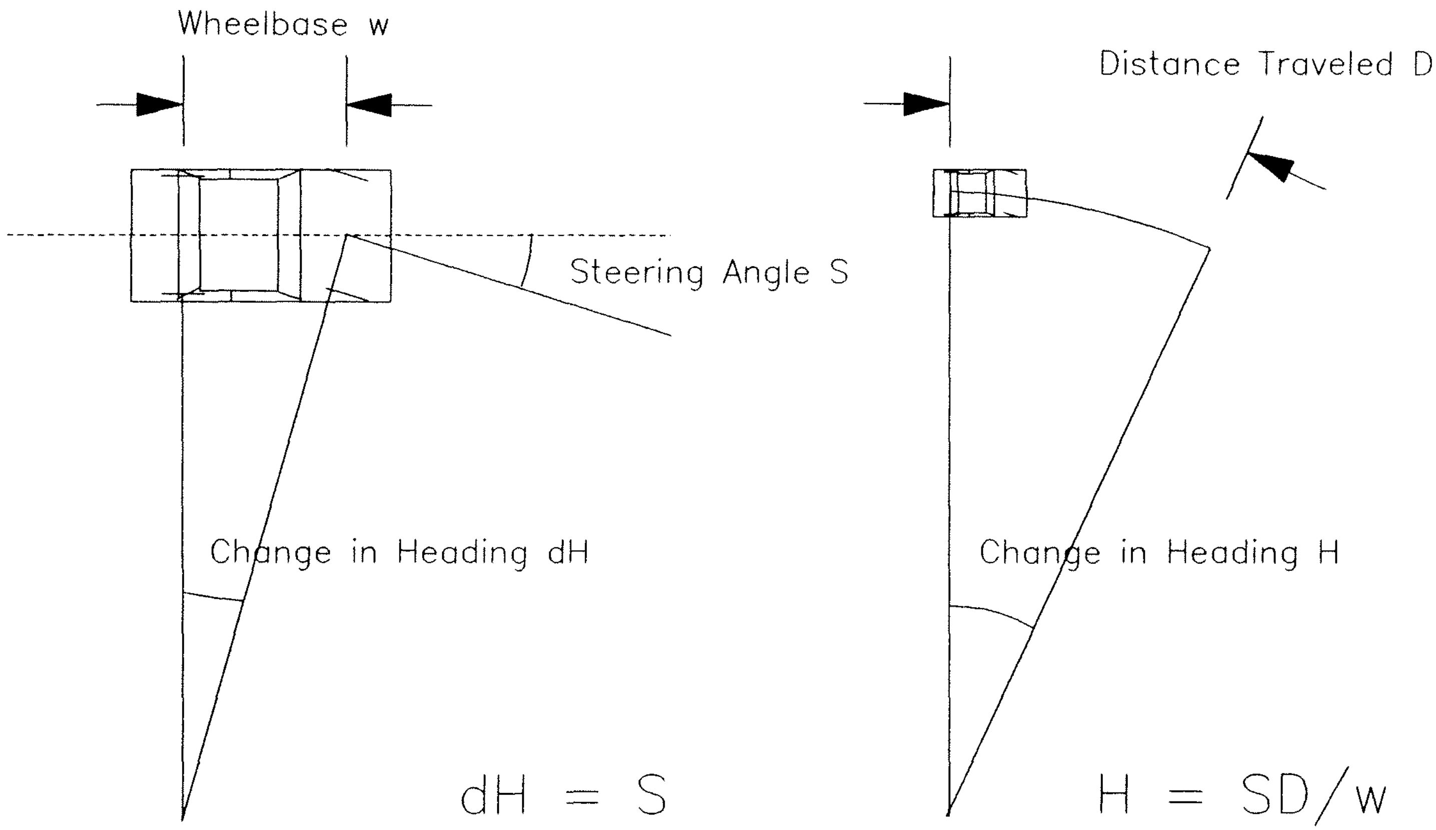


Figure 3.5 Vehicle Steering

4 THE ROAD NETWORK

In the traffic flow system model, the road's relationship with the vehicle is similar to that between the vehicle and the driver. It's role is to provide boundaries within which the vehicle can operate. This, in turn, provides a further limitation on the freedom of choice and action of the driver. The other limitations are those of vehicle acceleration, deceleration and steering. Basically, the road constrains the movement of vehicles to within its lanes and intersections.

The network bounds vehicles in their movements as shown in Figure 4.1. Vehicles can move to any point on a roadway. They can be entirely within a lane, straddling lanes or in opposing lanes. They can have any heading. Vehicles are out of bounds if they attempt to drive on the shoulders, or cross a barrier or divided median. The model does not prevent drivers, however, from choosing such actions. It removes vehicles which move out of bounds from the simulation.

The road network puts two other constraints on vehicles and drivers. The first is that of limiting speed on horizontal curves so centripetal forces do not exceed lateral friction forces. The second is limiting sight distance because of both roadway alignment and adjacent topography. These two constraints are discussed, respectively, in sections 4.1 and 4.2.

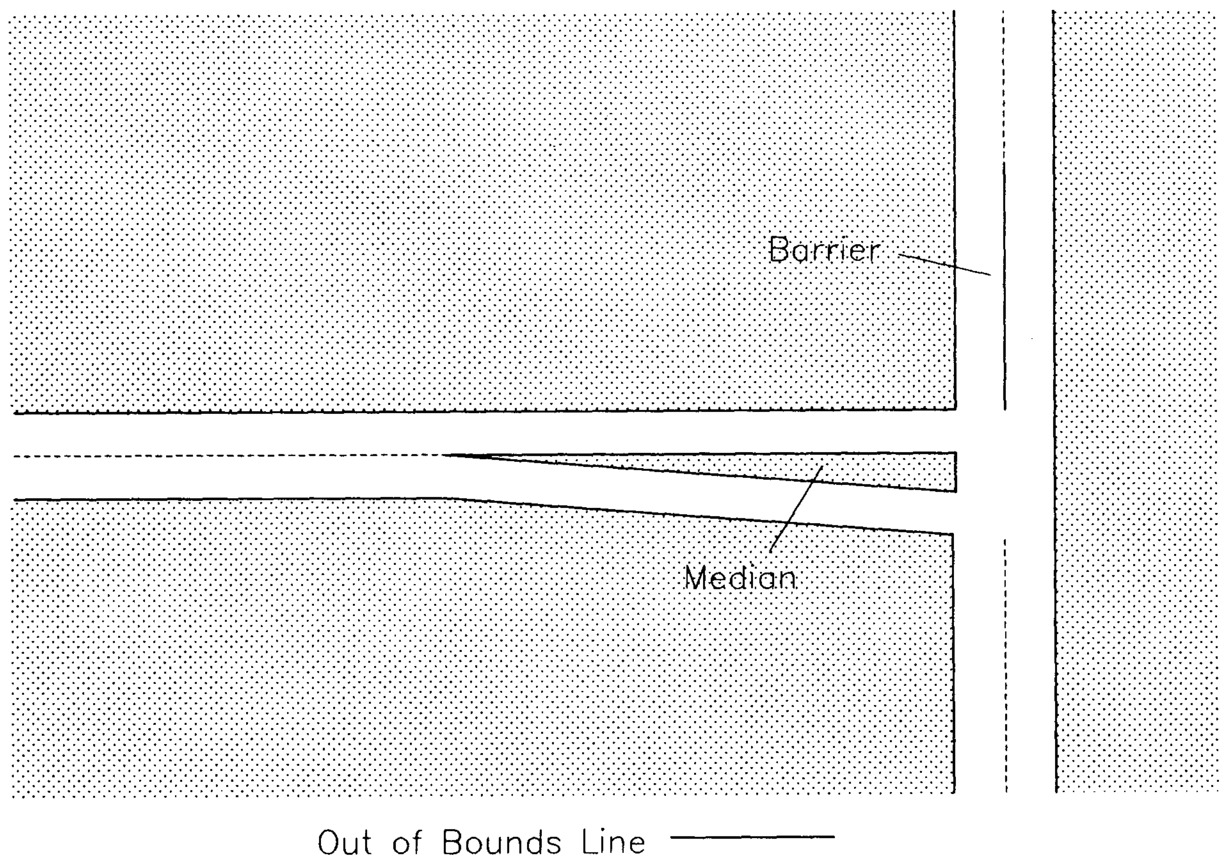


Figure 4.1 Road Bounds

4.1 Horizontal Curves

On tangent sections, the roadway does not limit vehicle acceleration or speed. On curves, however, there is a limit. Vehicle acceleration on curves has a component acting radially from the curve centre. The forces associated with this radial, or centripetal, acceleration limit vehicle performance. Figure 4.2 diagrams this. For system stability, the friction force acting radially inward from the centre of the curve must balance the force acting radially outward. The outward force is caused by the vehicle's centripetal acceleration. The interaction of the road surface and the vehicle tires limit the friction force, though. As vehicle speed increases so do the centripetal acceleration and associated radial force. At the same time, the friction force increases to balance the system until it is mobilized to its limit. Beyond this point, increased vehicle speed causes instability. The vehicle accelerates laterally off the side of the road outward from the curve centre.

There is thus a maximum speed attainable on horizontal curves. This speed is calculated as follows:

$$V_{max} = \sqrt{gR(f+e)} \tag{4.1}$$

where V_{max} = Maximum speed on horizontal curve
g = gravitational acceleration
= 9.81 m/s^2
R = Curve radius
f = Roadway coefficient of friction
e = superelevation

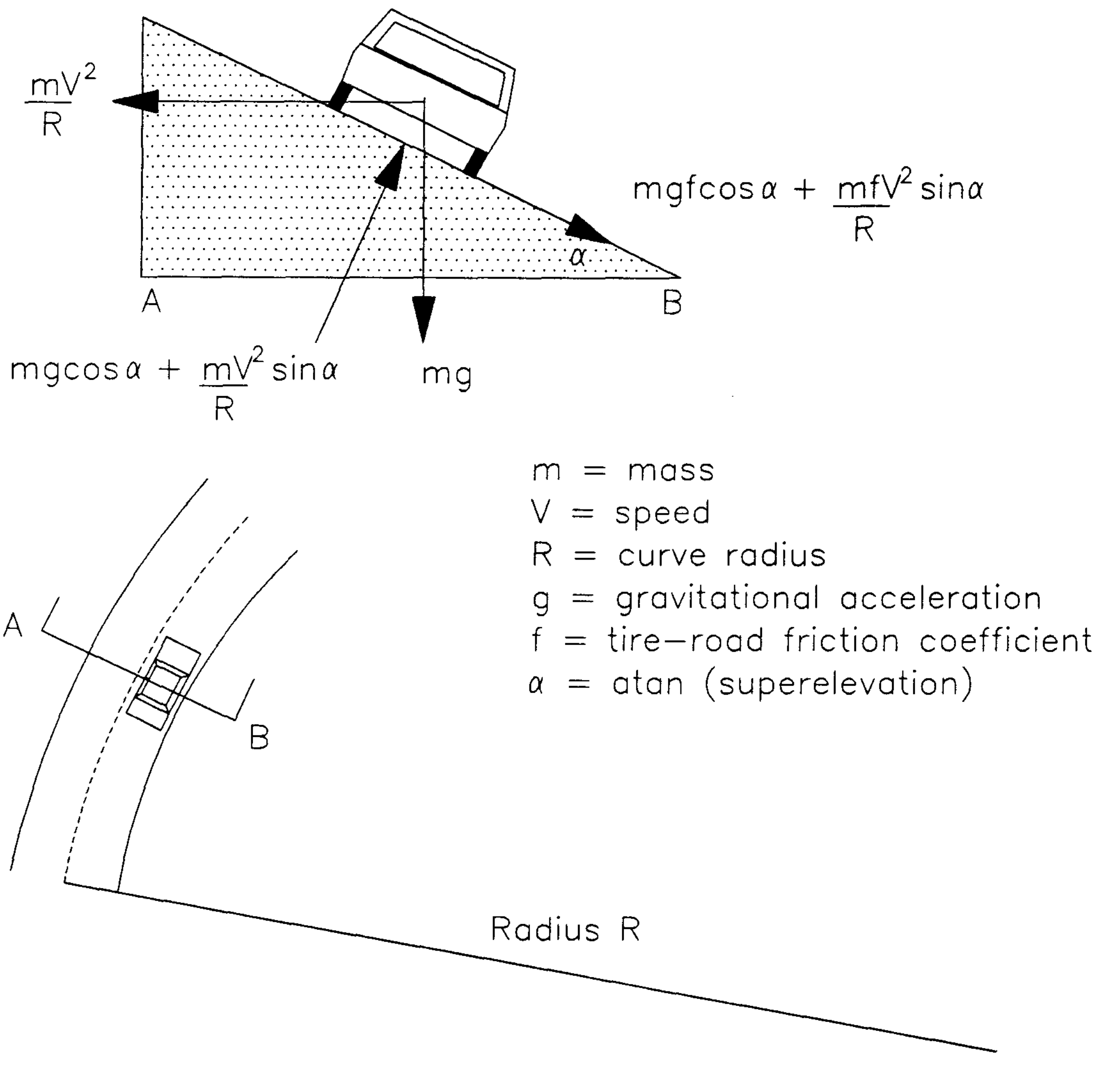


Figure 4.2 Horizontal Curves

As with driving out of bounds in other locations, drivers can choose to exceed the maximum speed on horizontal curves. They slip off the side of the road, and the model removes them from the simulation.

4.2 Sight Distance

The road network limitations discussed thus far are a set of consequences for particular driver actions. Sight distance, on the other hand, actually limits what a driver can do, or in this case, see. The roadway, specifically vertical alignment, and the topography adjacent to it can limit sight distance.

The horizontal curve limitation and vehicle limitations are a number of checks to make after the model determines driver output commands. The sight distance limitation acts at the start of the driver's processing. Sight distance, as discussed in section 2.2.1,determines the extent of the active environment around a driver. Thus it limits the region within which he obtains information from the environment.

5 SYSTEM MODEL

Sections 2 through 4 discussed the components of the traffic flow system as modeled. This section combines the driver, the vehicle and the road network and explores their interactions. Section 5.1 deals with the creation of vehicles and drivers for simulation, and section 5.2 deals with simulation scanning.

5.1 Traffic Generation

The traffic flow simulation starts with a single empty road network bounded by entrances and exit points as shown in Figure 5.1. The system model generates traffic at the entrances. The entrances form boundary conditions in the model. A user must carefully formulate their traffic generation functions in order not to adversely affect results.

Traffic generation consists of creating a set of vehicles and drivers and releasing them in succession according to a series of headways. The model must also assign an initial speed to each vehicle. For a particular entrance, the speed and headway distributions at the corresponding point in the real system are partially functions of roads upstream. The difficulty lies in generating traffic as would these unmodeled upstream roads.

There are two subsets of headways in traffic: free-flow and constrained. In the first set are

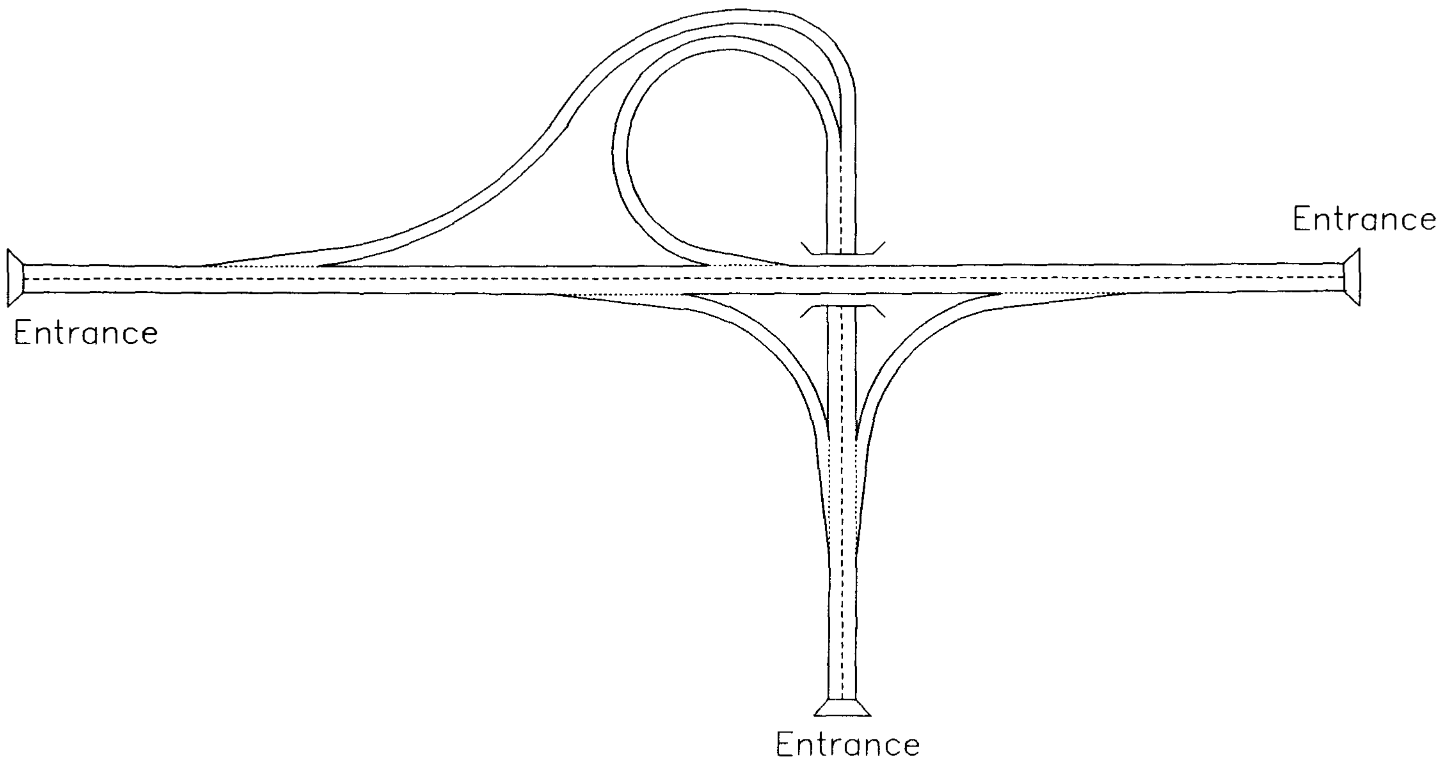


Figure 5.1 Traffic Generation Points

driver-vehicles whose actions are not influenced by vehicles ahead and have headways greater than about 4 seconds. In the second set are driver-vehicles who are travelling closer to the vehicles ahead and whose actions are influenced by those vehicles (Wu and Heimbach 1981).

The sizes of these subsets in a traffic stream are functions of the traffic volume. When the volume is low, there are proportionately more free-flow vehicles and fewer constrained ones. The opposite is true at high volumes. Grecco and Sword suggested a headway probability model which is a composite of the two subsets. As quoted by Gerlough and Huber (1975), it is as follows:

$$P(h>t) = 0.00115Ve^{-\frac{(t-1)}{2.5}} + (1-0.00115V)e^{-\frac{t}{24-0.0122V}} \qquad (5.1)$$

where P(h > t) = Probability that the headway h will be greater than time t
V = Average volume (vph)
t = time greater than 1s in seconds

The first term in this equation describes constrained headways and the second term describes free flow headways. Figure 5.2 shows plots of headway probability distributions for various traffic volumes.

During simulation, the model generates vehicles randomly at each entrance according to user specified average volumes and a probability distribution such as that of equation 5.1 (the default). Users can specify other distributions.

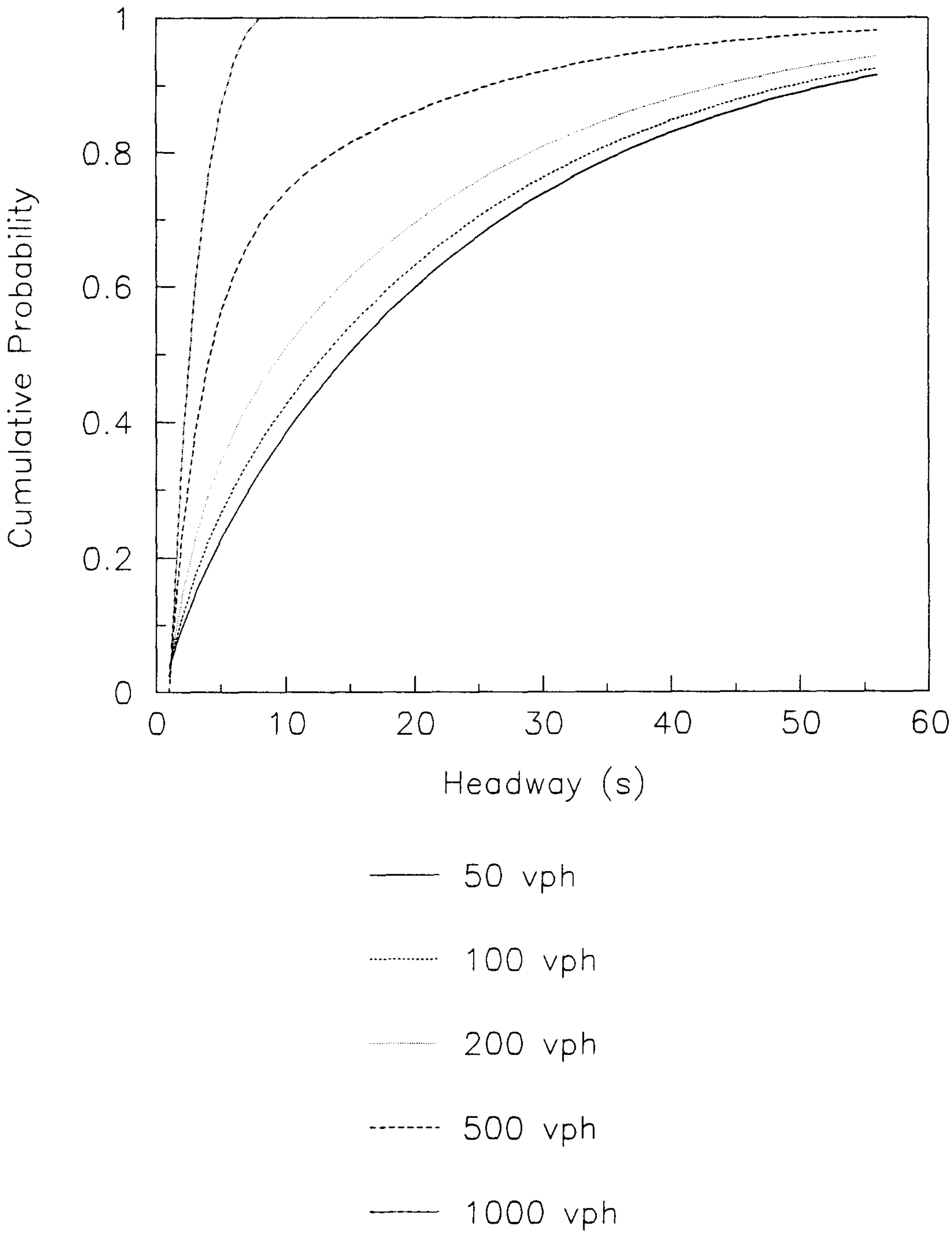


Figure 5.2 Vehicle Headway Distributions

Each generated vehicle is an instance of a vehicle prototype. For each vehicle, the model determines its prototype randomly according to average traffic composition percentages specified by the user for each entrance. Assigned to each vehicle is a driver, who, similarly, is an instance of a driver prototype. The model chooses the driver prototype randomly according to driver composition percentages specified by the user for the vehicle prototype. Thus, the model generates a succession of vehicles with randomly determined headways at each entrance. Each vehicle is an instance of a randomly determined vehicle prototype. It contains a driver who is an instance of a randomly determined driver prototype.

At the beginning of his life at an entrance, a driver has the following starting conditions:

- nothing in short term memory;
- under average stress;
- not undertaking any conscious driving programs; and
- focusing on the road ahead.

At the same time, the vehicle has these starting conditions:

- acceleration of zero;
- lights, signals, and flashers off; and
- speed determined by user specified distributions.

Because the driver has nothing in short term memory and is not consciously doing anything when he enters the system, he requires a short length of undemanding roadway to 'warm up'.

This involves filling his short term memory with some knowledge of the active environment as a basis for action.

The only one of these starting conditions which requires user input is the initial speed of the vehicles. The user specifies this using the speed of the preceding vehicle and the new vehicle's headway. Though the present model assumes no initial speed distributions, Lyons, Rainford, Kenworthy and Newman (1988) suggested the following rules for determining them.

> For vehicles on the freeway with an initial headway less than 1.5 s, their initial speed is set equal to that of the preceding vehicle, provided this speed is not greater than their individual desired vehicle speed. If the headway is greater than 1.5 s but less than 3.0 s, the initial vehicle speed is chosen randomly from a normal distribution centered on the speed of the leading vehicle with a maximum range of one-tenth of this speed. For all headways greater than or equal to 3.0 s, the initial vehicle speed is set at its desired vehicle speed. Thus at vehicle generation it is assumed that free flow exists for headways in excess of 3.0 s on the freeway.

The user specifies initial speed distributions as a function of preceding vehicle speed and headway for each driver prototype at each system entrance.

5.2 Simulation Scanning

Once the model generates driver-vehicle combinations at entrances, it determines their progression through the road network incrementally. The model performs scanning on a time rather than an event basis. In each time step, 0.2 seconds in length, the model advances thought and command processes of each driver. It also updates vehicle positions. The sequence it does this is as follows:

1. For each entrance, generate traffic as described in section 5.1;

2. For each driver,

   (a) add or update records in short term memory;

   (b) make new predictions if the driver is consciously driving and is giving attention to an object in the environment with time varying characteristics;

   (c) issue conscious and unconscious commands to the driver's vehicle delayed by appropriate reaction times;

3. For each vehicle,

   (a) check acceleration and steering commands issued by the driver to ensure that they fall within limitations;

   (b) update the vehicle's acceleration, steering and other characteristics;

   (c) calculate the vehicle's new speed, heading, and location.

The model performs these three steps at each time step, interrupted only by collection of output estimates as described in section 8. Once driver-vehicles reach system exits, it deletes them from the simulation.

6 COMPONENT CALIBRATION

The traffic flow model is generalized enough for use in modelling many different types of systems, including freeways, urban networks, intersections, and two lane highways. To demonstrate it, though, driver prototypes were developed for the last of these systems: two lane highways. Many of the calibration issues discussed in this section apply to all of the other systems.

Analysts undertake the calibration of traffic flow simulations at two levels: component calibration and aggregate calibration (McLean, 1989). A simulation model is a collection of sub-models each representing a system component. Component calibration is the adjustment of sub-model parameters so the behaviour of the sub-models accord with real-world behaviour. Unfortunately, the collectors of much of the empirical data available to do this did not structure them using the sub-model variables. A tuning process is required. Sub-model variables are repeatedly adjusted and simulations undertaken until the overall sub-model behaviour matches that of the data.

Aggregate calibration examines the predictive validity of the entire model. Typically, this involves comparing predicted properties of the model to those of an equivalent real traffic stream. Such properties include speed distributions, passing, platooning, travel times and speed-flow relations. A modeller adjusts parameters in the same way as for component calibration. He seeks a balance between the degrees to which he calibrates a model correctly on the component and aggregate levels. This section discusses component calibration, and section 7 treats aggregate calibration.

As section 2.2 described, the present model does not presuppose any fundamental set of driver characteristics upon which all behaviour is indexed. In this way, it does not pre-bias driver prototypes in any fashion. However, the use of such variables in older models did serve a purpose. The variables allowed one to construct a sample of composite drivers from statistics describing individual tasks and behaviours. There is extant in the literature considerable data on these behaviours, data which one can observe and measure. However, it would be a difficult and monumental task to measure and correlate all such behaviours with each other.

Fundamental driver characteristics, and the theories which spawn them, allow one to do this and generate composite drivers. Unfortunately, because of the psychological fact that individuals are unique, these characteristics introduce biases. By not presupposing any set of fundamental characteristics, the present model's achievement is to make the user responsible for them. At the same time, it gives him the power to change them and evaluate their effects.

In the development of two-lane driver prototypes discussed in this section, it was necessary to assume a set of fundamental characteristics. However, it is important to make the distinction that this was part of the example application. It was not part of the model itself. As with the rest of the work in this chapter, a user can easily choose to disregard the set of fundamental characteristics developed here and use one of his own.

Automobile drivers, in their seeming preference for constant speed open road driving, display tendencies toward constancy and control.

> [O]ne may conclude that each driver behaves in traffic so as to satisfy his own goals, independent of what is best for the overall traffic situation. His own goals are reflected above by the desired constant velocity input. The variables which determine this desired velocity are not known, but it is possible that this velocity is set even before a driver enters his automobile. He may driver at 20 mph regardless of where he is, or he may driver 5 mph over the speed limit to minimize travel time while not attracting the attention of the police, or he might have a fixed distance to travel in a fixed amount of time. Each example describes a different driving mode which could produce a constant desired velocity, and it is possible to say that in these modes, the driver is program bound ... One would also predict from the behaviour towards constancy, that each driver-vehicle performs as an inertial system, subject more to relative values in traffic flow than absolute values (Algea 1964).

Independent open road driving is the preference of most drivers. Some, though, are as happy following another vehicle travelling at a comfortable speed. This is a trade off between control (or perceived safety) and constancy. By following another vehicle, a driver gives up some control of the situation to the other driver in order to make it easier to achieve constancy. The other driver may do something unexpected or dangerous. Allowing the lead driver to set the speed means a following driver does not have to do it. Note that safety (degree of control over one's destiny) appears not to be the overriding concern or objective of drivers in normal situations. It is rather just one of the trade off variables.

To compromise safety is to take a risk. Risk taking is the yielding up of control in a situation. Control is the power to determine the outcome of a situation. In driving, elements in a situation with which a driver could not cope if events developed in a particular way limit the degree of control he has over his own destiny. A driver only has complete control if he can cope with *all* the contingencies which could develop. He can lose control by chance developments such as a tree falling across the road or by the actions of others. Another driver could make an unexpected movement. Drivers can also lose control because of their own intended or unintended actions. Control, just like its complement risk taking, is a subjective phenomenon (Taylor 1976).

Subjective estimates of risk are often less than objective risk. Most drivers probably feel that accidents only happen to other drivers and are sufficiently rare as not to cause worry. For example, in responding to reduced visibility, most drivers do not adjust their speeds downward sufficiently to maintain constant objective risk. This is a result both of a trade off with constancy and with an underestimation of risk in the situation.

> Drivers have become conditioned to expect that there will be no obstruction to their pathway, and therefore will only be affected by restricted visibility when it restricts their steering (and perhaps navigation) cues" (Hills 1980).

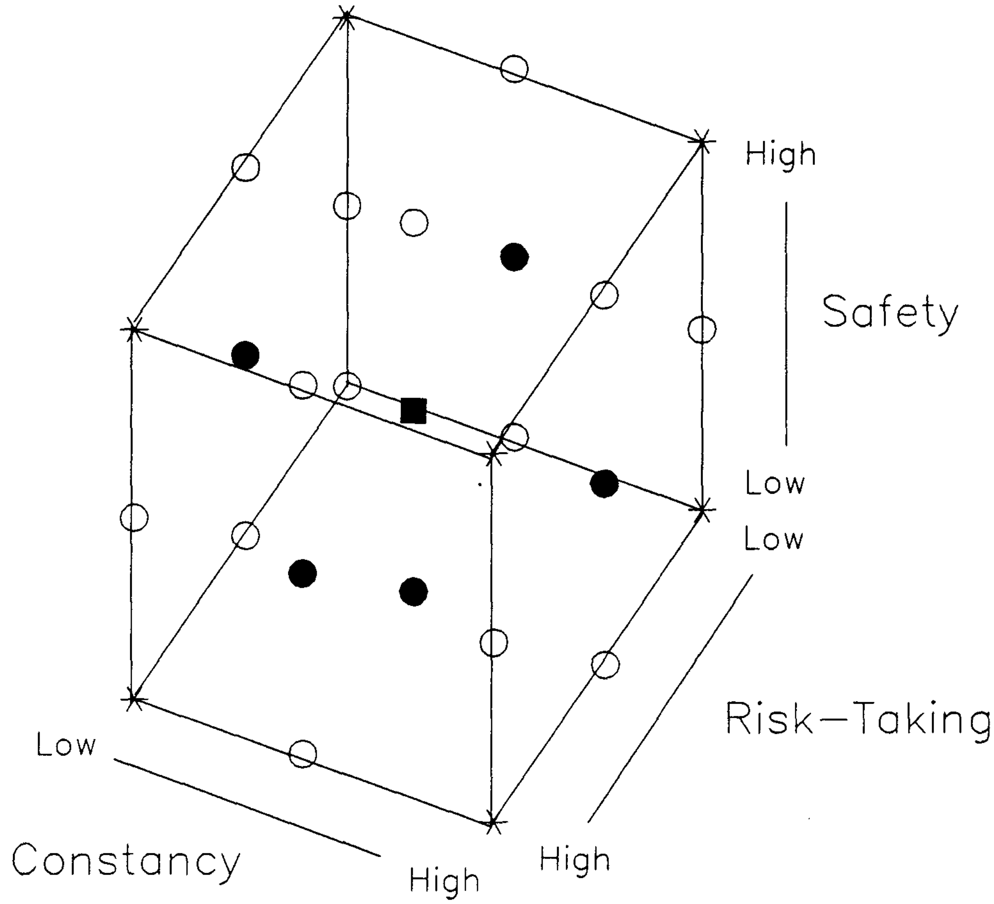
Because safety is not the only concern of the driver, he does not always seek a constant risk level. Road safety improvements are designed to lower risks. Sometimes drivers respond to them, however, by taking on other risks. They do this not to necessarily achieve the prior risk level. Rather, they see an opportunity to trade off some risk to satisfy another objective.

The effect of safety changes depends on how the changes affect all objective variables.

Thus far the objectives drivers weigh include constancy and safety. A third objective is risk taking, referring, of course, only to subjective risk. The objectives are considered from the point of view of the driver. Taylor postulated that people seek small subjective risks as means of arousal and excitement. The evidence for this lies in chosen pastimes such as skiing and horseback riding, and in risky but otherwise purposeless driving behaviour. He notes the example of seeing a risk-taking driver who "disappeared ahead in a cloud of dust taking his ease in the next ice-cream parlour along the road." Reducing travel time was obviously not this driver's objective.

This section has outlined three objective parameters upon which to create driver prototypes for two-lane driving. Twenty seven driver prototypes were created, each given a unique combination of weightings for the objective parameters (Figure 6.1). Each driver prototype was also arbitrarily assigned an expected frequency of occurrence within the driving population. Moderate drivers (those in the middle of the cube in Figure 6.1) were the most common.

For each driver prototype, driving programs were developed. The programs were structurally the same but were calibrated to different values reflecting the prototype objective parameter weightings. For each individual behaviour, this was done in the following way. For each variable in the behaviour which varies across the driving population,



Driver Prototypes	Frequency within Driving Population
✳	Low
○	
●	
■	High

Figure 6 . 1 Driver Prototype Matrix

such as gap acceptance in passing, the objective parameters being traded off were determined. This could be a one-for-one or a one-for-two trade-off. A function was then developed which would generate the population distribution for the variable using the objective parameters. The function had one of the forms shown in Figure 6.2 depending on the applicable trade-offs. Once the function was determined, values were assigned to the behaviours of each driver prototype, depending on where in the cube it lay.

The following sections describe the performance and behaviours of only one particular driver prototype, the average driver. In creating the other driver prototypes, the performance and behaviours of this prototype was varied according to the manner just described. Section 6.1 discusses the calibration of driver pattern recognition and prediction. Sections 6.2 through 6.6 discuss the calibration of programs, conscious and unconscious, used in two-lane highway driver prototypes.

6.1 Attention and Visibility

The information which a driver obtains from the environment is a function both of his perceptive abilities and of his allocation of attention. Researchers have directed much effort toward determining the limitations of the former and ensuring that the presentation of important elements of the environment fall within these limitations. Previous sections demonstrated, however, that it is the latter which is most important in determining how

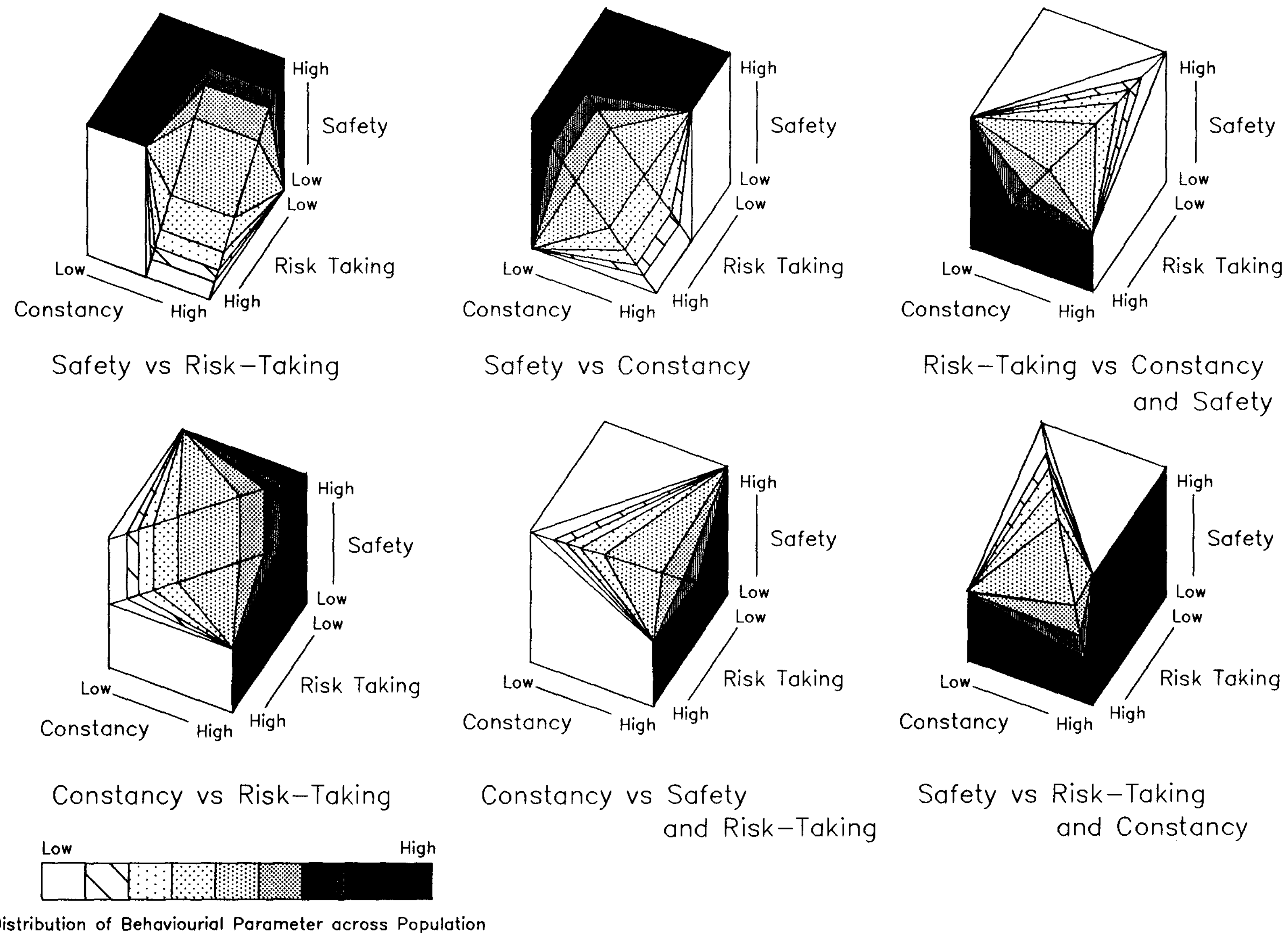


Figure 6 . 2 Behaviourial Functions on Driver Prototype Matrix

much and which information a driver obtains.

In a study of vehicle following, Colbourn, Brown and Copeman (1978) concluded that "the driver's main problem in safe vehicle following derives from difficulty in evaluating risk and hazard, rather than from the limitations on his sensory and perceptual abilities which underlie most previous explanations of human inadequacy in this situation."

Section 2.2.1 defined attention and search conspicuity. The former was assumed to occur in peripheral vision and the latter in foveal vision. In a study comparing these two types of conspicuity in typical driving environments, Cole and Hughes (1984) concluded that, on average, search conspicuities are three times higher than attention conspicuities. They measured them by recognition percentage. There is, however, a greater gain in conspicuity in the search mode for objects with lower rather than higher attention conspicuity. In the transformation functions used in pattern recognition in this model, this meant that fuzzy grades increased three times in the fovea compared to the periphery.

The angular size on the retina at which drivers first notice objects can be as small as one degree. Attention conspicuity drops off substantially beyond 10 degrees from the line of sight for stationary objects. For moving objects, the relative motion against the background extends the area in which an object is conspicuous to higher angles. In a moving environment such as that of an automobile, the eccentricity angle effects on conspicuity are more important than those of size. Thus, a driver is more likely to notice an object if it has a small eccentricity and small visual size than if it has a large eccentricity and large visual size. This differs from the static environment situation in which eccentricity and size tend to balance one another (Cole and Hughes 1984).

What do drivers pay attention to in two-lane highway driving? This is dependent both on the experience of a driver and his familiarity with the route he is travelling. In open road driving, a driver's mean eye focus location is higher in the scene than it is in a car-following situation. The distribution of focus locations is also more compact in the latter. Compactness increases with route familiarity in both situations (Mourant and Rockwell 1970).

In both open road and car following situations, drivers make few direct fixations on lane marking lines. They rather rely on peripheral vision to guide their steering. They also use peripheral vision to monitor the positions of other vehicles. Much of the time, drivers look ahead toward the focus of expansion in open road driving. They fixate on the lead vehicle in vehicle following situations (Mourant, Rockwell and Rackoff 1969). Table 6.1 shows fixations and their corresponding durations in open road driving. Table 6.2 gives similar values for vehicle following.

Much of the discussion in this section thus far has dealt with attention. However perception abilities are also a factor in generating pattern recognition functions for use in the model. There have been many studies which have shown that people are notoriously poor at judging

Table 6.1 Eye Focus Locations in Open Road Driving
(Source: Mourant, Rockwell and Rackoff 1969)

	Unfamiliar with Route		Familiar with Route	
Focus Object	Viewing Time Percentage	Fixation Duration (s)	Viewing Time Percentage	Fixation Duration (s)
Looking Ahead	50.4	0.26	58.3	0.26
Road and Lane Markers	2.2	0.31	2.0	0.26
Vehicles	5.0	0.28	6.7	0.25
Road Signs	7.5	0.33	5.4	0.30
Bridges	8.0	0.30	7.1	0.29
Rear View Mirror	6.1	0.61	4.7	0.61
Side Mirror	2.2	0.66	1.7	0.66
Speedometer	5.5	0.72	4.2	0.72
Blinking	4.8	0.16	3.7	0.16
Other	8.2	-	6.2	-

distances and speeds in the driving environment. The distances and speeds are measured in object terms, though. For behaviour, this is unimportant because people do not drive on the basis of objective measurements. Rather, they drive on the basis of fuzzy terms such as far, near, fast and slow. *They have done so from the first time they sat behind the wheel.* The importance of this last statement is that drivers learn how to drive based on their perceptions. Their actions therefore adjust and compensate for any objective perception errors. The upshot is that the ability to judge 100 metres as 100 metres is of no consequence in driving.

Table 6.2 Eye Focus Locations in Car-Following
(Source: Mourant, Rockwell and Rackoff 1969)

Focus Object	Unfamiliar with Route		Familiar with Route	
	Viewing Time Percentage	Fixation Duration (s)	Viewing Time Percentage	Fixation Duration (s)
Looking Ahead	31.2	0.30	30.7	0.31
Road and Lane Markers	2.2	0.28	1.8	0.23
Lead and Other Vehicles	38.8	0.34	44.3	0.32
Road Signs	4.9	0.33	2.5	0.31
Bridges	5.8	0.32	5.4	0.28
Rear View Mirror	3.9	0.61	7.0	0.61
Side Mirror	1.4	0.66	2.5	0.66
Speedometer	3.5	0.72	6.3	0.72
Blinking	3.1	0.16	5.5	0.16
Other	5.2	-	9.4	-

Perceptual errors do arise, however, but they consist of incorrect perception of fuzzy measures, for example interpreting 'far' as 'not very far'. Such errors are due entirely to lack of experience of a particular perception. This is why some drivers have difficulty judging safe passing manoeuvres at high speed. It is not because they are better at judging 100 metres than 400 metres. It is because they have had less experience in high speed passing. Their perceptual judgement in these situations is therefore more suspect but suspect using fuzzy yardsticks.

In the present model, pattern recognition functions with which a driver has had considerable experience, consist of a measurement of attention bounded by perceptual limits. For example, a driver can detect the speed of another vehicle only when the vehicle is within a certain distance. Within this distance, if the speed is in the range normally found on the road, its pattern recognition is only a function of the attention paid to it. The fuzzy graphs are centred on the correct value and are flatter and broader with decreasing attention. If the speed is not in the range normally found on the road, its pattern recognition is a function of the attention paid to it but may additionally be biased toward the expected range (Figure 6.3).

6.2 The Driver and the Road

The two-lane highway driving programs were classified according to the interactions a driver might have with specific elements of the driving environment. This section describes the programs dealing with road sections.

6.2.1 Speeds on Tangent Sections. In non-stressful situations in which other vehicles or the roadway alignment do not constrain a driver's progress, he will tend to travel at or near a particular 'comfortable speed'. This speed varies from driver to driver but was considered to be unconscious behaviour, as described in section 2.2.6,in all but novice drivers.

McLean (1981) suggested that for a particular driver on a particular route, desired speed is a function of road classification, trip purpose and length, proximity to urban centres and

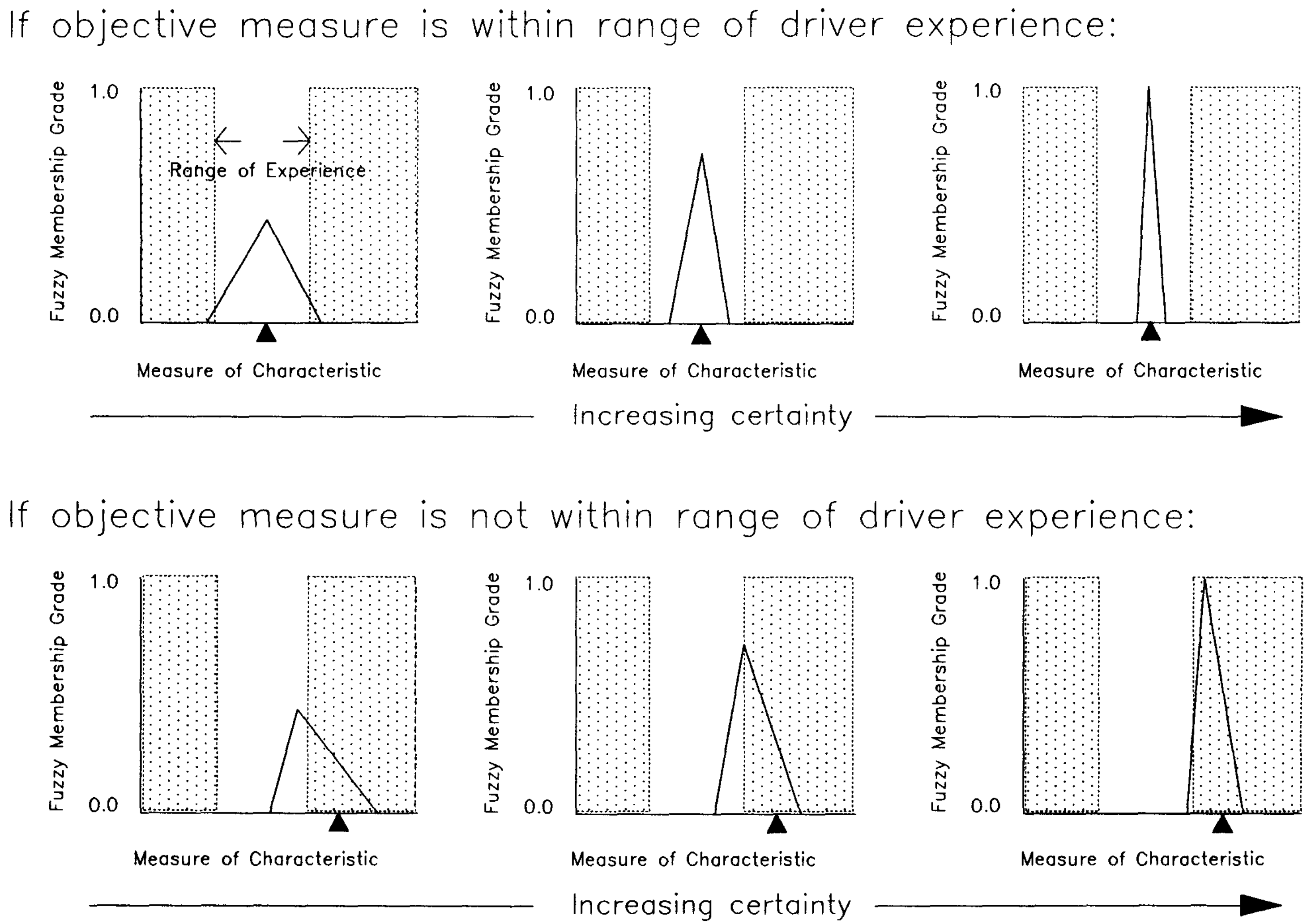


Figure 6.3 Effect of Experience on Pattern Recognition

overall alignment standard, specifically design speed. From measurements taken in Australia on two-lane rural roads with a State Highway classification, he also determined that desired speeds for a sample of passenger car drivers were normally distributed with means and 85th percentile values as given in Table 6.3.

It has been shown that desired speeds are not a function of shoulder type (grass, gravel or pavement) or lane width for widths normally found on two-lane highways. Taragin (1944) studied traffic on two-lane roads with lane widths ranging from 5.4 m to 7.2 m and determined that lane width had no discernable effect on the speed of free moving vehicles, regardless of whether they were facing opposing traffic. This held both for day and night conditions. Shelby and Tutt (1958) also reported that lane width and shoulder type had no significant effect on travel speeds on tangent sections. They obtained 85th percentile speeds in the ranges of 94 to 110 kph for passenger cars and 72 to 94 kph for trucks on long tangent sections of a two-lane suburban highway.

Consider the effects of speed adaptation on attempting to maintain a desired speed. As noted in section 2.1.2,senses become more sensitive in the absence of stimulation and less so in its presence. In open road driving in which the environment presents a driver with a constant stimuli, sensitivity to the moving scene increases. Higher speeds begin to seem lower than they actually are. In the present model, stress is proportional to stimulation and, as a consequence, sensitivity increases as stress decreases. The effect on the maintenance of desired speed is such that at low stresses, drivers allow their speeds to increase.

Table 6.3 Desired Speeds on Two-Lane Highways (Source: McLean 1981 and McLean 1989)

Overall Design Speed (kph)	Desired Speed (kph)*		
	Flat	Rolling	Mountainous
40 - 50			63 (70)**
50 - 70		79 (90)	
70 - 90		88 (100)	
90 - 120	100 (115)	96 (110)	
>120	104 (120)		

* Mean desired speeds are listed with 85th percentile speeds in parentheses.
** Under these conditions, tangent lengths are too short for a meaningful measure of 'desired' speed. The values given are based on the typical speed distributions measured on available tangents, which might well be lower than the speeds which might have been observed had the tangents been longer.

The average driver prototype incorporated three unconscious instructions to account for desired speed behaviour. Figure 6.4 shows them. The first instruction satisfies the desire to increase speed to a desired level, in this case about 75 kph, suitable for an average driver on a mountain highway. The second and third instructions are the same except the speed portion of the antecent. The high stress instruction attempts to keep the speed down to about 75 kph. In the low stress instruction, speed adaptation effects raise this value to 80 kph.

Referring to Figure 6.2, the desired speed objective function across driver prototypes had

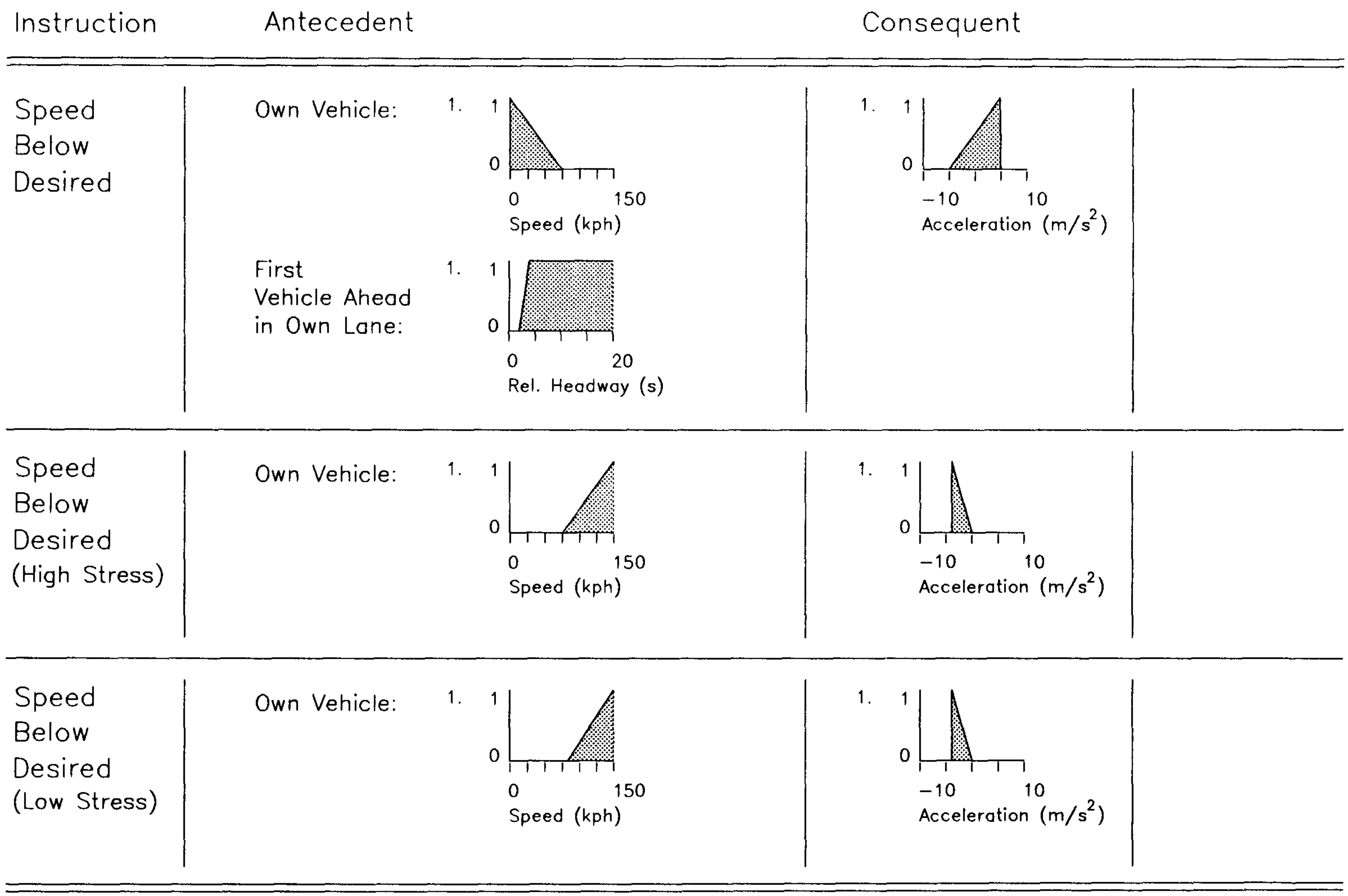


Figure 6.4 Desired Speed Program Instructions

a safety versus risk-taking trade-off. This was even though adopting a desired speed is in itself an example of constancy behaviour. Choosing higher speeds reduces safety and increases risk-taking. The rate at which the environment presents information to the driver is directly proportional to speed, and the human capacity to retrieve this information is rate limited. A tendency to slow down in situations where the driver must process more information is evidence of this latter point.

6.2.2 Lane Keeping on Tangent Sections. Donges (1978) modelled steering within a traffic lane as two parallel processes, one of guidance and one of stabilization. At the guidance level, a driver uses information from the environment to lay out a desired path for the vehicle. At the stabilization level, he uses the actual motions of the vehicle related to the desired path to make small corrections as necessary.

First, consider the guidance level. Donges modelled steering as a "control process with the desired path as forcing function ... and the vehicle's position and attitude relative to the forcing function as output variables." On tangent sections in the absence of other traffic, this forcing function consists of a driver's desired lateral position in his lane, which as demonstrated below, is a function of the driver and the roadway cross section.

The key to handling the forcing function is a driver's perception of the road ahead. The perspective view of this road allows drivers to act in an anticipatory manner to handle forcing function changes. Because this view plays an integral role in steering, Kramer and

Rohr (1982) have suggested that it is necessary that the road ahead of a driver be incorporated into a model of steering behaviour in this form. Conventional models present the road to a driver using geometrical parameters describing the course from a bird's-eye view. The present model objectively defined the road in the conventional form. But the subjective representation used by the driver consists of the objective form modified (fuzzified) to account for the perspective from which he saw it. Thus the driver in the present model and one in that of Kramer and Rohr obtain the same information from the road ahead given its relative orientation.

On tangent sections, a driver's desired lateral position, in the absence of other traffic, is a function of a number of variables. One of these is lane width. Shelby and Tutt (1958) observed traffic on two-lane suburban highway sections in Texas. As lane width increased, the average desired lateral position moved away from the road centre line by an amount equal to about one quarter of the added width. It thereby stayed relatively close to the centre line (Figures 6.5 and 6.6). In a similar study, Taragin (1944) obtained approximately the same results. He stated that as width increased, drivers used at least two-thirds of the added portion to increase edge clearance both for passenger cars and trucks.

Drivers also adjust their desired lateral position according to shoulder type. Figures 6.5 and 6.6 show the results of both Shelby and Tutt, and Taragin for day and night, respectively. The desired lateral position moves closer to the shoulder as the type of shoulder is improved from grass to gravel to hard surface. Unlike for shoulder type, Shelby and Tutt

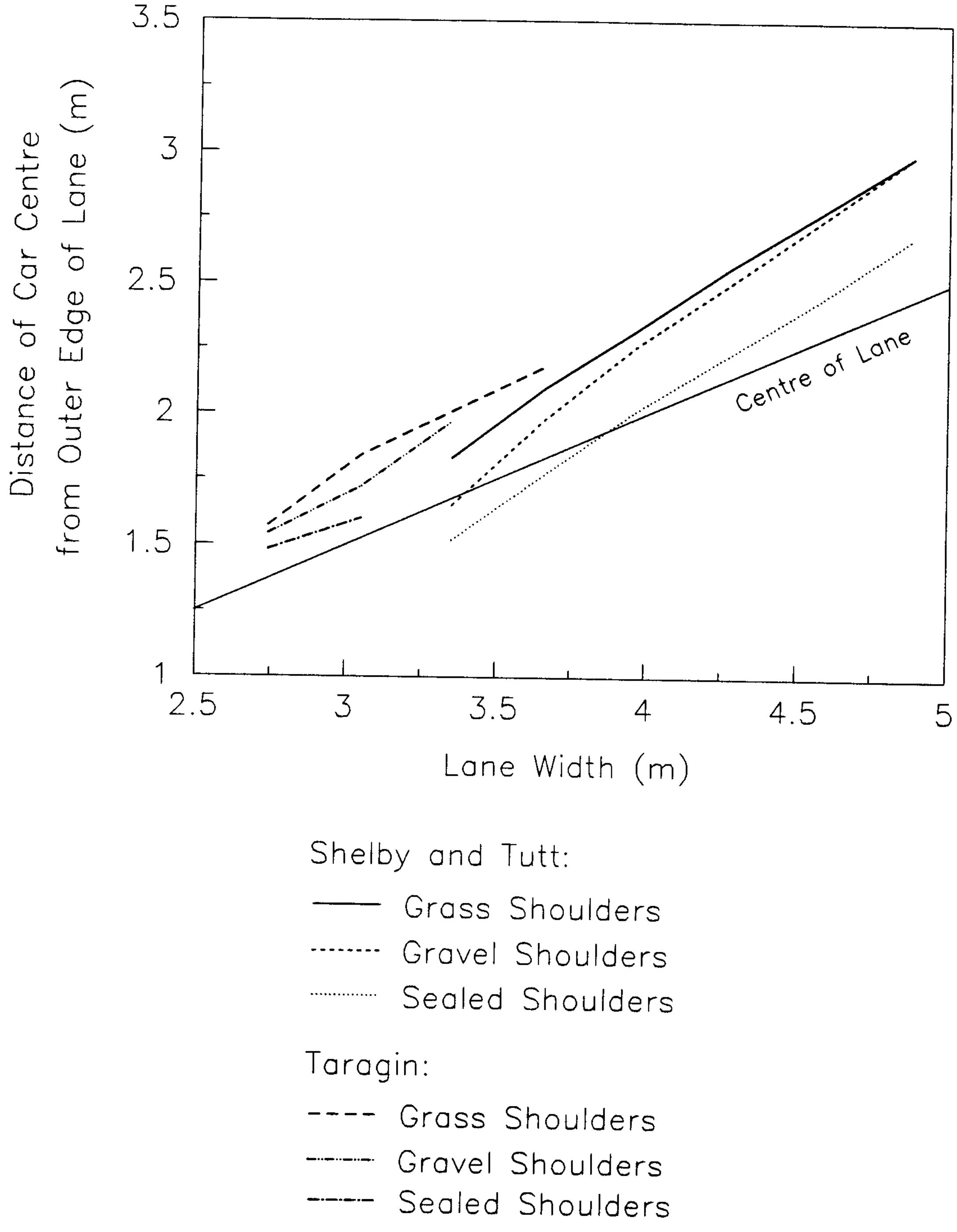


Figure 6.5 Daytime Free-moving Lateral Position

(Sources: Shelby and Tutt 1958; Taragin 1944)

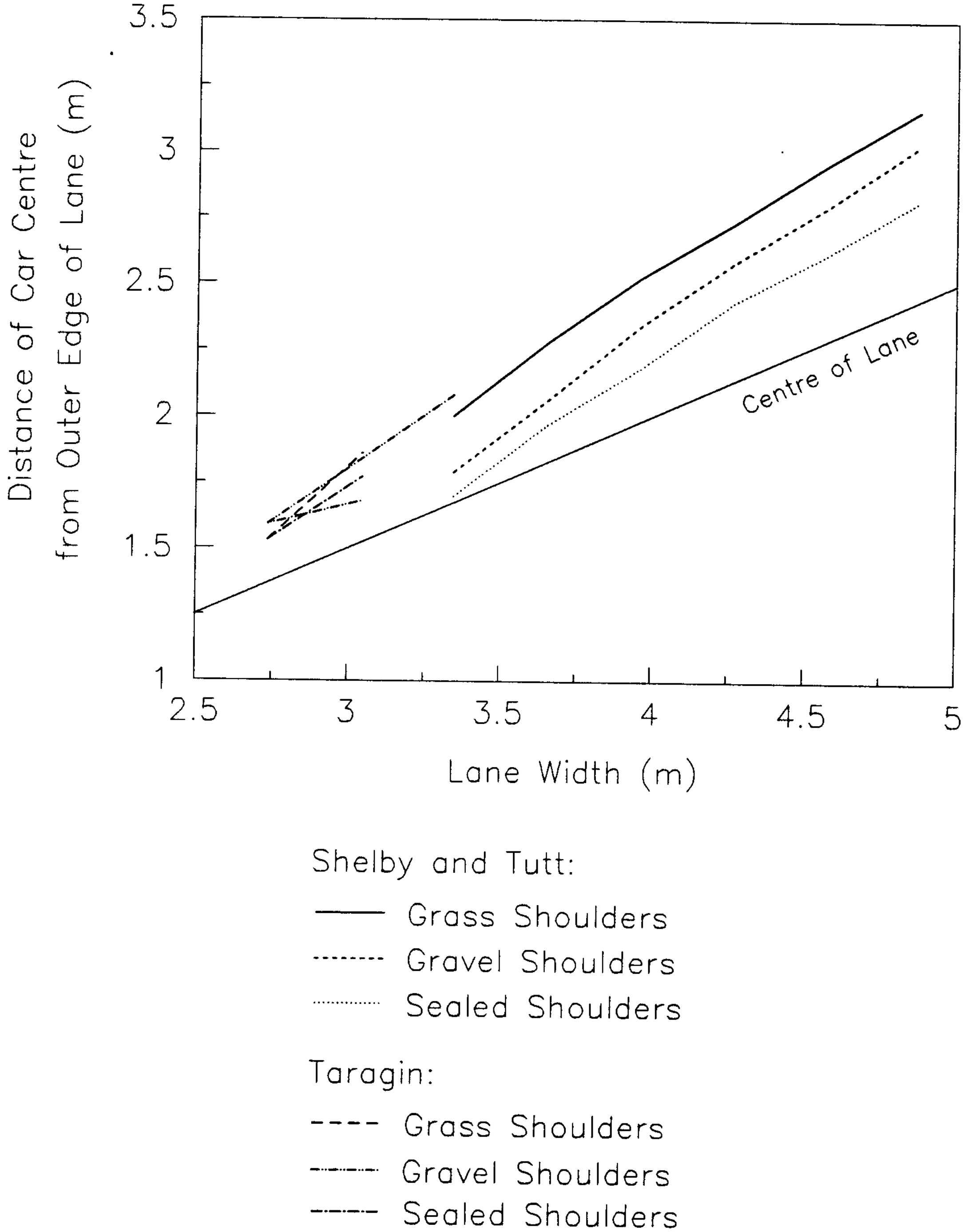


Figure 6.6 Nighttime Free–moving Lateral Position

(Sources: Shelby and Tutt 1958; Taragin 1944) concluded that shoulder widths, at least above 1 metre, had little or no effect on average lateral placements. Taragin produced similar results, suggesting that shoulder widths between 1.2 and 3.0 metres do not affect average lateral placement.

Analysis by Taragin showed no definite relationship between vehicle speed and desired lateral position, though at a few of the locations he studied, free-moving passenger car drivers who maintained higher speeds travelled somewhat closer to the centre line than slower drivers by a distance on the order of only 0.15 m.

Shelby and Tutt showed that the general trends for trucks were similar to those for passenger cars, though the former tended to drive a little closer to the edge of the pavement, especially in wider lanes, possibly because, as the authors suggested, they tried to stay out of the way of faster vehicles.

Now, consider steering stabilization. This is the process by which drivers make small corrections to align the actual path of their vehicles to their desired ones. The evidence of both Shelby and Tutt (1958), and Taragin (1944) suggest that in absence of other traffic, drivers tend to be rather lax in their maintenance of the desired lateral position. This is demonstrated by the fairly high frequency with which they drive over the centre line or onto the shoulder. The time frame for stabilization control movements is on the order of 0.5 seconds (Allen, Lunenfeld and Alexander).

At the stabilization level, Donges (1978) suggested three key variables: lateral deviation of the vehicle relative to the desired position, heading of the vehicle relative to that of the desired path, and difference between the curvatures of the vehicle's and the desired path.

At both the guidance and stabilization levels, the steering required in most driving situations is heavily overlearned and can be carried out unconsciously (Allen, Lunenfeld and Alexander). In the driver prototypes developed here, steering stabilization consisted of two instructions as shown in Figure 6.7, correcting respectively for veering right and left from the desired lane position provided by steering guidance. Guidance was modeled by duplicating these two stabilization instructions and adding extra conditions onto each pair. The example shown in Figure 6.7 is for gravel shoulders and a narrow lane width. Because there exist only average guidance and stabilization values in the literature and because open road steering is considered of lesser importance in terms of model outputs, average values were used for all driver prototypes.

6.2.3 Horizontal Curves. Drivers adjust both their speeds and steering when negotiating horizontal curves. Both involve anticipatory actions based on the view of the road ahead as described in section 6.2.2.

First, consider speed adjustment. Taragin (1954) has shown that drivers do not change their speeds appreciably after entering a horizontal curve. Rather, most of the adjustment is made on the approach to the curve. Drivers attempt to achieve a particular speed when

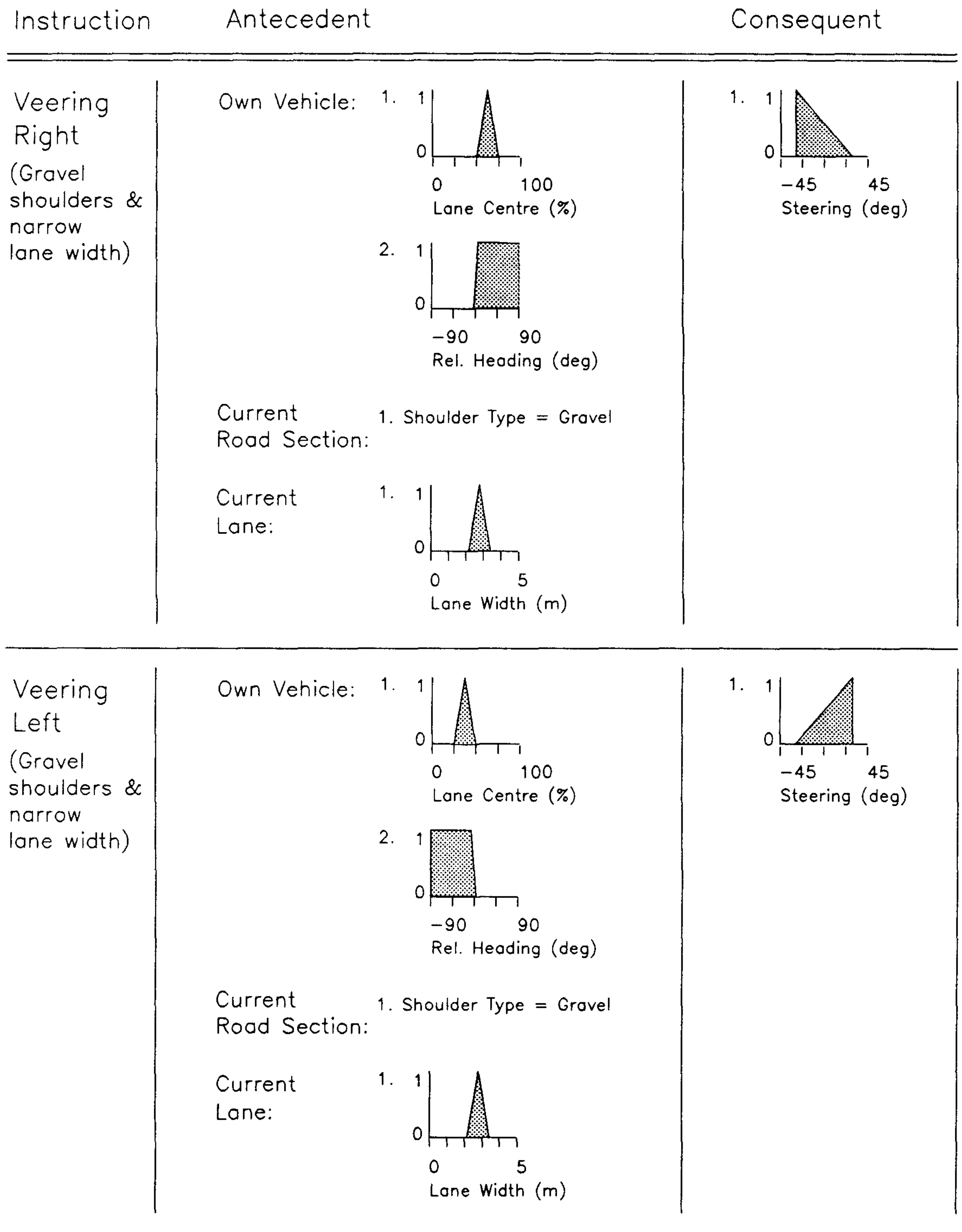


Figure 6.7 Tangent Section Steering

approaching a curve and then maintain it. This speed is primarily a function of the roadway curvature and, to a much lesser degree, of the available sight distance.

The regression analysis of McLean (1981) showed that, in addition to road curvature, desired speeds as described in section 6.2.1 also influenced horizontal curve speeds. His analysis indicated that superelevation, lane width and shoulder width had insignificant effects and that the effects of sight distance were significant but accounted for less than 1 percent of the variability in observed 85th percentile speeds. Taragin's work also showed that superelevation had no effect. McLean's resulting regression equation for 85th percentile passenger car curve speeds was:

$$V_c(85) = 53.8 + 0.464V_f - \frac{3260}{R} + \frac{85000}{R^2} \tag{6.1}$$

where $V_c(85)$ = 85th percentile car curve speed (kph)
V_f = desired speed of the 85th percentile car (kph)
R = curve radius (m)

This equation explained 92 percent of the variance of the dependent variable. Figure 6.8 shows the family of curve speed prediction relations developed by McLean. He cited other research by the Road Safety and Traffic Authority in Victoria, Australia which indicated that mean truck speeds on open sections of rural highway are typically 80 percent of mean car speeds. Based on this and Equation 6.1, he derived the following equations for mean car and truck speeds on curves:

$$\bar{V}_c = 0.833V_c(85) + 4.4 \tag{6.2}$$

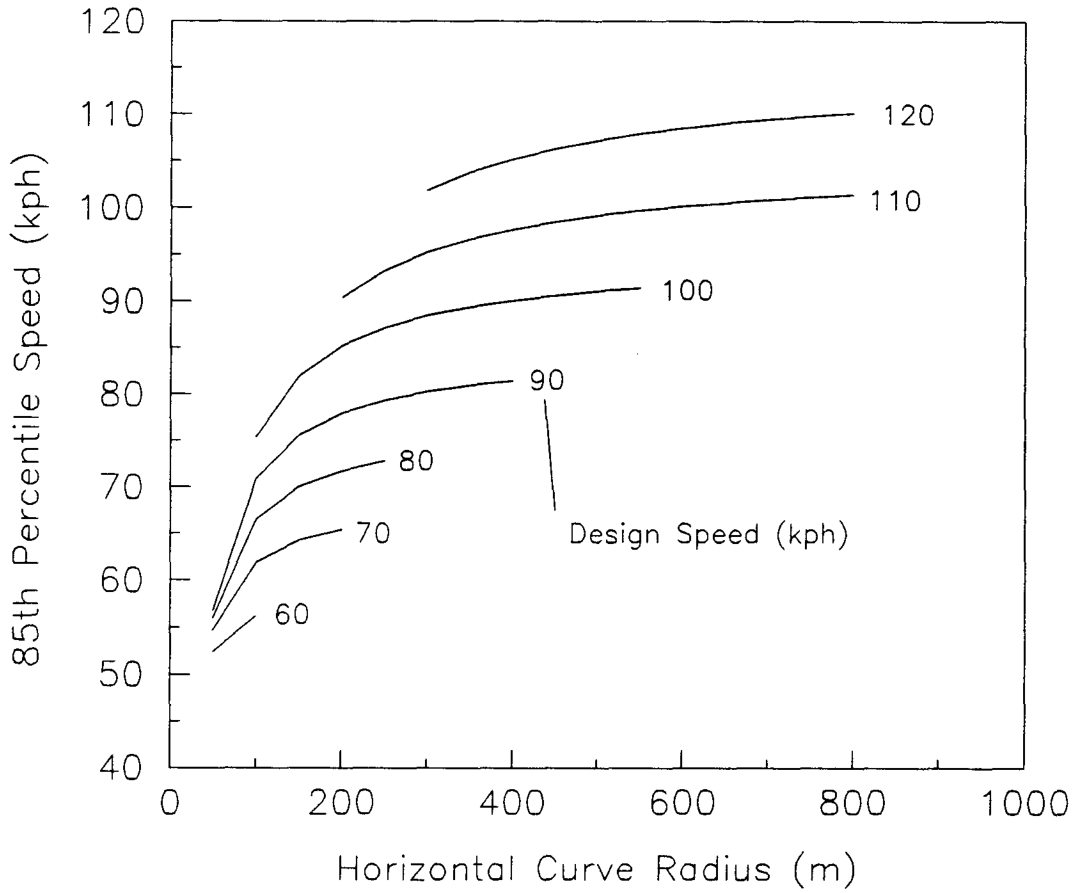


Figure 6.8 Speeds on Horizontal Curves

(Source: McLean 1981)

$$\bar{V}_t = 0.536V_c(85) + 19.0 \tag{6.3}$$

where $V_c(85)$ = 85th percentile car curve speed (kph) from Equation 6.1

Equation 6.2 explains 98 percent and Equation 6.3 explains 62 percent of the respective variances.

Based on his observations of speeds on curves in the eastern United States, Taragin (1954) concluded that speeds are not a function of the direction of curvature, left or right, when road curvatures are measured at the centre line. However, a more recent study by Lindqvist (1992) suggested that drivers tend to use a somewhat higher speed on left-hand curves than on right-hand curves. This may be because the outer lane has a lower curvature than the inner lane and drivers judge the curvature of a road section not by that of the centre line but by that of their travel lanes.

In developing the driver prototypes, an instruction for speed reduction on curves was formulated and is shown in Figure 6.9. This instruction is based on the observation that average drivers select their curve speeds so that their lateral acceleration will be less than about 3 m/s^2, and is modeled as an unconscious behaviour. The objective function across driver prototypes was developed using a safety versus risk-taking and constancy trade-off (Figure 6.2). Speed reduction in a curve is loss of the constancy afforded on tangent sections. The more a driver reduces his speed, the more constancy is lost. At the same time, reducing the speed on curves decreases the enjoyment a driver may receive from

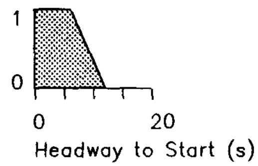
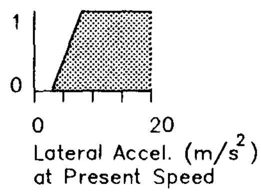
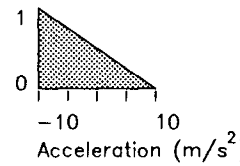
Instruction	Antecedent		Consequent
Curve Speed	Any Road Section Ahead in Own Lane:	1. Horiz. Alignment = not Straight 2.  3. 	1. 

Figure 6.9 Speed Reduction Instruction in a Horizontal Curve

higher centripetal acceleration (risk-taking). Balanced against these two is safety. The faster one drives on a curve, the greater seems the loss of control.

Now, consider steering on a horizontal curve. Godthelp (1986) examined the steering actions of drivers in negotiating curves and found that, with reference to Figure 6.10,

> [t]he driver will start the steering action with an anticipation time, T_a, before the actual curve begins (t_b). This anticipatory steering-wheel action will be finished at a short period after t_b. Then a period of stationary curve driving begins, during which the driver generates correcting steering-wheel movements. Finally, the steering wheel is returned to the central position in a period covering the endpoint of the curve (t_e). Godthelp (1986)

As described in section 6.2.2, steering can be modelled as the combination of guidance and stabilization processes acting in parallel, the latter being used to correct for inaccuracies in the former. In steering through a horizontal curve, the guidance process is active in stages 1 and 3 as shown in Figure 6.10. Godthelp's measurements indicate typical guidance inaccuracies in terms of steering wheel angle standard deviation are about 9 percent of the steering wheel angle amplitude.

The structure of curve steering instructions used in the driver prototypes is shown in Figure 6.11. As with tangent section steering, the curve steering instructions were not varied from one driver prototype to another and were considered unconscious behaviour.

6.2.4 Vertical Alignment. Though vehicle capabilities serve to limit performance on upgrades,

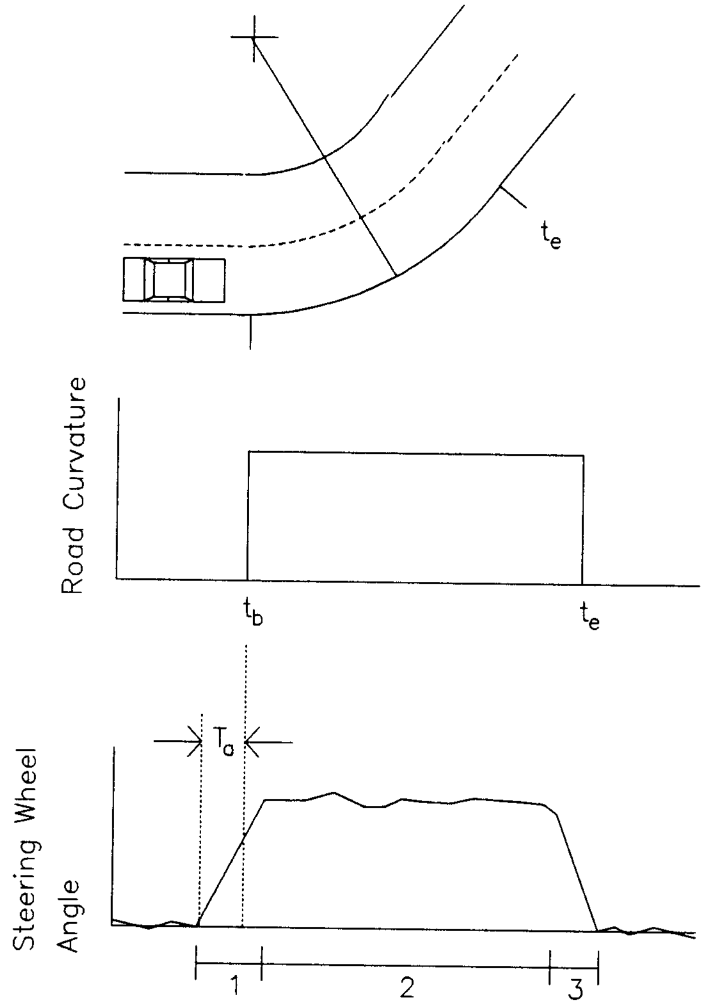


Figure 6 . 10 Steering in a Horizontal Curve

(Source: Godthelp 1986)

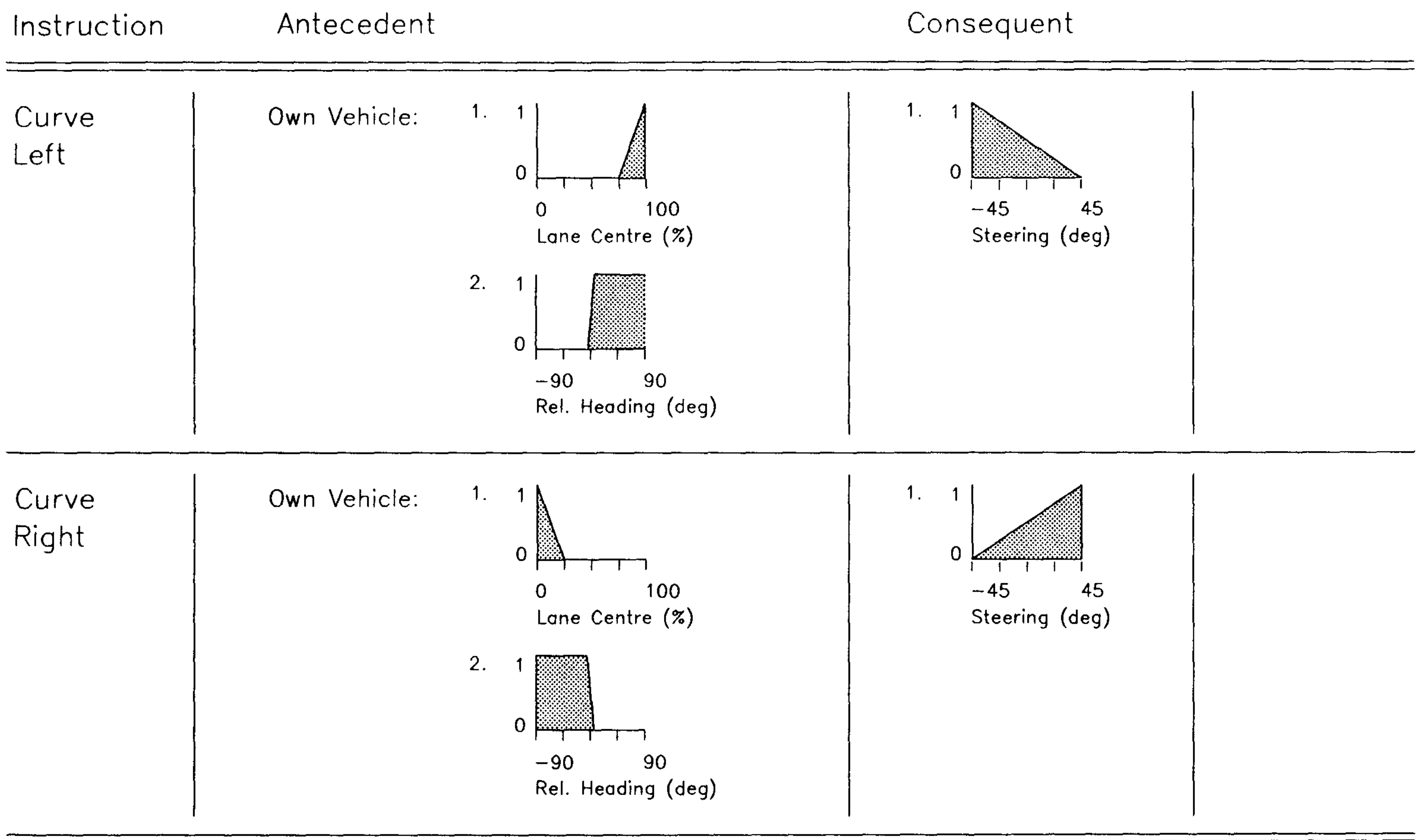


Figure 6.11 Steering Program Instructions in a Horizontal Curve

drivers also make speed adjustments for other reasons on varying vertical alignments. One of these reasons is limited sight distance. Drivers approaching a crest vertical curve which limits sight distance tend to reduce their speeds somewhat before reaching the point of minimum sight distance (Lefeve 1953). The amount of speed reduction increases as sight distance decreases as shown in Figure 6.12. The speed reduction is also a function of the approach speed such that percentage reductions for a particular sight distance tend to be more or less constant. Lefeve showed that it is not a direct result of an approach upgrade because, in his study, the speed reductions on upgrades did not vary from those on level approaches.

Though the average driver slows down a bit when approaching a crest vertical curve, he still overdrives the curve by a considerable margin in terms of stopping sight distance requirements. Though there is a tendency for average speeds on crest curves to decrease with decreasing sight distance, there is no definite relationship between minimum sight distance and speed or speed distribution. Rather, the speed distribution maintains its normal distribution found at locations with greater sight distance. From this, Lefeve concluded that "[a]pparently, the drivers feel that the reduction in speed is greater than it actually is, or individual drivers so seldom encounter critical situations on vertical curves that they are not aware of the hazard involved."

Driver behaviour when approaching crest vertical curves where sight distance is limited was modeled as shown in Figure 6.13. The objective function over driver prototypes was

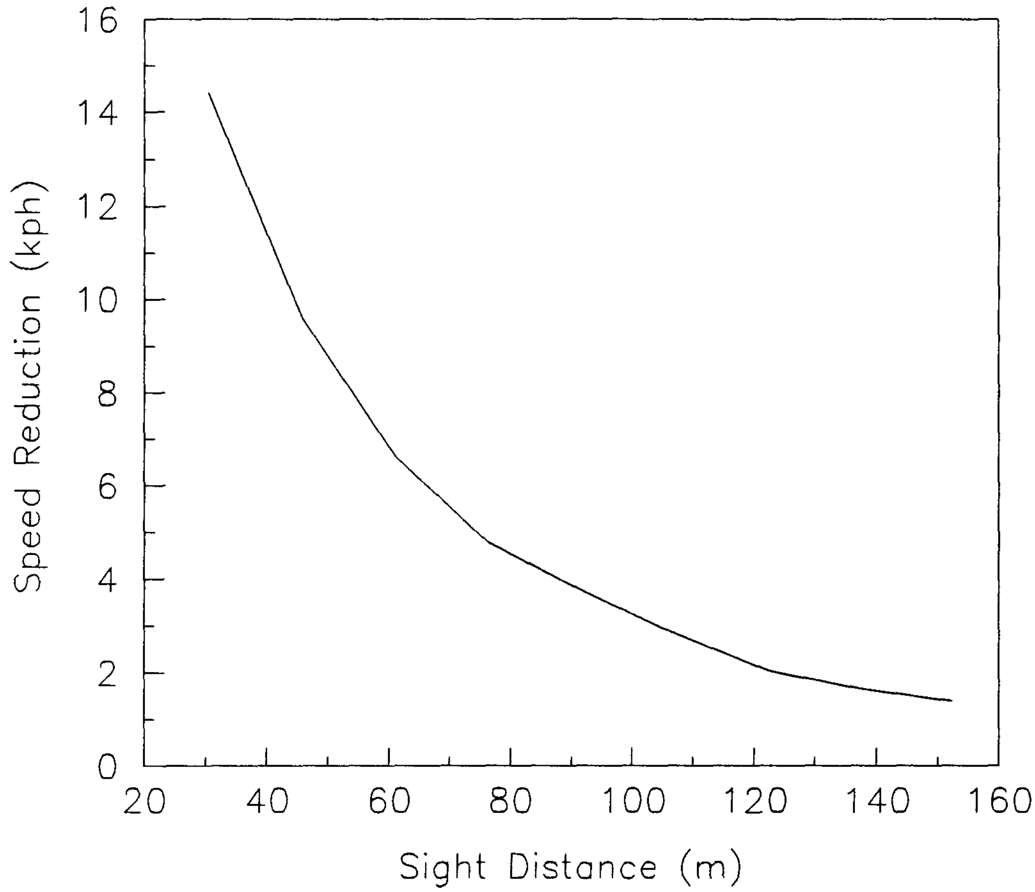


Figure 6 . 12 Speed Reduction on Crest Vertical Curve

(Source: Lefeve 1953)

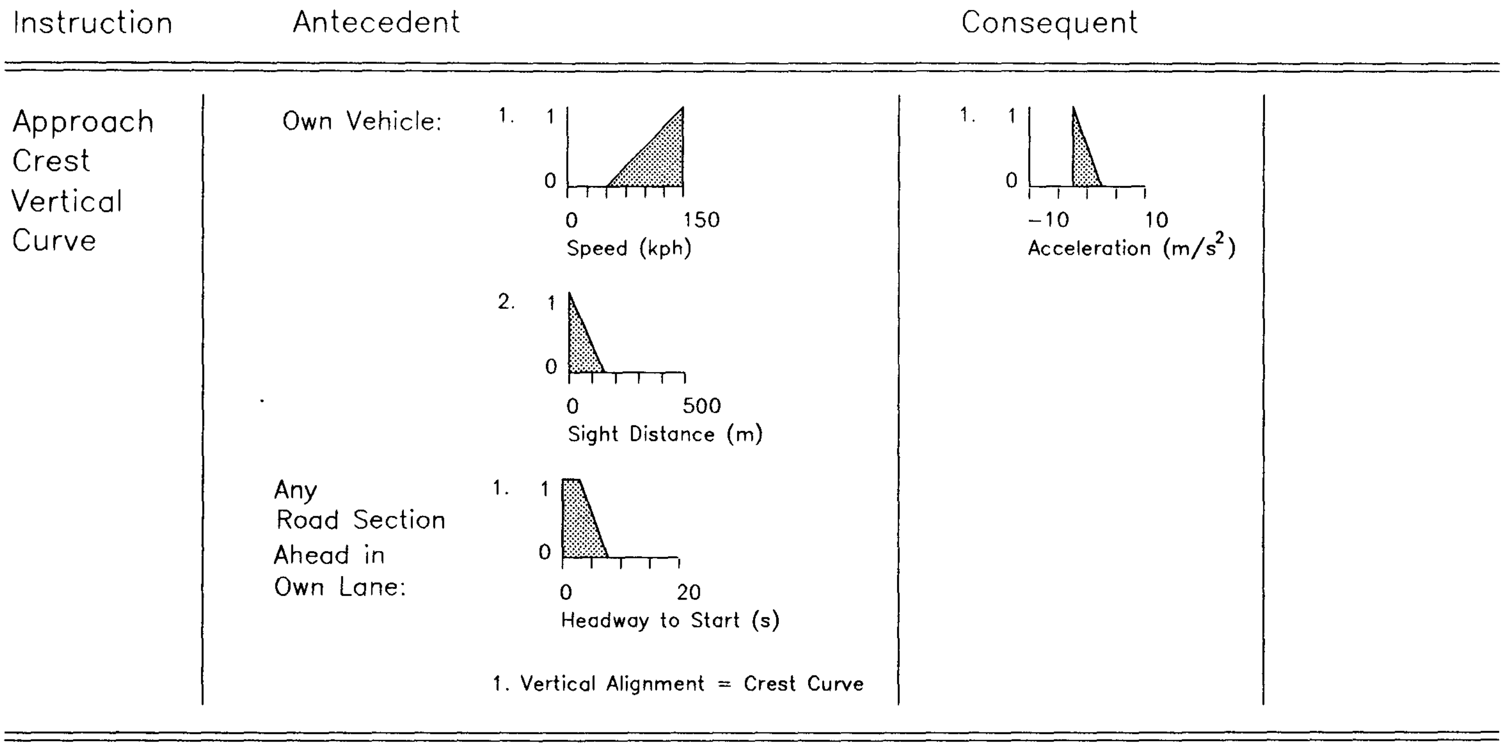


Figure 6 . 13 Crest Vertical Curve Speed Reduction Instruction

developed using a constancy versus safety trade-off. As with speed reductions on horizontal curves, speed reductions on vertical curves allow for increased safety (although not nearly sufficient as already shown) at the expense of constancy. It is argued that drivers do not perceive speeding over a crest curve as a risk-taking venture because of their perception that safety is compromised to only a small degree. This is considered an unconscious behaviour.

Except for their behaviour when approaching a crest vertical curve, most drivers on upgrades attempt to maintain their desired speeds and performance is limited mainly by vehicle capabilities. On downgrades, drivers of passenger cars behave in the same manner, typically allowing their vehicles to accelerate to the upper part of their desired speed ranges before braking. The drivers of commercial vehicles, on the other hand, must limit their speeds on long or steep grades to prevent overheating of their brakes.

As a truck descends a grade, its potential energy must be converted into kinetic energy and heat, the first by virtue of the truck's speed and the second by virtue of the truck's brakes. To maintain control, a driver cannot just let grade acceleration convert all potential energy into kinetic energy. Therefore, the brakes must be applied to limit speed to a safe value. The power demand on brakes to do this is a function of the truck speed (Fancher and Winkler 1983) as follows:

$$P_B = \frac{WGV}{367} - P_N - P_R \qquad (6.4)$$

where P_B = brake power (kW)
W = vehicle weight (kg)

G = grade (rad)
V = vehicle speed (kph)
P_N = natural retardation from aerodynamic drag and rolling resistance (kW)
P_R = power dissipation by retarders

Over the length of a grade, brake temperature rises as a function of the energy dissipated, which equals the product of brake power and the time taken to drive down the grade. This time is directly proportional to the length of the grade and inversely proportional to the speed, so for a particular slope and length of grade, there will be a range of operating speeds within which brake energy dissipation is minimised and brake temperature is held within acceptable limits. The optimum speed is the highest in this range.

It has thus been shown that the optimum speed for a commercial vehicle on a downgrade is a function both of slope and grade length. However, field evidence suggests that drivers, in the absence of aiding signs, adjust their speeds only for the perceived slope such that they tend to driver slower than necessary on moderate grades and too fast on severe grades (Fancher and Winkler 1983). They tend not to account for grade length. In their model of grade effects on traffic flow, St. John and Kobett (1978) used the following approximation of commercial driver behaviour on downgrades:

$$V_{crawl} = -\frac{321}{G} \qquad (6.5)$$

where V_{crawl} = mean downgrade crawl speed (kph)
G = grade (percent)

They used a standard deviation of 11 kph and noted that commercial vehicle speeds were normally not influenced unless the grade was at least 1.5 km long and had a slope in excess of 4 percent.

The program for commercial drivers on a downgrade was modeled as a conscious behaviour because of the grade length and slope judgement required. Figure 6.14 shows the starting conditions for the program and the instructions are shown in Figure 6.15. The objective function for this behaviour was developed using a safety versus constancy and risk-taking trade-off.

6.2.5 Surface Effects. The road surface can have a marked effect on drivers' speed choices. Average desired speeds tend to decrease as the road surface changes from high to low quality. The highest average desired speeds are found on concrete highways. Those on asphalt highways tend to be 5 to 6.5 kph lower. On gravel surfaces, desired speeds are some 15 kph lower than on concrete highways. And, on unsurfaced roads, they are a further 8 kph lower than on gravel roads (Oppenlander, 1966).

For a given surface, though, a decrease in surface friction resulting from rain causes very little decrease in average speeds. Stohner (1956) concluded that, for concrete and asphalt highways, "drivers of free-moving passenger cars operated at about the same speed on wet pavements as on dry pavements." This conclusion was based on his studies of straight sections and curves with a radius greater than 200 metres.

Any
Road Section
Ahead in
Own Lane:

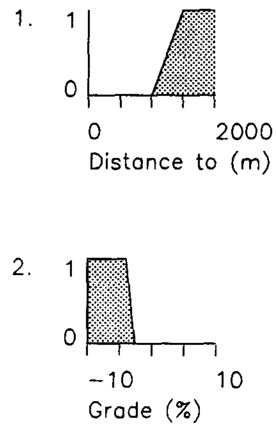


Figure 6.14 Commercial Vehicle Downgrade Program Starting Conditions

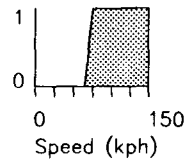
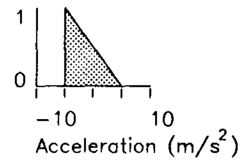
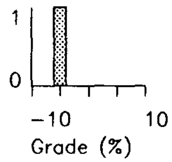
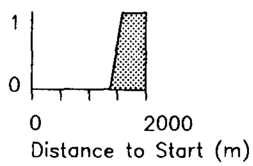
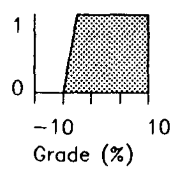
Instruction	Antecedent		Consequent	Flags
Speed on 5% Downgrade	Own Vehicle:	1. 	1. 	
	Current Road Section:	1. 		
... Similar Instructions for 7, 9 and 11 percent grades ...				
Check if Nearing End of Grade	Any Road Section Ahead in Own Lane:	1. 		Start Required
		2. 		

Figure 6 . 15 Downgrade Program for Commercial Vehicles

The driver prototypes developed here were for asphalt surfaces exclusively.

6.3 The Driver and Road Signs

The most important factors in the perception and use of traffic signs are not the visibility, size or conspicuity of the signs but the motivation and attention of the drivers.

> Several studies have shown that when requested drivers are able to detect and identify traffic signs quite effectively, at least on an open highway ...Thus drivers seem to have perceptual skills sufficient for the adequate use of traffic signs...it is known, however, that drivers, when stopped on the road, report the signs which they have just passed inadequately. Furthermore, these studies revealed substantial differences between the signs in this 'registration capability'. If 78% of drivers reported a speed-limit sign correctly, 63% a special police-control sign and 55% a warning of damage in the road surface, the percentages for a general warning and a pedestrian crossing ahead were only 18 and 17 respectively.
>
> It is likely these differences are due to motivational aspects of the signs. The registration rank order indicates the significance of each sign to the drivers very clearly and, on the other side, it does not correspond to the perceptual rank order in traffic sign conspicuity measured in tachistoscopic studies... (Summala and Hietamaki 1984)

It has been demonstrated that the effectiveness of a sign is much more dependent on the importance of its message to a driver than on its physical characteristics. Johansson and Backlund (1970) studied drivers reactions to six different physically similar signs yielding the results quoted above. Their analysis showed that sign registration efficiency was dependent primarily on the sign message and the driver, specifically his degree of familiarity with the road. Other factors had no significant effect, although, for signs considered less important by drivers, registration efficiency appeared to increase as road conditions worsened, decrease with poor visibility and decrease with increased traffic density. These effects were not observed for signs considered important.

The Johansson and Backlund study consisted of stopping drivers after they passed a sign and quizzing them on their recollection of it. Some of the stopped drivers knew about the study and some of these had been actively looking for the signs. These drivers were considered to be more motivated to look for the signs and had higher registration efficiencies for all the signs, but these efficiencies still showed the same sign importance effects, though to a lesser degree (Figure 6.16).

Shinar and Drory (1983) noted that drivers paid considerably more attention to directional signs than warning signs and that direct fixations on warning signs are rare. They advanced the following explanation of this poor sign registration performance by drivers:

> [I]n daylight driving on an open road, drivers can obtain most of the information contained in warning signs from their perspective view of the road ahead. Therefore, only the information that they consider relevant *and which they are unable to obtain directly from the roadway* is registered...
>
> The implication of the preceding explanation is that sign registration performance could almost paradoxically, improve under conditions of degraded visibility, when information that would otherwise be directly retrieved from the roadway and adjacent areas is unavailable. (Italics added)

The study by Shinar and Drory used a method similar to that of Johansson and Backlund. They employed a different set of signs (stop ahead, side road, winding road and general warning) and obtained lower registration efficiencies, averaging 4.5 percent during the day and 16.5 percent at night. Their analysis showed that the difference between day and night

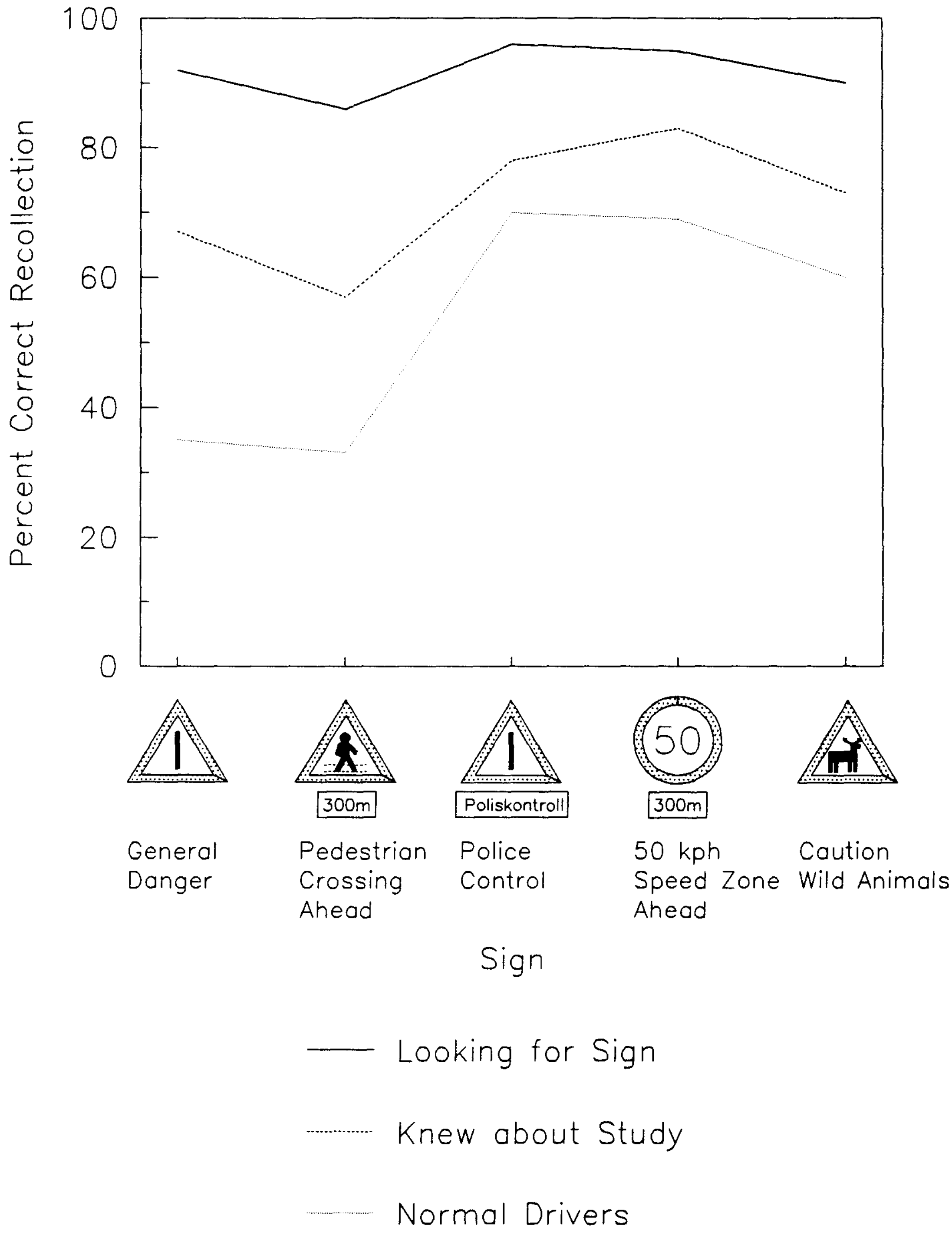


Figure 6 . 16 Conscious Attention to Road Signs

(Source: Johansson and Backlund 1970)

results was significant and opposite to what one might expect given visibility differences between night and day. An explanation is that at night, drivers shift more attention to road signs because of their decreased ability to see the road ahead. In terms of the Johansson and Backlund study, their motivational levels increase and consequently so do their registration efficiencies.

Shinar and Drory also examined the effect of the road on the approach to a road sign and found that there was no significant difference in driver sign registration between approaching on a demanding windy road and approaching on a long straight one. In both cases, it seems, drivers obtained the same amount of information from their perspective view of the road and therefore resorted to looking at signs to the same extent.

6.3.1 Reactions to Some Particular Signs. The Council on Uniform Traffic Control Devices for Canada (1985) classifies road signs as regulatory, warning or information signs. Regulatory signs include those for right of way control (such as stop and yield signs) and those for road use control (such as maximum speed signs and turning signs). Generally, they require a specific behaviour in all situations and are considered the most important by most drivers.

As an example of the driver prototype programs developed for regulatory signs, consider the stop sign. Zwahlen (1988) examined driver speed behaviour on the approach to a stop sign controlled intersection on a rural highway. He varied the conditions using a 2x2x2 matrix.

The three conditions were (a) day versus night, (b) with a stop ahead sign versus without one, and (c) good approach visibility versus approach visibility obscured by an intervening hill. For the obscured visibility condition, the stop ahead sign gave an early warning to drivers. All of the drivers were unfamiliar with the road. The results are shown in Figure 6.17. From them, one can note that the stop ahead sign had negligible effect during the day but served to slow vehicles down on the approach at night. Also, it is noted that speeds were highest on approaches at night without the stop ahead sign. Under these conditions, reduced visibility meant that the drivers did not become aware of the intersection as soon as their daytime or forewarned (by a stop ahead sign) counterparts.

Zwahlen also compared the results of familiar and unfamiliar drivers on the approach to an intersection and found that familiar drivers travelled about 3 to 6.5 kph faster than unfamiliar drivers both at night and during the day. This suggests that unfamiliar drivers maintained lower than normal speeds because of their wariness of the unfamiliar roadway.

The program for stopping at a stop sign, like those for most regulatory signs, is considered a conscious one, the starting conditions of which are shown in Figure 6.18. The actual program is shown in Figure 6.19 and contains vehicle following instructions which ensure that the driver attempts to stop his vehicle behind one he is following.

The second category of signs are warning signs, including those for physical conditions and those for traffic regulations ahead. Summala and Hietamaki examined the reactions of

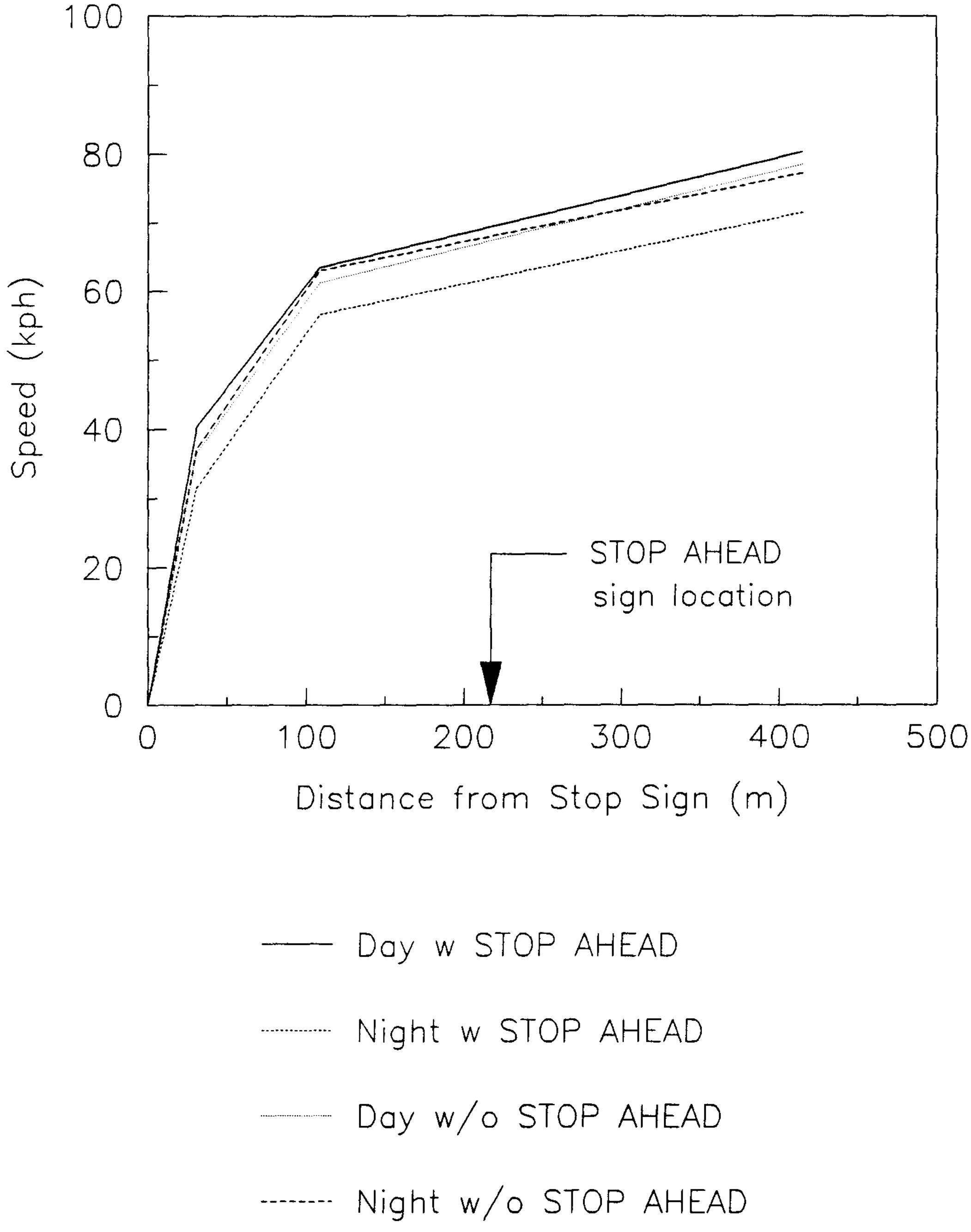


Figure 6 . 17 Speeds on Stop Sign Approach

(Source: Zwahlen 1988)

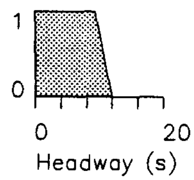
Any Sign Ahead:	1. Sign Type = RA-1 2. 

Figure 6.18 Starting Conditions for Stopping at a Stop Sign

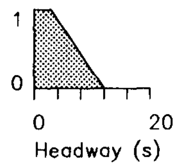
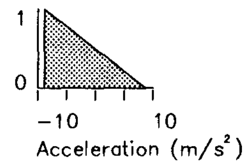
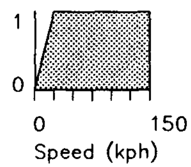
Instruction	Antecedent		Consequent	Flags
Vehicle Following Instructions (Figure 6.26) ...				
Stop at Stop Sign	Any Sign Ahead:	1. Sign Type = RA-1 2. 	1. 	Stop Required
	Own Vehicle:	1. 		

Figure 6 . 19 Stopping at a Stop Sign Program Instructions

drivers to 30 kph speed zone (a regulatory sign), and to danger and children crossing signs (warning signs) on a roadway with an average travelling speed of 50 kph. The signs were placed so that they became visible only 30 metres or about 2 seconds before their location to control for reaction time. For all the signs, the average driver slowed down somewhat and then in short order resumed his initial speed (Figure 6.20). There was a speed reduction for the control as shown in Figure 6.20 because the measurements were taken on the upgrade of a crest vertical curve. The speed reduction was the greatest for the speed limit sign (the regulatory one) and lowest for the danger sign. It did not seem to be a function of the initial travel speed.

By measuring the performance of a test vehicle reproducing observed decelerations on the hill, Summala and Hietamaki were able to conclude that for all the signs studied, most drivers merely released their accelerators slightly on approach, perhaps while evaluating the urgency of the sign message and then accelerated to resume speed beyond it. They cited similar results in another of their studies in which signs were viewed from a longer distance.

An explanation of the rank order was offered as follows. The danger sign had the smallest effect possibly because it is too ambiguous and without supporting evidence on the road ahead, it was all but ignored. The children crossing sign produced a slightly higher speed reduction but it, too, is not perceived as an important sign because most of the time, there are no children on the road. The speed limit sign evoked the largest response possibly because neglecting it could incur a fine.

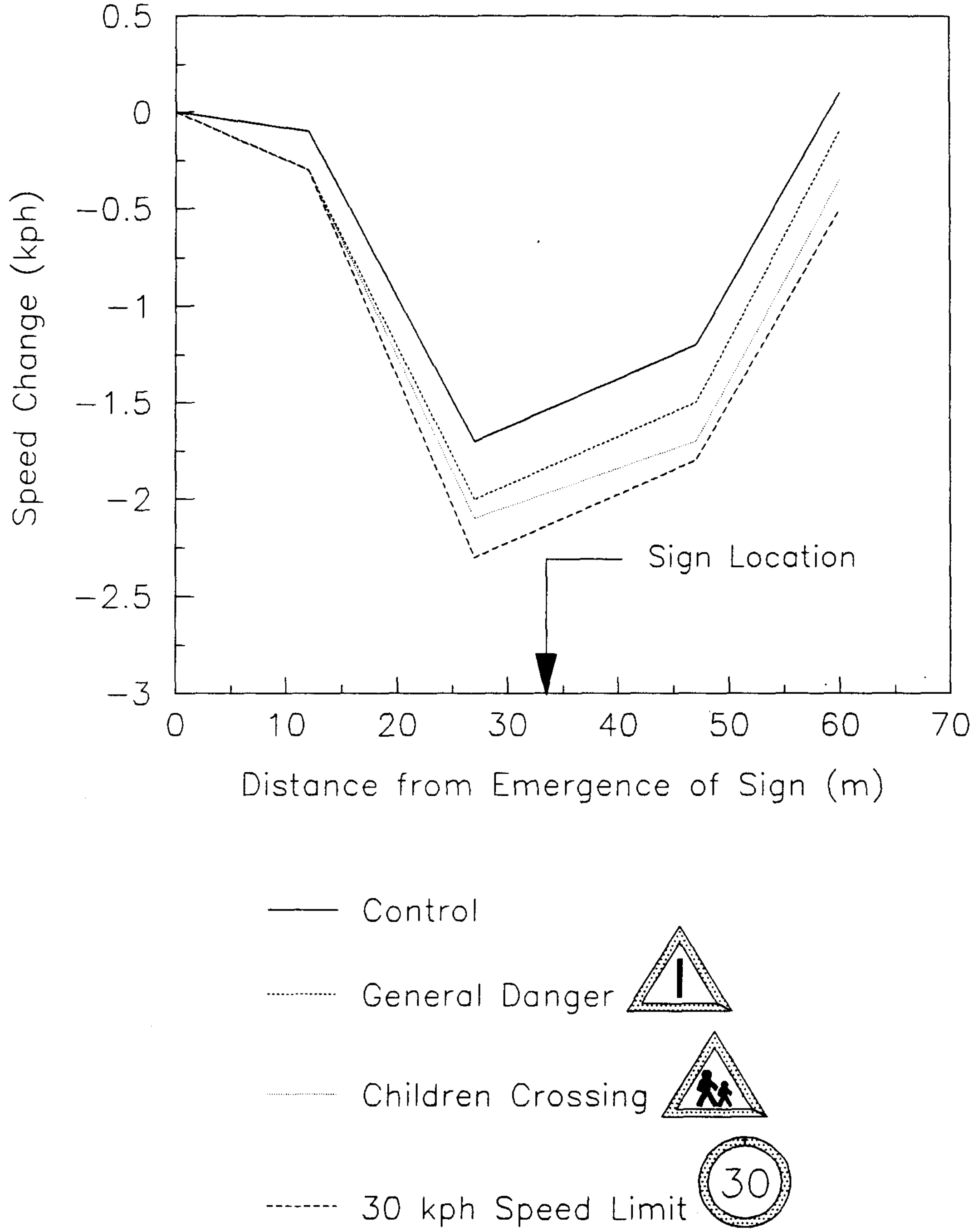


Figure 6.20 Driver Reaction to Selected Signs

(Source: Summala and Hietamaki 1984)

Because of the low conscious recognition of warning signs, the driver prototype instructions for dealing with these signs were all modeled as unconscious behaviour. Because drivers do not respond to the messages of these signs without some verification of the signs' information from another source (the road), the modeled responses to these signs consisted of momentarily suppressing positive acceleration and directing attention to where information supporting the sign's message may be found. Assuming the supporting information is there, drivers then react to that information using other (appropriate) programs. In this way, the function of warning signs is to help the driver direct his search for important information in the environment rather than prompt any action because of their messages. A typical program is shown in Figure 6.21.

Information signs, the third class as defined by the Council on Uniform Traffic Control Devices for Canada, are used in higher order navigation in driving. As discussed at the beginning of section 2, this is outside the scope of this model. For the purposes of the present model, time spent studying navigation signs or planning routes is considered a distraction from the situational level of driving and is included as such.

6.3.2 Reactions to Flashing Beacons. The importance of road signs is a function of the experience a driver has had with them in the past. Signs which 'cry wolf' too much are soon ignored. Also, people notice novelty in their environment. These two principles mean that the effectiveness of certain traffic signs and signals can be enhanced by their rarity and stringent use. This is the source of the effectiveness of flashing beacons, which signify

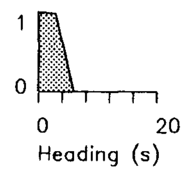
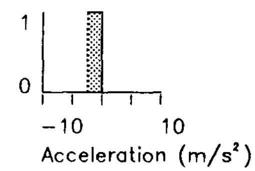
Instruction	Antecedent		Consequent
Direct Focus and Cap Acceleration	Any Sign Ahead:	1. Sign Type = WA-2R 2. 	1.  2. Focus = Road Ahead

Figure 6.21 Instructions for Responding to Curve Sign

enhanced danger and usually carry through on their promises.

Zwahlen (1988) cited study results which demonstrated that flashing beacons placed on the approaches to intersections caused a reduction in both the average approach speed and speed variance. The effectiveness of the beacons was diminished though if drivers again encountered them at a downstream intersection.

Summala and Hietamaki (1984) studied the effects of flashing beacons in conjunction with warning signs along with the effects of warning signs alone (the latter having been described in section 6.3.1). They found that flashing beacons elicited a greater speed reduction than warning signs alone, except under sunny conditions when they were not as visible. The importance rank order of the signs remained the same both with and without the beacons.

6.4 The Driver and Traffic

This section describes the interactive behaviour among drivers.

6.4.1 Opposing Traffic. On a two lane highway, opposing traffic, because of its proximity, can have an effect on the lateral position chosen by drivers. Drivers seek to reduce the risk of a head-on collision either consciously or, more likely, unconsciously. They can do this either by lowering their speeds or by moving away from the roadway centre line. The average driver does the second of these first, perhaps because it requires less effort, and only reduces his speed if the lane is too narrow to allow movement sufficiently away from the path of oncoming vehicles.

Shelby and Tutt (1958) showed that the average desired lateral position of drivers increased with increasing opposing traffic volume, regardless of lane width, shoulder width and shoulder type. Taragin (1944) had similar results, but made a further analysis based on whether vehicles were free-moving (unaffected by traffic opposing or ahead) or meeting oncoming vehicles (only unaffected by vehicles ahead) at a particular time. His results showed that the average desired lateral position of drivers tended to move toward the pavement edge when meeting an oncoming vehicle. Figure 6.22 shows the effect of traffic volume on the desired lateral positions of free-moving and meeting vehicles where one direction carries significantly more traffic than the other. Figure 6.23 shows a similar graph for the situation where the traffic volumes are of similar size. Based on these, Taragin concluded that as opposing traffic volumes increased and the percentage of time drivers spend as free-moving consequently decreased, "the drivers shifted their vehicles from the free-moving position to the meeting position more frequently ... until a condition was reached when the average driver found it more convenient to remain in a position almost coinciding with his position when meeting other vehicles." Thus, as shown in Figures 6.22 and 6.23,as traffic volumes increase, the desired lateral positions of free-moving and meeting vehicles come closer together.

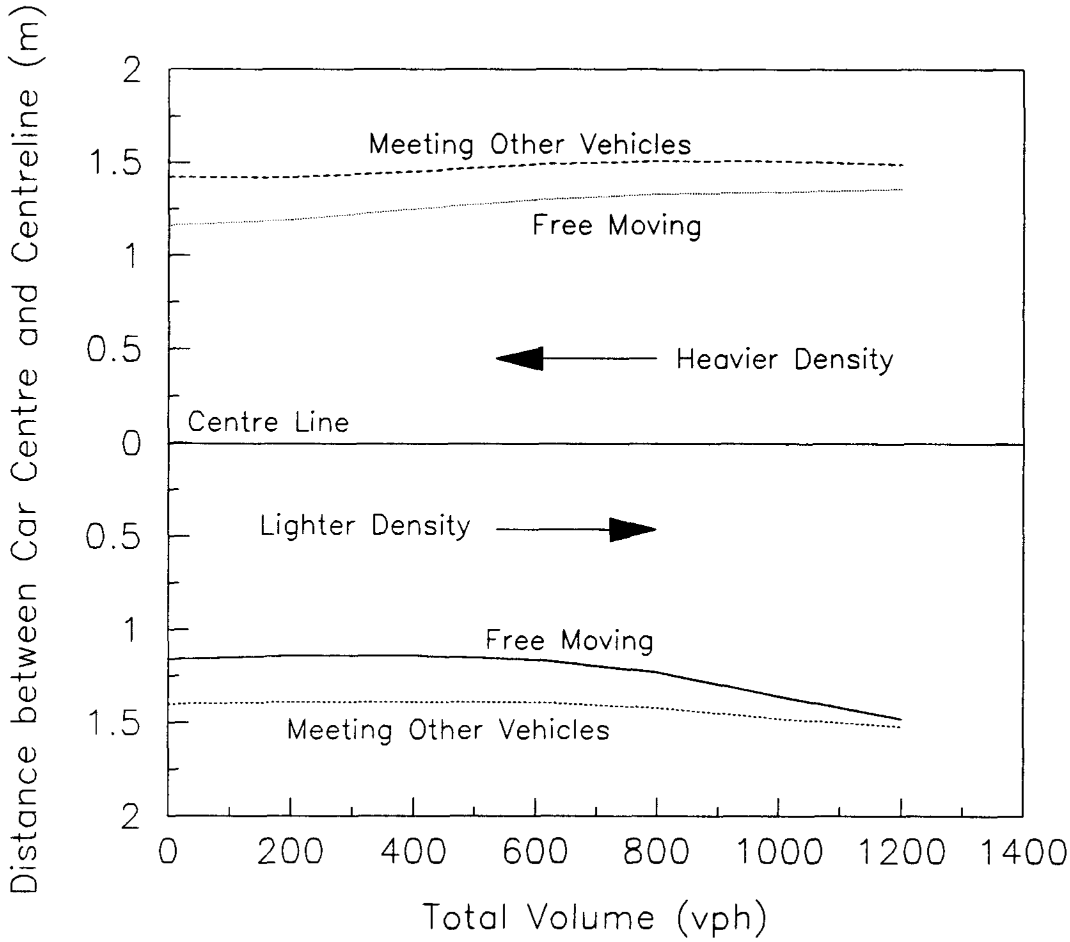


Figure 6 . 22 Lateral Position and Unbalanced Traffic

(Source: Taragin 1944)

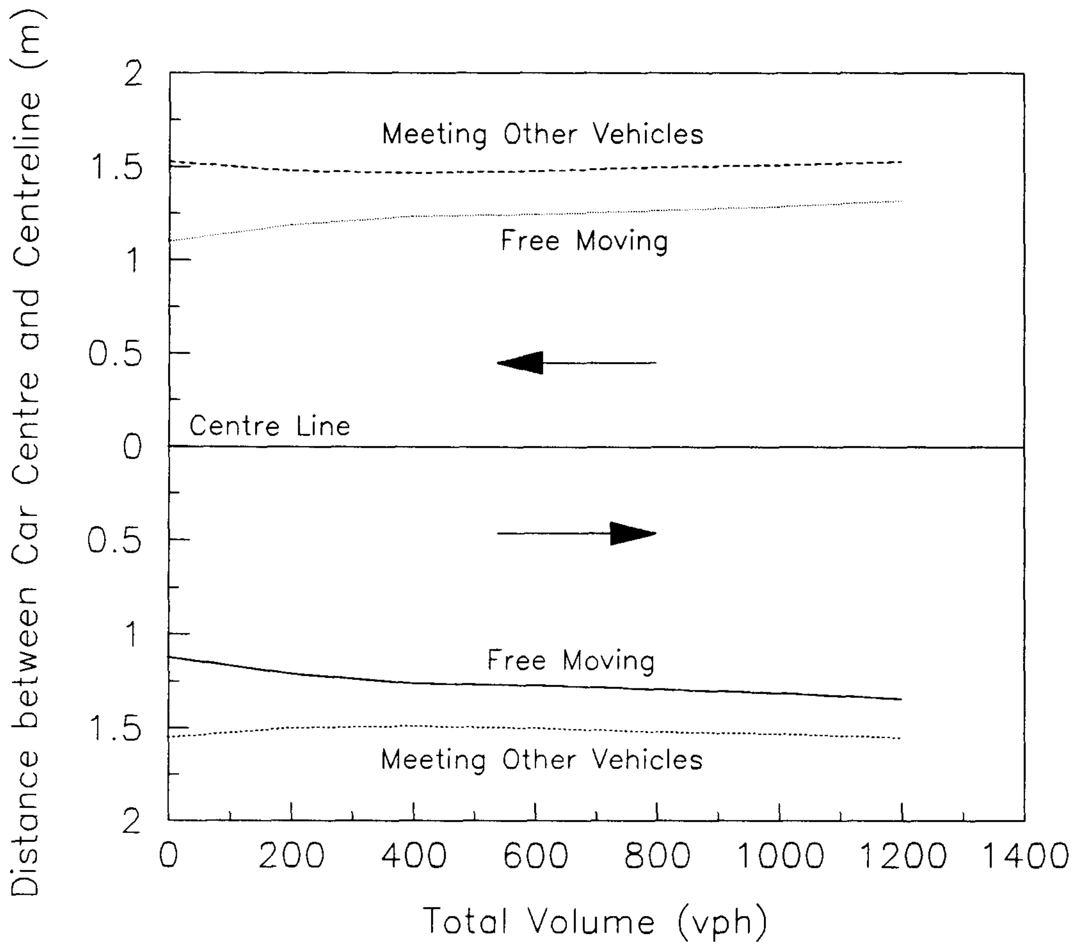


Figure 6 . 23 Lateral Position and Balanced Traffic

(Source: Taragin 1944)

6.4.2 Vehicle Following. Vehicles in a stream of traffic are distinguished as being either in a free flow or in a car following situation. The distinction is that in the latter situation, the behaviour of the driver behind is affected by the actions of the driver ahead. Definitions of a following vehicle are usually given in terms of the time headway separating it from the vehicle ahead of it, and values range from 2 to 5 seconds (Batz 1989).

When in the car following situation, the distance at which a particular driver chooses to follow another vehicle is almost entirely a function of the speeds at which the vehicles are travelling. Such variables as duration of driving, day versus night, and the actions of other nearby traffic seem to have no effect (Lerner, Abbott and Sleight 1964). Figure 6.24 shows typical average following distances as a function of speed.

While the desired following distance of a particular driver is mostly a function of speed, there is considerable variation among drivers. Field measurements have shown that at 55 kph, the mean following distance is about 36 metres with a standard deviation of 19 metres. At 90 kph, it is about 54 metres with a standard deviation of 30 metres (Lerner, Abbott and Sleight 1964). Colbourn, Brown and Copeman (1978) suggested that experienced drivers tended to maintain lower following distances than inexperienced drivers especially at higher speeds.

While there is considerable variation in following distances among drivers, there is relatively little within-driver variability at a given speed. Much of that variability, which increases with

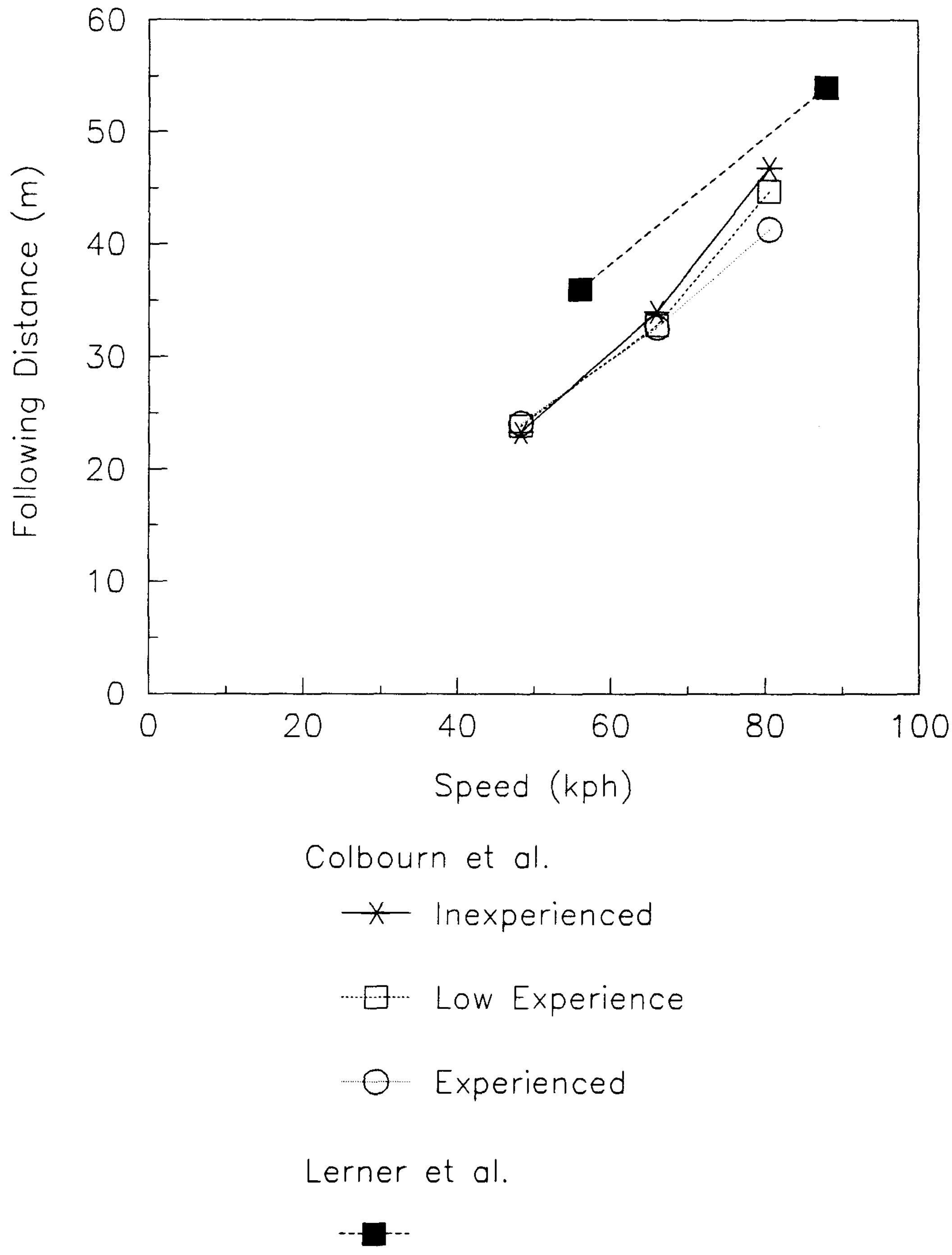


Figure 6.24 Average Vehicle Following Distance

(Sources: Colbourn, Brown and Copeman 1978;
Lerner, Abbott and Sleight 1964)

increasing speed, is a function of the difficulty of maintaining a constant headway. Colbourn et. al. obtained within-driver standard deviations as shown in Figure 6.25 for experienced, low experience and inexperienced drivers.

Lerner et. al. also studied the effects of the actions of a third car on the following distance chosen by a driver. Tail-gating by the third car caused the following driver to be annoyed but did not prompt him to speed up and reduce his following distance. Cutting in by the third car caused the following driver to fall back but he assumed a following distance behind the car that cut-in similar to his original one and did not increase the spacing, as Lerner et. al. had hypothesized.

In adjusting to a comfortable following position, drivers use varying accelerations and decelerations. Algea (1964) suggested that they increase with increasing relative velocity and decrease with increasing headway.

For most drivers, vehicle following is considered an unconscious behaviour. Its instructions in the average driver prototype are shown in Figure 6.26. In developing the objective function for this behaviour, the assumed trade off was risk-taking versus safety and constancy. One the one hand, driving close behind another vehicle can heighten stimulation and excitement. One the other, it is safer to drive at a longer following distance to allow time to react to the lead drivers actions, and following at a longer distance means the following driver need not react to every perturbation in the lead driver's speed and therefore

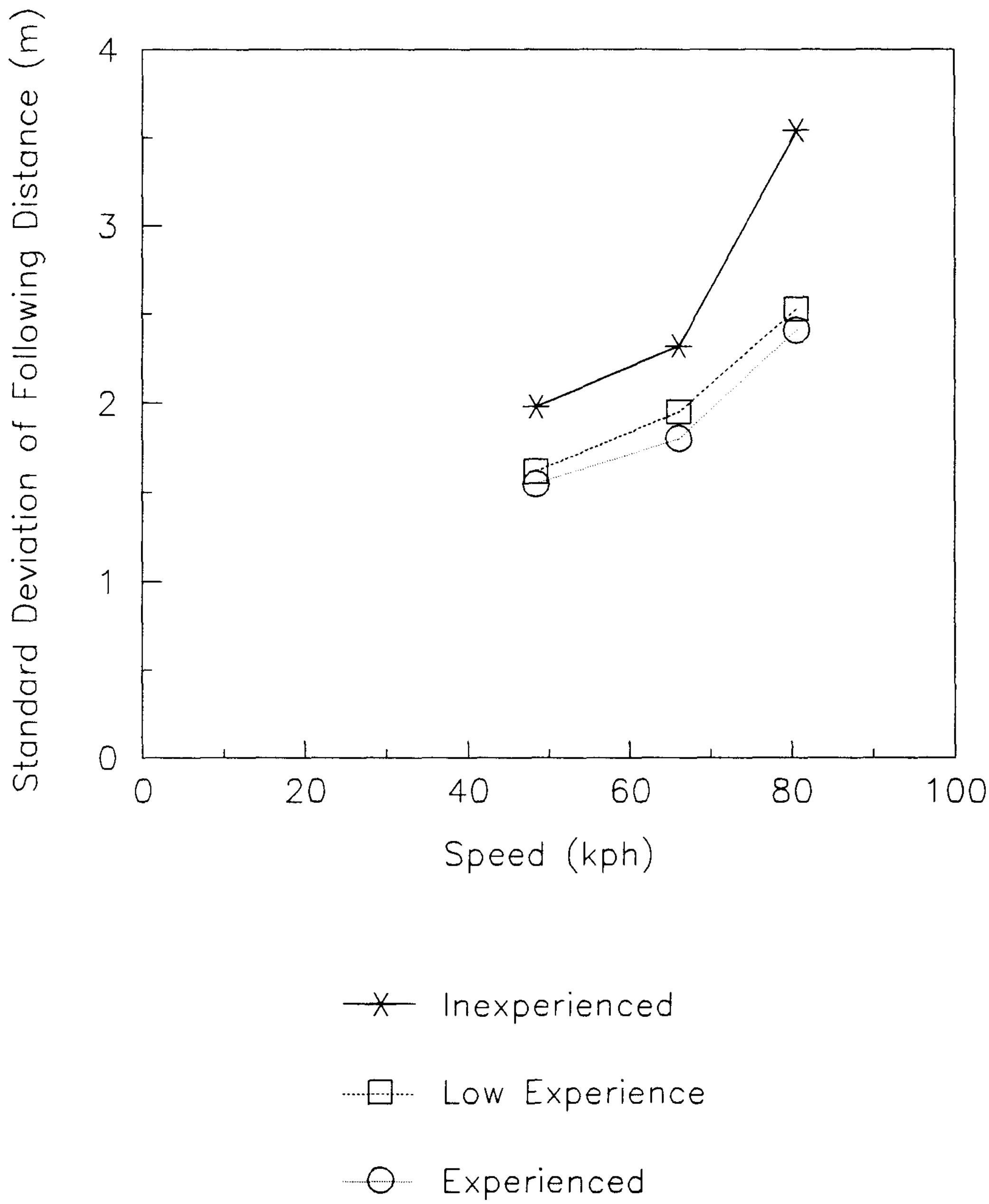


Figure 6.25 Vehicle Following Distance Variability

(Sources: Colbourn, Brown and Copeman 1978)

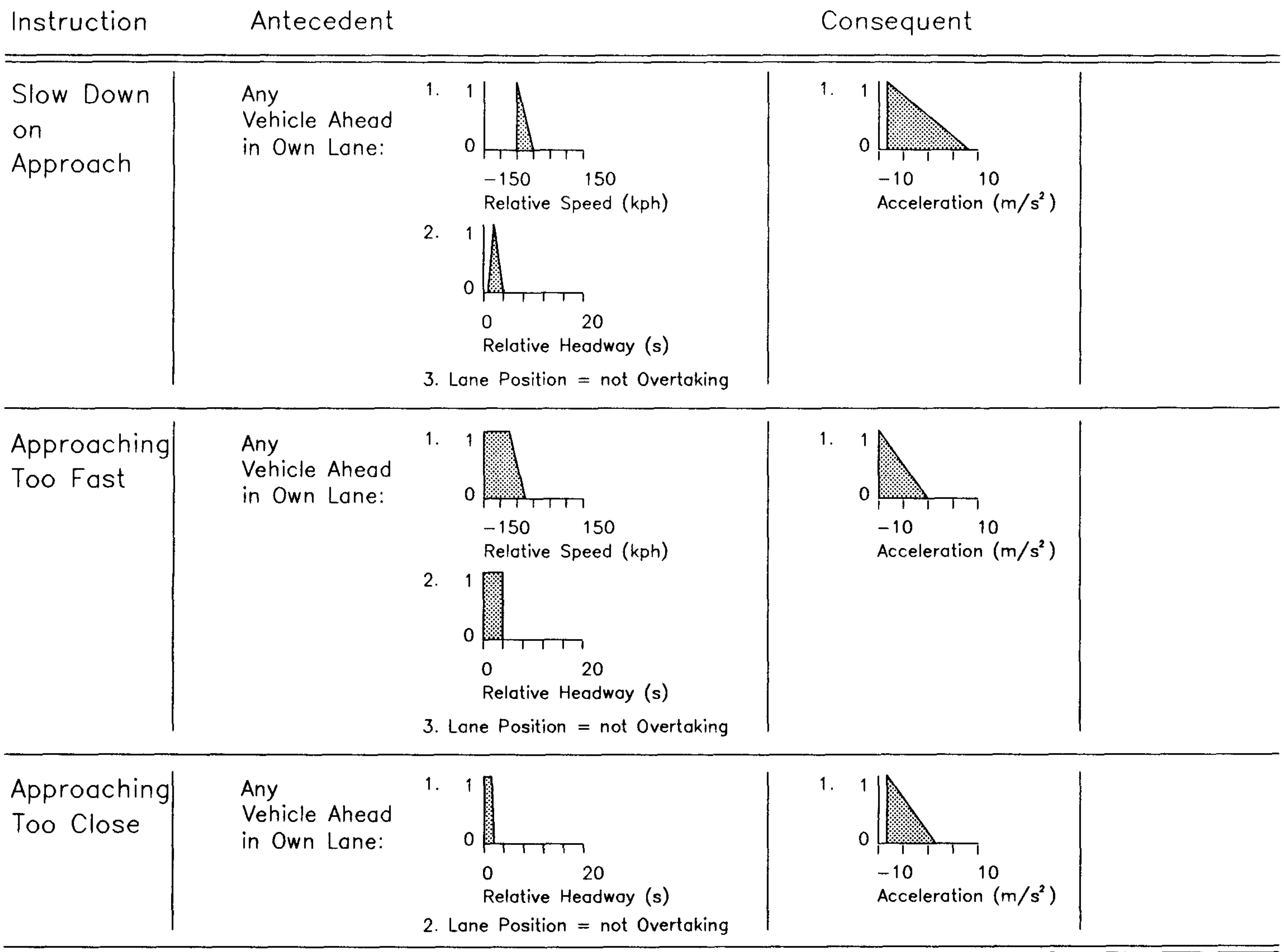


Figure 6.26 Vehicle Following Instructions

can maintain a more constant speed himself.

6.4.3 Merging and Lane Changing. The lane change manoeuvre is part of the merging, turning and passing processes. The path of a driver's vehicle in such a manoeuvre can be roughly described as a sine wave (Godthelp 1985) with its proportions being a function of the speed of travel. Figure 6.27 shows average time histories of lane changes at different speeds along with standard deviations of lateral position.

On a two lane highway, merging occurs at the terminal end of sections with an added passing or climbing lane and at on-ramps. In both cases, the process is that of a forced merge, that is one in which normally no queues are formed. Over the length of a merging section, an average driver in either the merging or through lane gradually adjusts his car following strategy so that he falls in behind the nearest vehicle ahead regardless of which lane it is in. In this way, the traffic in both lanes can 'zipper' together. In merging, drivers tend to temporarily accept smaller headways (about one second) than in single lane car following. Presumably their heightened attention to the driving task allows for this. (Lyons, Rainford, Kenworthy and Newman 1988). The majority of drivers merge within 30 metres of the end of the lane taper (less than 2 seconds travel time) (Homburger 1987).

Because lane changing behaviour does not vary a great deal across the driving population, a single instruction was used in all driver prototypes. This instruction, a variation of which is depicted in Figure 6.28, is used as part of several other conscious programs. The

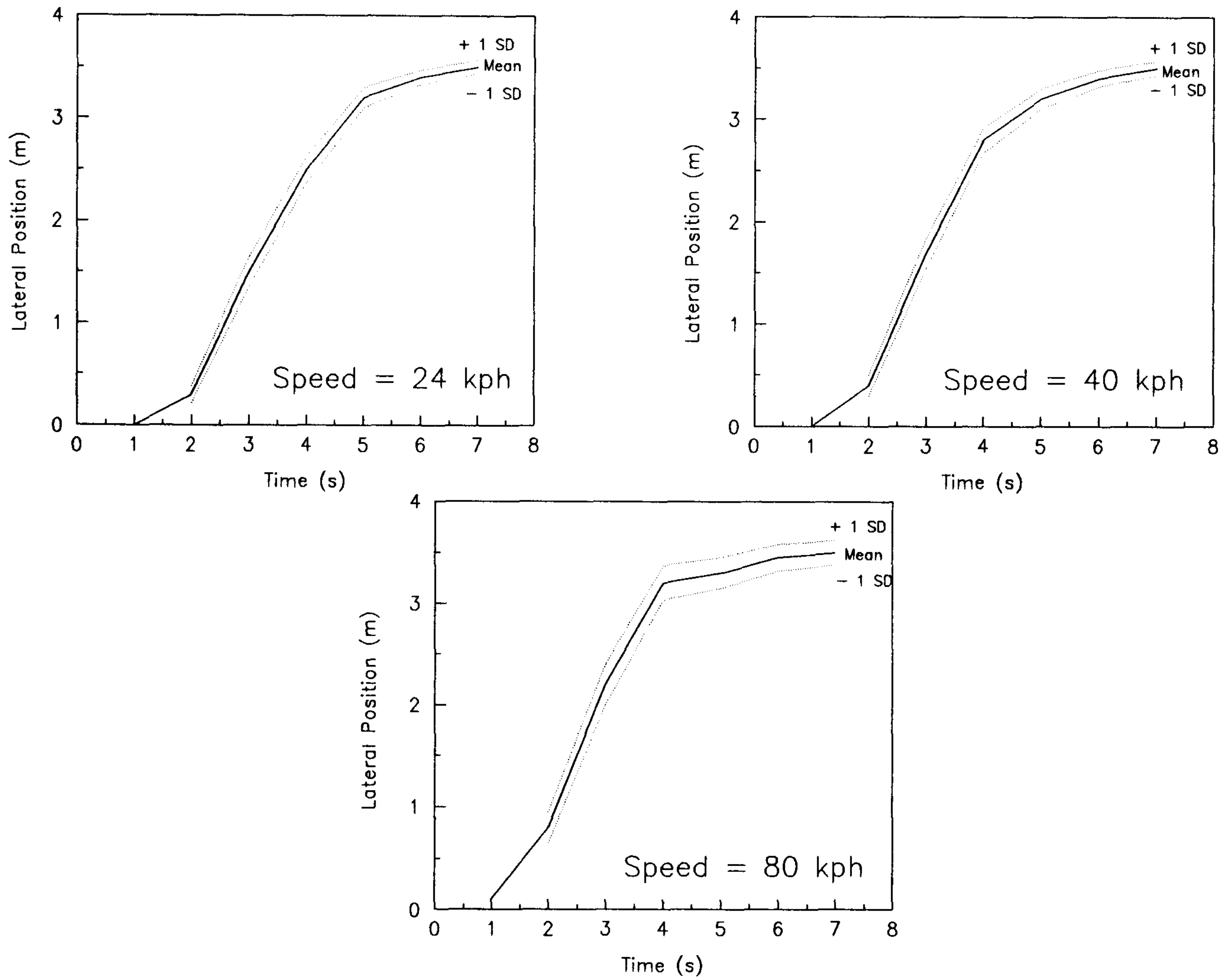


Figure 6 . 27 Lane Change Manoeuvres
(Source: Godthelp 1985)

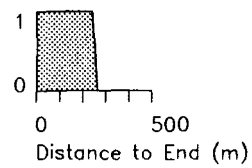
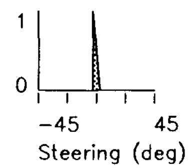
Instruction	Antecedent	Consequent	Flags
Other Instructions ...			
Change Lanes Left	Any Lane Ahead in Own Lane: 1.  2. Type = Merge Left	1. 	Variable
Other Instructions ...			

Figure 6.28 Example Lane Change Instruction

instruction in Figure 6.28 accomplishes a lane change manoeuvre to the left before the current lane tapers out.

6.4.4 Passing. On a two-lane highway, passing is undertaken by a driver wanting to travel at a higher speed than one immediately ahead of him. It involves travel in the opposing lane so before attempting to pass, a driver must ensure that there is a sufficient gap in the opposing traffic. Judgement of these gaps is performed on the basis of the distance to an opposing vehicle and the speed of the driver's own vehicle. The speed of the oncoming vehicle is not used because at typical oncoming car speeds and distances, the rate of change of the visual angle is below human detection threshold (Farber 1969). Drivers assume that oncoming vehicles are travelling at the same speed as they are (Wilson and Best 1982). Crawford (1963) posited that every driver has a threshold gap size, above which he considers passing safe. Over the driving population, there exists a threshold zone within which some drivers will attempt a pass and others will not. Gap acceptance thresholds are graphed in Figure 6.29. The thresholds at low speeds allowed for more than adequate time to complete an average performance passing operation. At high speeds, however, drivers were forced to hurry passes at their thresholds to complete them safely. This is perhaps because an average driver has had less experience passing at high speed than at low speed and therefore has poorer judgement regarding the former.

Crawford analyzed gap acceptances as a function of the passing driver's vehicle type and found that there was no statistically significant difference among small cars, large cars and

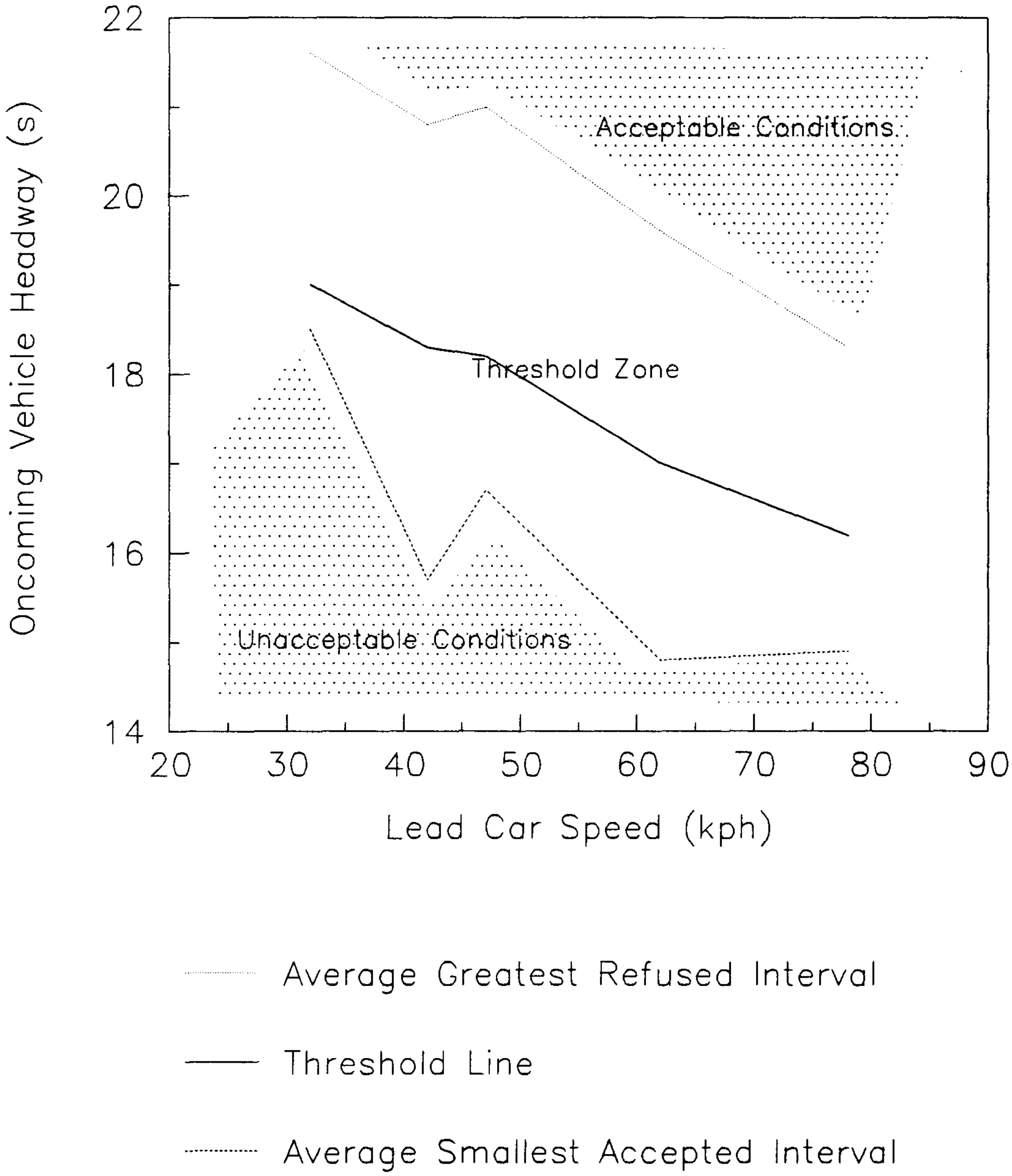


Figure 6.29 Accelerative Passing Gap Acceptance

(Adapted from: Crawford 1963)

light commercial vehicles, their average threshold being 11.5 seconds. For heavy commercial vehicles, the average threshold was 14 seconds. The type of oncoming vehicle affected thresholds only slightly such that thresholds increased for larger oncoming vehicles. The type of impeding vehicle affects thresholds such that they are larger for commercial vehicles than for passenger cars (Wilson and Best 1982) presumably because the former restrict sight more and are longer.

In cases where there is no visible opposing vehicle, drivers must judge whether there is sufficient sight distance to allow for aborting a pass should an opposing vehicle suddenly come into view. Generally, gap thresholds for sight distance limited passing are smaller than those for passing against an opposing vehicle (Wilson and Best 1982).

The time taken to make a decision whether or not to pass is important because while drivers decide, the vehicles continue to move. Response times vary with the offered gap size as shown in Figure 6.30. Drivers take longer to decide when offered gaps are close to their thresholds and also when offered gaps are large. In the former case, the larger risk of a 'close call' means drivers take longer to be sure they are making the right decision. Ironically, by taking longer to decide, they allow themselves smaller gaps and increase the risk still further. In the latter case, the large gaps allow a more leisurely, less rushed pass. In some situations, a gap can be seen before it is available and a decision can be made whether or not to use it beforehand. In these situations, only a last confirming glance need be made once the gap is available and the effect of decision time on performance is

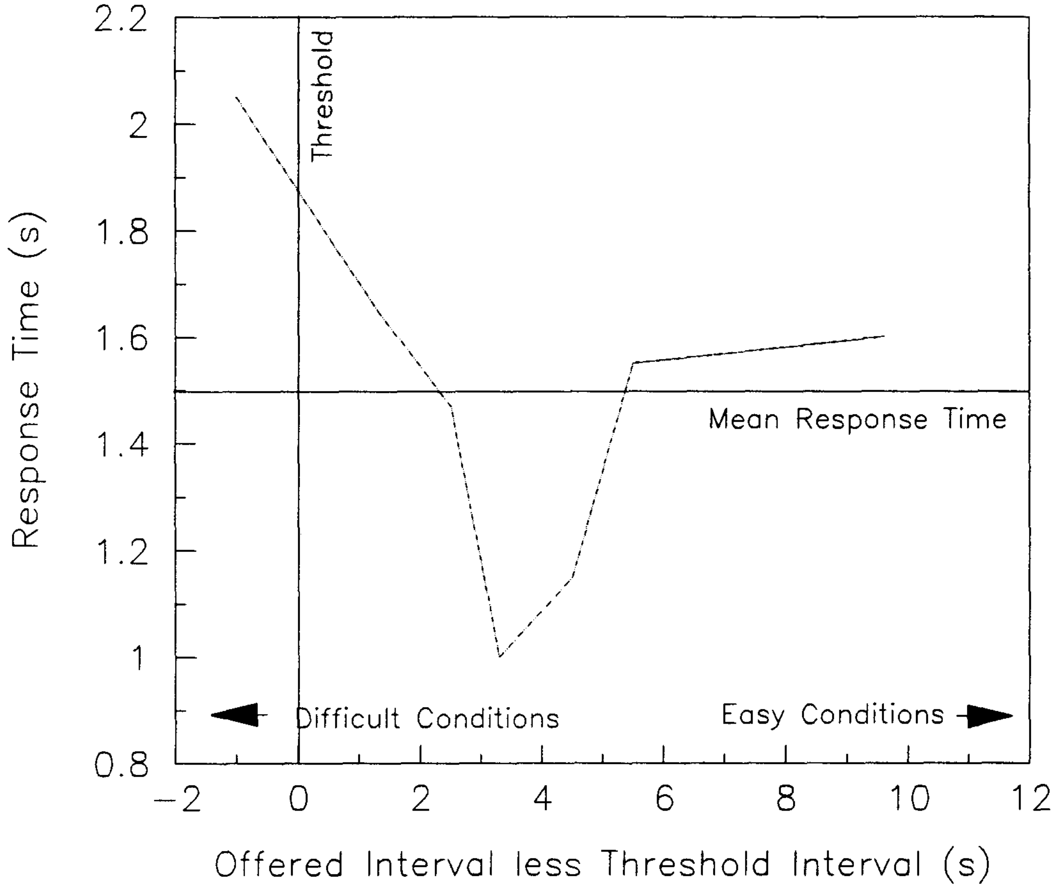


Figure 6.30 Passing Decision Time

(Source: Crawford 1963)

consequently diminished. Also, in these situations, a driver can accelerate close behind the impeding vehicle thus beginning his pass before the opportunity arises.

Gap acceptance thresholds are a function of the number of passing opportunities presented. In terrain with limited sight distances, thresholds tend to decrease as drivers realize that opportunities are smaller and less frequent (Matson and Forbes 1938).

One should also consider the situation in which a second driver piggybacks on the passing manoeuvre of another, following him into the oncoming lane and passing the same impeding vehicle or vehicles. The gap acceptance thresholds of piggyback drivers tend to be lower than those of drivers in ordinary passing situations. This may be because of a feeling of invulnerability due to the presence of the lead car (Wilson and Best 1982).

Besides gap acceptance, a driver must determine whether there is sufficient room to return to the right-hand lane after completing the pass. It may be that there is a sufficient gap in opposing traffic but that the headway between the impeding vehicle and the vehicle ahead of it is too small to allow a comfortable return after passing.

Once deciding to pass, a driver accelerates, if necessary, to a speed which is between 16 and 19 kph faster than that of the impeding vehicle and then maintains this speed through the duration of the manoeuvre. Average acceleration in this manoeuvre ranges from 0.4 to 0.7 m/s^2 (Prisk 1941). Passed vehicles do not appreciably change their speed during the manoeuvre.

Matson and Forbes classified passing into two types:

> In the first class, will be found the "flying"type of manoeuvre, wherein the overtaking vehicle proceeds at constant or nearly constant speed so as to complete the entire pass without slowing down. In the second class, will be found the type of manoeuvre which may be defined as the "accelerative" type, wherein the overtaking vehicle is following behind the overtaken vehicle and by acceleration increases its speed so as to complete the pass.

The paths of passing vehicles in different situations are shown in Figure 6.31. These curves are constant speed curves where there is no interference from opposing traffic. Curve C shows the path of following vehicles seeking an opportunity to pass and, in the meantime, travelling the same speed as the vehicle ahead. As the headway drops below 2 seconds, the following driver edges toward the centre line to check for a passing opportunity. Curve B shows a passing manoeuvre in which the passing vehicle travels between 6 and 14 kph faster than the impeding vehicle. The passing driver crosses the centre line at a headway of 1.5 seconds behind the impeding vehicle and returns to the right-hand lane at a headway of 1.2 seconds ahead of it. Curve A shows a pass in which the passing vehicle travels more than 16 kph faster than the impeding one. In this situation, the passing driver crosses into the opposing lane at a headway of 3 seconds behind the impeding vehicle and returns to the right-hand lane at about 2.5 seconds ahead of that vehicle.

In a flying pass, a driver may maintain a constant speed throughout and thus follow one of the curves in Figure 6.31. In the course of an accelerative pass, he may switch from curve C to curve B and then from curve B to curve A as he increases his speed differential versus

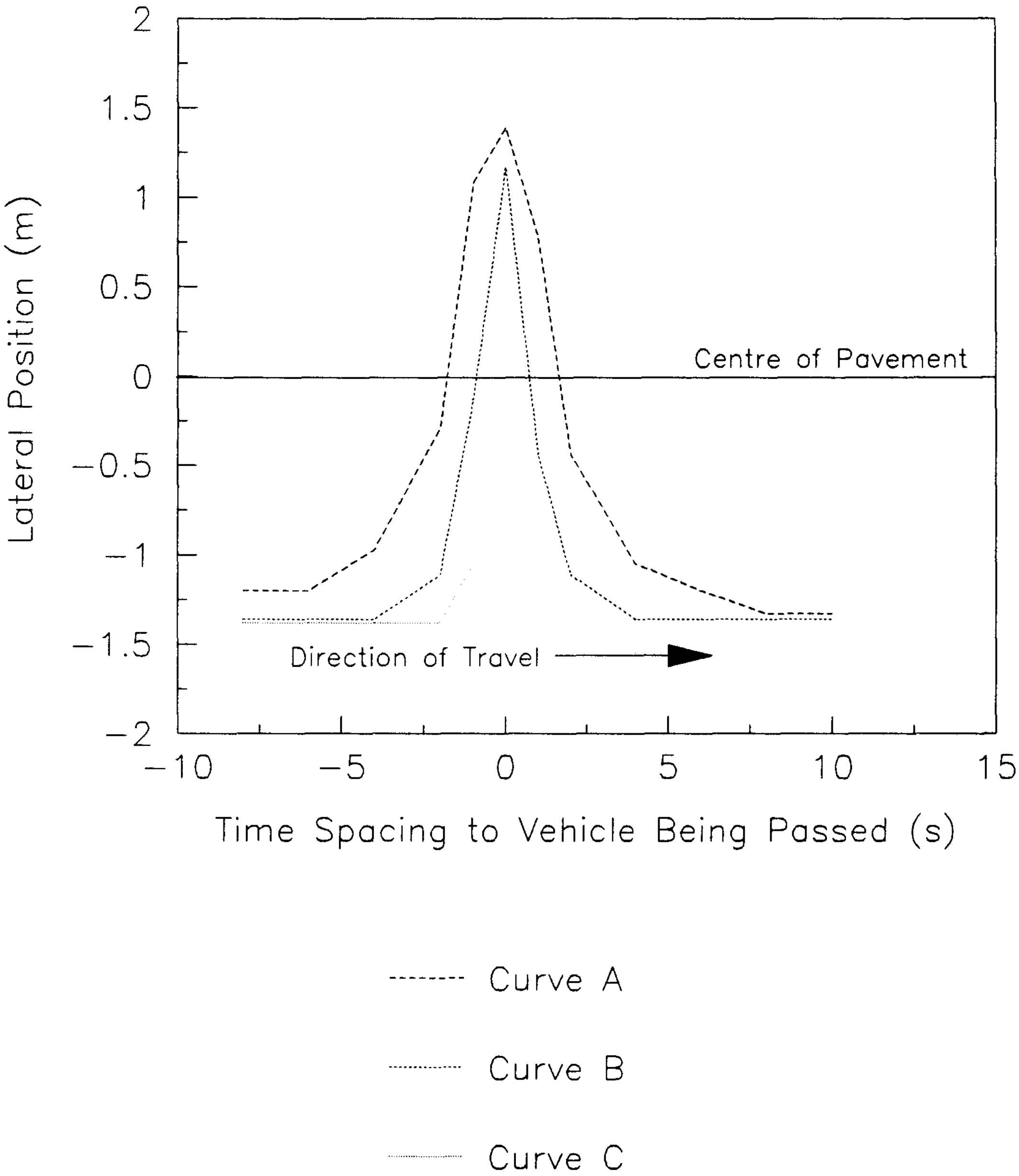


Figure 6.31 Path of Passing Vehicle

(Source: Taragin 1944)

the impeding vehicle.

In the opposing lane, passing drivers judge their lateral position by the centre line and assume a position such that their right wheels are about 0.5 metres from the centre line regardless of pavement width. This is when they are passing cars. When passing commercial vehicles, they allow slightly more distance from the centre line (0.6 metres). While being passed, drivers shift their lateral position only about 0.15 metres to the right. The lateral clearance between vehicles is less than that used when meeting opposing traffic likely because of the much lower speed differential between vehicles (Taragin 1944).

After going by the impeding vehicle in the opposing lane, drivers do not return to the right-hand lane until they have driven some distance past the other vehicle. This distance is relatively constant (between 14 and 18 metres) and not a function of speed. The exception is for passes near the threshold in which drivers are more rushed. In these situations, they allow only about half this clearance before returning to the right-hand lane.

As an example of the passing programs developed for the driver prototypes, Figures 6.32 and 6.33 show that for an accelerative pass of one vehicle. Trade offs in the objective function were made between constancy and risk-taking versus safety, the former two tending to give low gap acceptances and the latter tending to give high gap acceptances.

6.4.5 Passing and Climbing Lanes. Two lane highways in hilly or mountainous terrain often

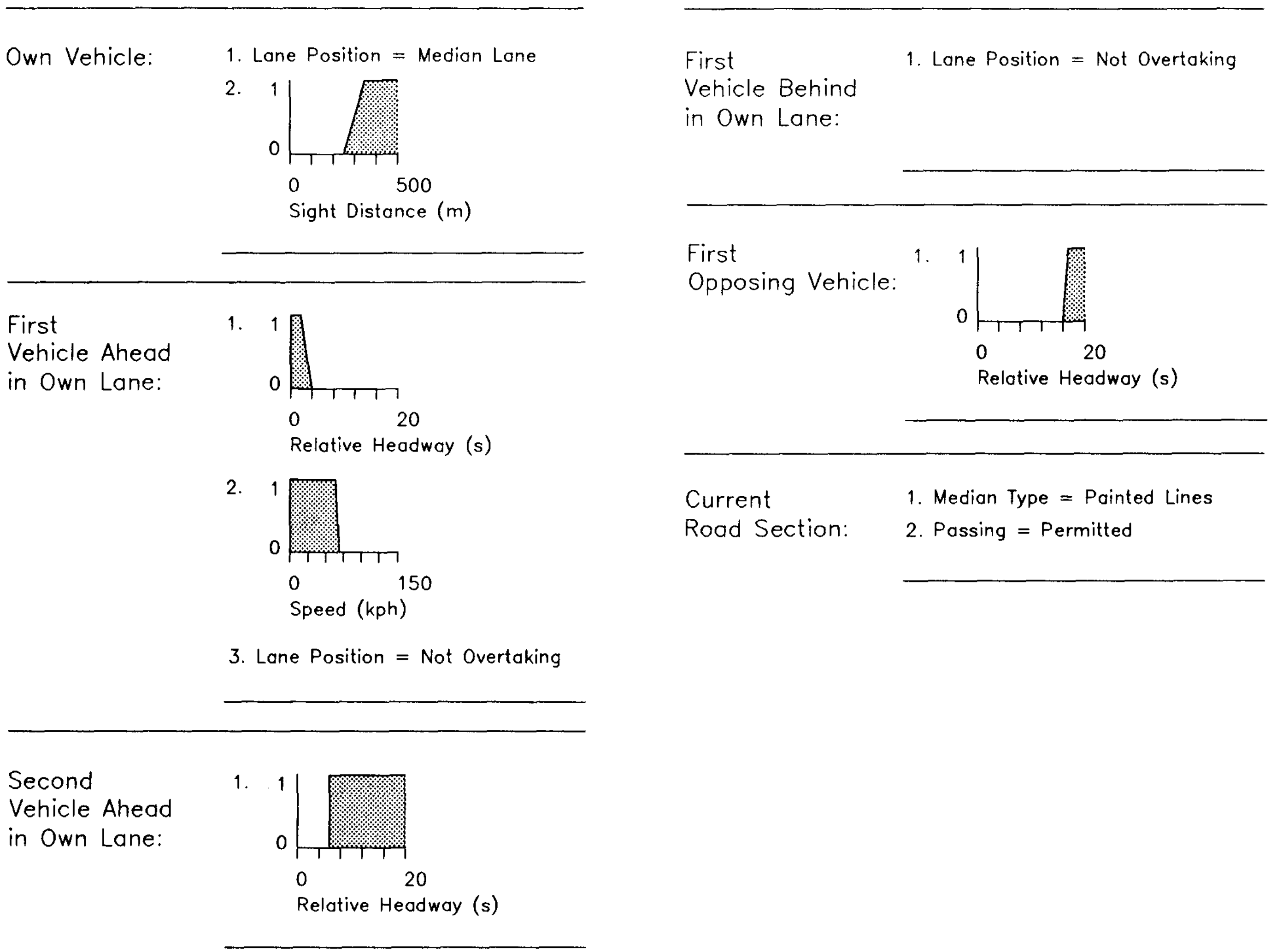


Figure 6.32 Starting Conditions for an Accelerative Pass

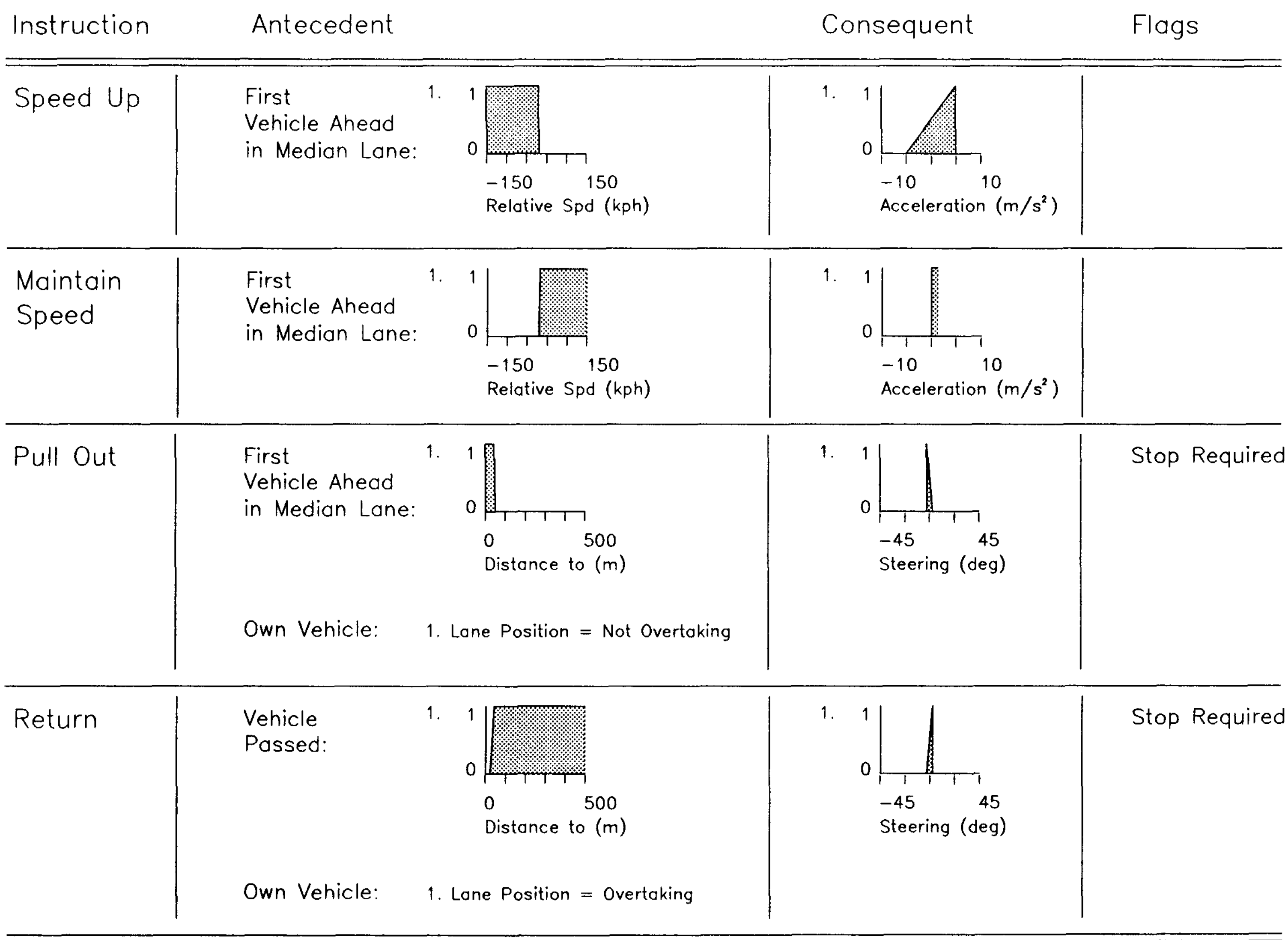


Figure 6.33 Accelerative Pass Program Instructions

do not provide very many passing opportunities because of restricted sight distance. Hence, the use of periodic passing lanes. Passing lanes are additional lanes added over a short distance to allow faster vehicles to get by slower ones. Climbing lanes are similar in form but serve a slightly different purpose. They are provided on long or steep uphill sections so that faster passenger cars can pass larger and slower commercial vehicles. Passing and climbing lanes are usually added as outer lanes and are designed so that slower vehicles use them, leaving the main travelling lane for faster vehicles.

In driving through a section with a passing lane, drivers face two tasks in addition to ordinary driving. First they must choose which of the two lanes to use and, second, they must merge back into a single lane at the end of the section. The latter merging operation was described in section 6.4.3. The former is discussed here.

The traffic stream entering a passing lane section usually consists of a series of platoons led by slower moving vehicles. Passing lanes would be most effective if the platoon leader always chose the right-hand lane of travel. This is not always the case, however, and the choice seems to be partially a function of the roadway curvature at the start of the section. On roads that curve to the right, less than half of platoon leaders initially choose the right lane. On roads curving to the left, four out of five leaders initially choose the right lane, and, on straight sections, about two thirds choose the right lane (Batz 1989). This suggests that drivers, on curved sections, seek the path of least curvature, perhaps because this is the one requiring the smallest speed reduction. Of the vehicles which initially choose the left lane, some change to the right lane part way through the passing section, presumably upon noticing that they are being followed. Table 6.4 shows lane choice as a function of curvature.

Table 6.4 Platoon Leader Lane Choice in Passing Lane Sections
(Source: Batz 1989)

Curvature	Right Lane	Left Lane	Left Initially, Then Right
Right	47%	46%	7%
Straight	66%	29%	5%
Left	81%	18%	0%

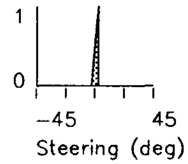
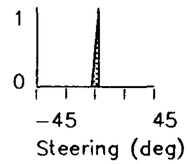
Because not all platoon leaders use the right lane, some followers decide to pass on the right. Table 6.5 shows the frequency with which this occurs when the platoon leader chooses to use the left lane. These values suggest that drivers are more confident in passing on the right on straight sections where they can accelerate to higher speeds and have a better view of the roadway ahead, specifically to note the remaining distance in the passing lane section. Also, passing on the right may be more prevalent on right curves than on left curves because one must travel a smaller distance to get by on the 'inside track'.

Table 6.5 Frequency of Passes on the Right When Platoon Leader Uses Left Lane (Source: Batz 1989)

Curvature	Passes of Platoon Leader	Other Passes Within the Platoon
Right	29%	29%
Straight	62%	67%
Left	17%	8%

The choice of lanes in a passing or climbing lane section was modeled using two programs, one conscious and the other unconscious. The unconscious program, shown in Figure 6.34, is used to determine the 'instinctive' or initial choice of lanes, based on the curvature of the road. Trade offs in its objective function were between constancy and safety, the latter of which in this case is arguably a surrogate for courtesy. The second, conscious program, shown in Figures 6.35 and 6.36, applies only to those who initially choose the left-hand lane. It models a driver's either giving way or not giving way to a car approaching from behind. The same trade-off was used here.

A third program, this one also conscious, was developed for passing on the right. The starting conditions are shown in Figure 6.37. Note the similarities and differences between them and those for accelerative passing (Figure 6.32). Program instructions are listed in Figure 6.38. In developing an objective function for this behaviour, a trade off was made between risk-taking and safety. The manoeuvre gives pleasurable risk because in many places it is illegal, and contravenes safety because the passed driver does not expect it and because of the imminent and necessary merge lying ahead.

Instruction	Antecedent		Consequent
Straight Road	Own Vehicle:	1. Lane Position = not Curb	1. 
	Current Road Section:	1. Horizontal Alignment = Straight	
Left Curve Road	Own Vehicle:	1. Lane Position = not Curb	1. 
	Current Road Section:	1. Horizontal Alignment = Left Curve	

Right Curve – no instruction for average driver prototype (ie. keep left)
Some driver prototypes with larger safety objectives and lower constancy have a right curve instruction similar to those above.

Figure 6.34 Lane Choice Instructions in Diverging Section

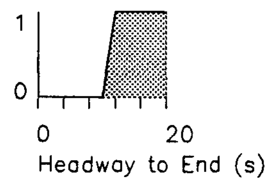
Own Vehicle:

1. Lane Position = not Curb

Any Lane Ahead and Right 1:

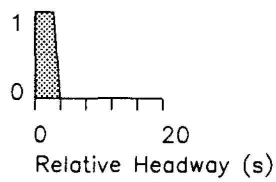
1. Type = Merge Left

2.



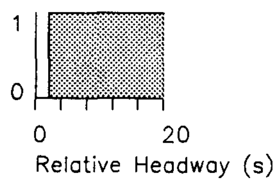
First Vehicle Behind in Own Lane:

1.



First Vehicle Behind One Lane to the Right:

1.



First Vehicle Ahead One Lane to the Right:

1.

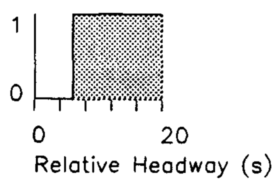


Figure 6.35 Starting Conditions for Giving Way to Right

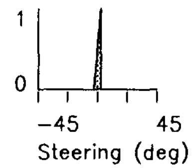
Instruction	Antecedent		Consequent	Flags
Steer Right	Own Vehicle:	1. Lane Position = Median	1. 	Stop Required

Figure 6.36 Program for Giving Way to Vehicles Behind

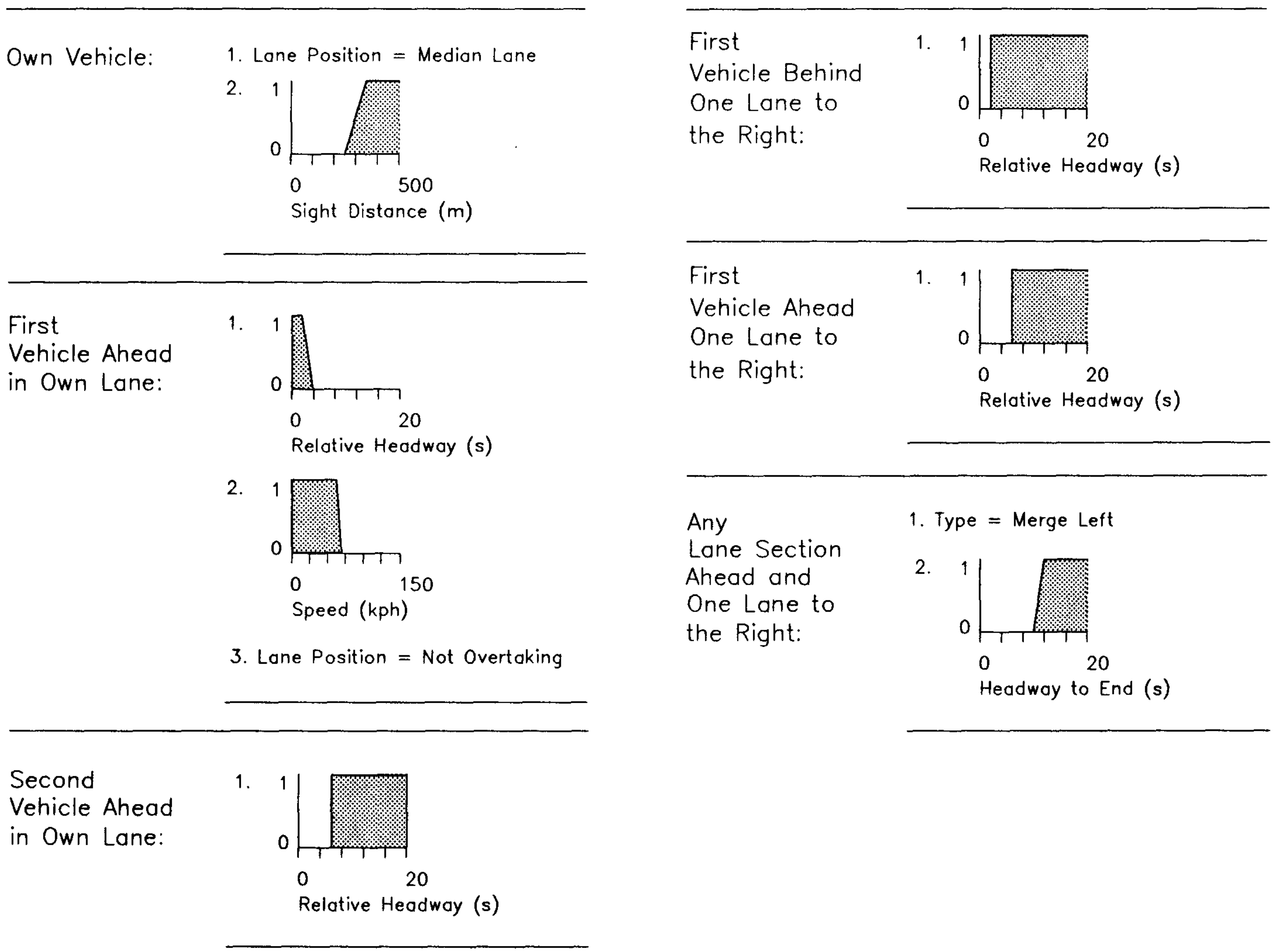
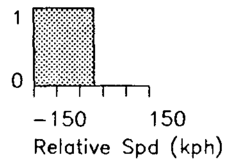
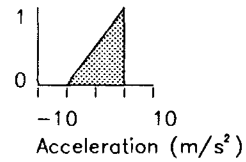
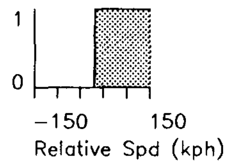
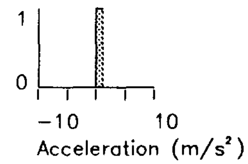
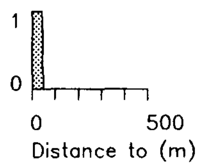
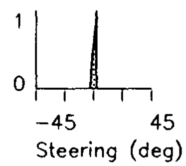


Figure 6.37 Starting Conditions for Passing on the Right

Instruction	Antecedent		Consequent	Flags
Speed Up	First Vehicle Ahead in Median Lane:	1. 	1. 	
Maintain Speed	First Vehicle Ahead in Median Lane:	1. 	1. 	
Pull Out	First Vehicle Ahead in Median Lane: Own Vehicle:	1.  1. Lane Position = Not Curb	1. 	Stop Required

Return – use merging program (Figure 6.28)

Figure 6.38 Passing on the Right Program Instructions

6.5 Emergencies

Most emergency situations in two-lane highway driving can be classified into two categories. The first occurs when one driver is following another and the one in front performs an erratic manoeuvre which requires either severe braking or steering or both on the part of the following driver. The second occurs in passing when an oncoming vehicle forces the premature completion or abortion of the manoeuvre by a passing driver. Other emergency situations such as skidding off the side of the road were considered too unique in their causes and consequently required variables or too infrequent to model effectively.

6.5.1 Car Following Emergencies. In the first type of emergency situation, that of car following, the following driver responds to certain clues which tell him that emergency manoeuvres are required. Among these are visual detection of swerving, deceleration or the application of brake lights of the vehicle ahead. Whether or not and how quickly a following driver picks up on these cues is dependent on the magnitude of the cue signal and, more importantly, whether or not he is applying a high level of attention to his driving. If the driver is not attentive, any cues will be unexpected and therefore elicit a slower response. For example, in situations where sudden braking is required, reaction times have been shown to be a function of expectation. The brake reaction times of drivers expecting to stop are as shown in Figure 6.39. Johansson and Rumar (1971) determined that in non-expectant situations, reaction times are some 35 percent higher.

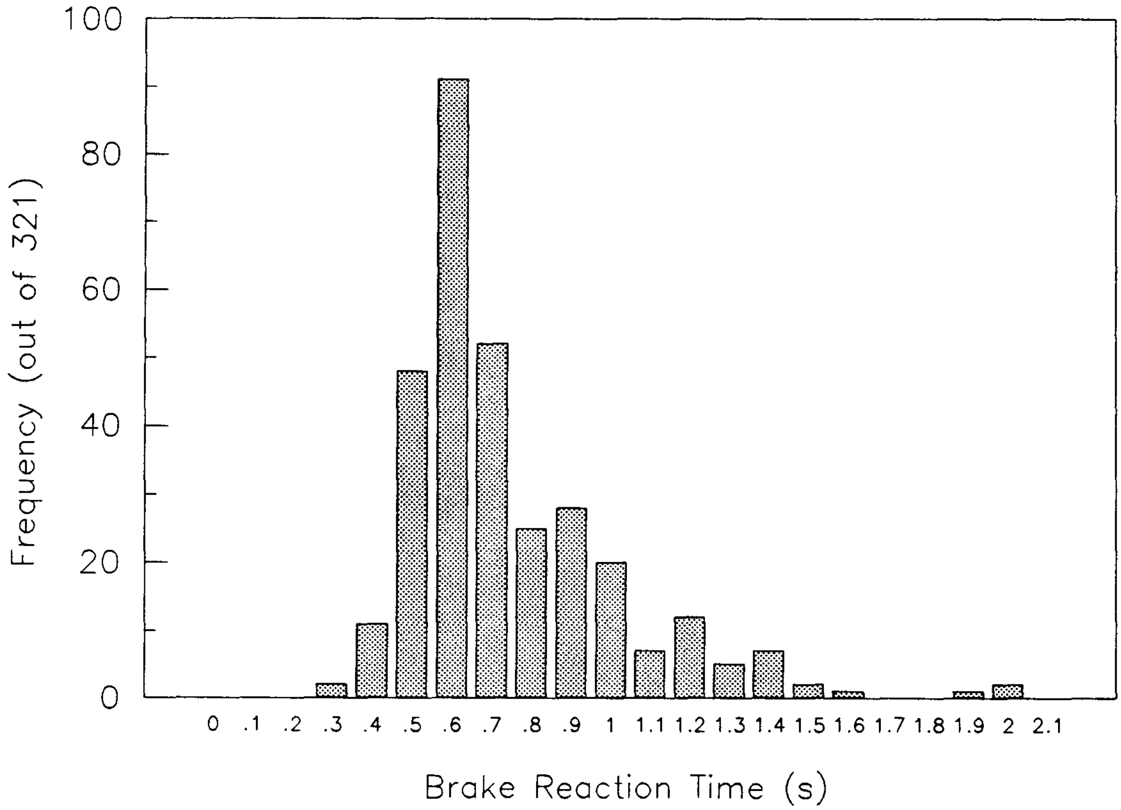


Figure 6.39 Brake Reaction Time of Expectant Drivers

(Source: Johansson and Rumar 1971)

The presence of more than one cue enhances the effectiveness of the others. In an experiment by Sivak, Post, Olson and Donohue (1981), the brake lights of a lead car were applied without decelerating to determine the response of the unsuspecting following drivers. Response rates within the first 3 seconds after application ranged from 31.4 percent for ordinary brake lights to 54.8% for ordinary lights supplemented with a single high mounted light. Among those drivers that did respond, the average response time was a high 1.38 seconds. Without the accompanying deceleration, the other following drivers may only have let up on the accelerator. In comparison, Johansson and Rumar (1971) obtained an average reaction time of 0.9 seconds in unanticipated braking situations in which the lead vehicle actually decelerated.

In addition to showing the effectiveness of multiple cues, this may also suggest that drivers, even when not anticipating an emergency, react in proportion to what they think a situation demands. This requires quick development of a more heightened attention to the driving task than merely recognizing a problem and instinctively giving an uncomplicated response, such as stamping down on the brake. Such responses point up the importance of having at least a certain amount of attention paid to the driving task prior to an emergency in order to avoid a collision.

Once drivers are alerted to an emergency situation in car following, their responses are sudden braking and steering actions. Braking responses can range as high as 8.4 to 8.9 m/s^2, limited of course by the deceleration capability of the vehicle on the particular road surface.

However, because rapid deceleration is a low probability event for most drivers, they lack experience with these situations and consequently may produce inappropriately low deceleration levels in emergency situations (Colbourn, Brown and Copeman 1978). This may be a consequence of the aforementioned desire among drivers to attempt a proportional reaction even in an unfamiliar emergency situation. The results of Koppa and Hayes (1976) in a test of driver abilities in emergency situations led to the same conclusion. In their study, drivers used up to 80 to 90 percent of vehicle deceleration limits, producing slightly larger values in expectant situations than in surprise ones.

A steering response in a car following emergency situation is often a secondary response initiated when the driver determines that braking alone will not prevent a collision with the vehicle ahead. In steering past the vehicle ahead, a lane change manoeuvre is made, either to the right perhaps onto the shoulder or off the road, or to the left, perhaps into the oncoming lane. Typical emergency lane change manoeuvres are shown in Figure 6.40.

Because of their rarity and high demands, avoidance programs are conscious ones. One for stopping and, if necessary steering to the right, is shown in Figures 6.41 and 6.42.

6.5.2 Passing Emergencies. Because drivers are not infallible in their passing decisions, the need to abort a pass may arise before its completion. This can involve one of two courses of action, either dropping back behind the impeding vehicle and returning to the right-hand lane or hurrying completion and returning to the right-hand lane ahead of the impeding

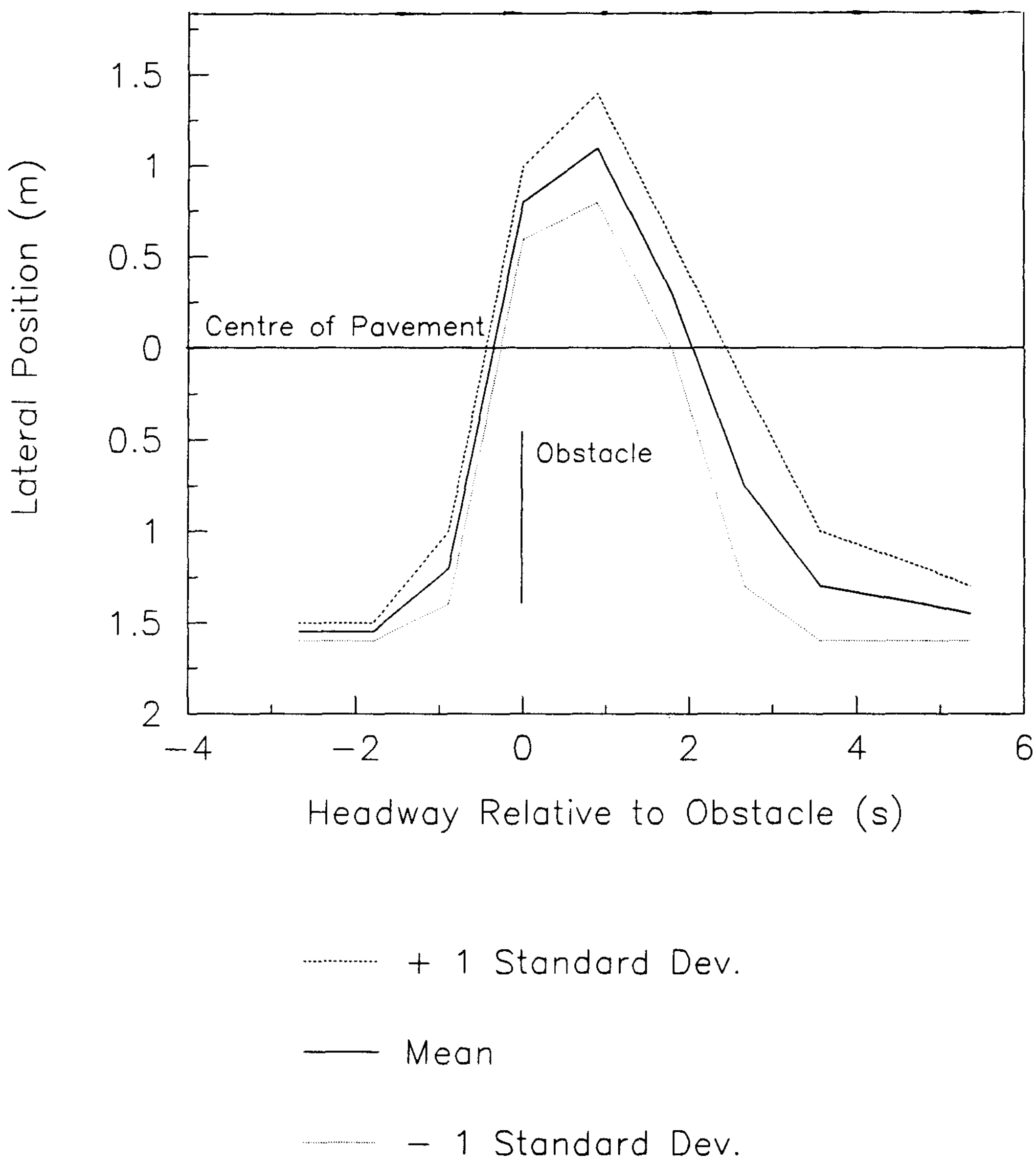


Figure 6.40 Obstacle Avoidance Manoeuvres

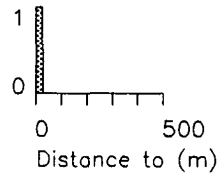
(Source: Reid, Solowka and Billing 1981)

Own Vehicle:

1. Lane Position = not Overtaking

First Vehicle Ahead in Own Lane:

1\.



2\. Lane Position = not Overtaking

Figure 6.41 Starting Conditions Rear-End Avoidance

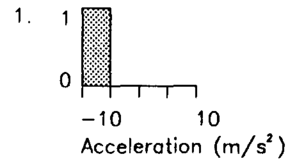
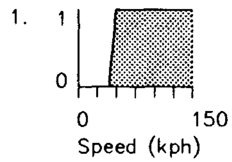
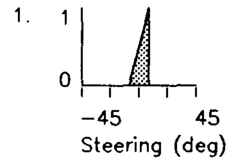
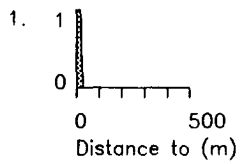
Instruction	Antecedent		Consequent	Flags
Slow Down				
Steer Right If Cannot Stop	Own Vehicle:			
	First Vehicle Ahead In Own Lane:			

Figure 6.42 Rear-End Accident Avoidance Program Instructions

vehicle. Glennon (1989) showed that there is a point in the passing manoeuvre, which he termed the critical position, beyond which it requires less time to hurry a pass completion and before which it is better to back off and return to a position behind the impeding vehicle. Glennon calculated the critical position rather than observing one drivers actually use. It is very likely that drivers assume a critical position different from that calculated by Glennon, however, perhaps erring on the side of backing off when they should complete the pass. This is because of the perceived added risk in accelerating toward an oncoming vehicle. It is postulated here that the cues used by a driver in deciding how to abort a pass are his position relative to the impeding vehicle and the reserve accelerative capacity of his vehicle. Further, it is suggested that the following rules govern the situation for an average driver:

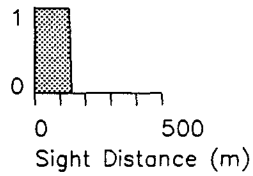
1. If the passing vehicle is ahead of the impeding vehicle such that the passing driver cannot see the latter in his peripheral vision, the passing driver accelerates further and returns to the right-hand lane at a small clearance ahead of the impeding vehicle.
2. If the passing vehicle is abreast of the impeding vehicle, the passing driver completes the pass as above provided he has sufficient reserve acceleration capacity.
3. If the passing vehicle is behind the impeding vehicle, the passing driver backs off and returns to the right-hand lane, allowing a small clearance to the impeding vehicle.
4. In situations where a passing driver believes that neither abort option will work, he returns to the right-hand lane immediately cutting off the impeding vehicle.

In the process of backing off, Cassel and Janoff (1968) suggest a deceleration rate ranging up to 6 m/s^2. The rules just described were developed into two conscious programs for the driver prototypes. They are shown in Figures 6.43 through 6.46.

Own Vehicle:

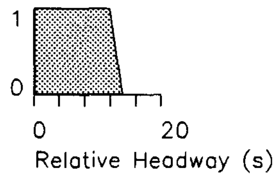
1. Lane Position = Overtaking

2.



First Opposing Vehicle:

1.



Vehicle Passed:

1.

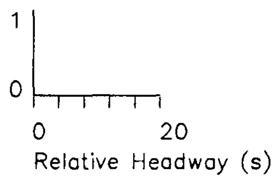


Figure 6.43 Starting Conditions for Aborting Pass

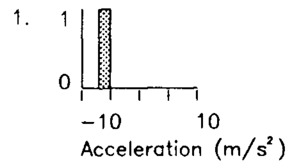
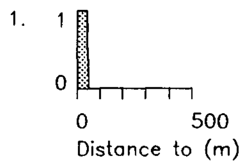
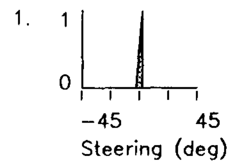
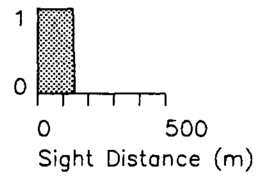
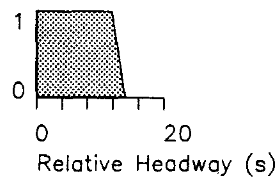
Instruction	Antecedent	Consequent	Flags
Slow Down			
Return to Right−Hand Lane	First Vehicle Ahead in Median Lane:  Own Vehicle: 1. Lane Position = Overtaking		Stop Required

Figure 6.44 Abort Pass Program Instructions

Own Vehicle:

1. Lane Position = Overtaking
2. 

First Opposing Vehicle:

1. 

Vehicle Passed:

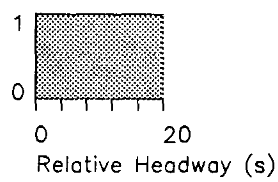
1. 

Figure 6.45 Starting Conditions for Hurrying A Pass

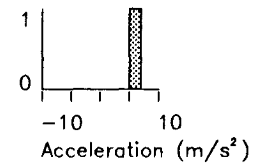
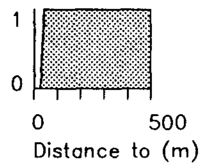
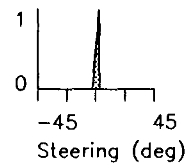
Instruction	Antecedent	Consequent	Flags
Speed Up		1. 	
Return to Right-Hand Lane	Vehicle Passed: 1.  Own Vehicle: 1. Lane Position = Overtaking	1. 	Stop Required

Figure 6.46 Hurry Pass Program Instructions

7 AGGREGATE CALIBRATION

Aggregate calibration, as opposed to component calibration, involved running a series of tests with the model, seeing how different input sets mapped into corresponding expected outputs. The input sets were for two lane rural highways and used the driver prototypes described in section 6.

The tests used three methods: (a) comparison to field data; (b) comparison to established standards; and (c) examination for reasonability. Table 7.1 shows the matrix of input-output mappings tested, and, for each, notes the test method. The Highway Capacity Manual

Table 7.1 Aggregate Calibration Tests

Input Variables / Output Measures:	Speed	Speed Variance	Percent Following	Passing
Traffic Volume	FD/HCM	R	FD/HCM	FD
Trucks	HCM	R	R	R
Directional Split	HCM	R	R	R
Grades	HCM	R	R	R

FD = Field Data
HCM = Highway Capacity Manual
R = Reasonability Examination

(1985) of the U.S. Transportation Research Board (hereafter known as the HCM) was chosen as the established standard for comparison. The field data came from Morrall and Werner (1982) and from Dommerholt and Botma (1988). The following sections discuss each of the input-output tests, grouping them by method.

7.1 Highway Capacity Manual Comparisons

The HCM provides procedures for determining levels of service under different road and traffic conditions, and, for two lane roads, uses three parameters to do this: (a) average travel speed; (b) percent time delay; and (c) capacity utilization. Percent time delay is defined as follows:

> "Percent time delay" is the average percent of the total travel time that all motorists are delayed in platoons while travelling a given section of highway. Motorists are defined to be delayed when travelling behind a platoon leader at speeds less than their desired speed and at headways less than 5 sec. For field measurement purposes, percent time delay in a section is approximately the same as the percentage of all vehicles travelling in platoons at headways less than 5 sec. (HCM 1985)

Separate calculations are provided for determining level of service on flat sections and sections with grades. A number of model runs under varying conditions were made and the outputs compared to these HCM calculations.

7.1.1 Level Sections. According to the HCM level of service is determined using relationships for a highway section under 'ideal' conditions coupled with a set of adjustment factors for deviations from these conditions. Ideal conditions are defined as follows:

1. Design speed greater than or equal to 60 mph.
2. Lane widths greater than or equal to 12 ft.
3. Clear shoulders wider than or equal to 6 ft.
4. No "no passing zones" on the highway.
5. All passenger cars in the traffic stream.
6. A 50/50 directional split in the traffic stream.

7. No impediments to through traffic due to traffic control or turning vehicles.
8. Level terrain. (HCM 1985)

The HCM defines six levels of service (A to F) using ranges of average travel speed, percent time delay and capacity utilization (volume to capacity ratio). It further defines volume to capacity ratio as a function of average travel speed, percent no passing zones and terrain type. It is this latter relationship which models traffic flow and it is to this that model outputs were compared. In the HCM, service flow rate on level sections is calculated using this equation:

$$SF_i = 2800(v/c)_i f_d f_w f_{HV} \tag{7.1}$$

where
SF_i = service flow rate (both directions) for prevailing conditions and level of service i (vph)
v/c_i = ratio of flow rate to ideal capacity for level of service i
f_d = adjustment factor for directional split
f_w = adjustment factor for narrow lanes or shoulders
f_{HV} = adjustment factor for heavy vehicles in the stream

Under ideal conditions, each of the factors f_d, f_w and f_{HV} are equal to one. A number of model runs were made first under ideal conditions and then varying one of these factors at a time. The road section used for comparison was a straight four kilometre section in which the middle two kilometres were used to obtain outputs. The middle section conformed to the ideal conditions listed previously. The two end sections, where generated traffic stabilized itself from starting conditions, were no passing sections.

The first set of runs was made for ideal conditions, varying only traffic volume. Plots of space mean speed versus service volume for the model and the HCM are shown in Figure 7.1. Note that space mean speed decreased with increasing volume for both, though the model speeds were consistently higher, perhaps for the following reason. The driver prototypes used in the model were structured using the concept of desired speed, which is defined as the speed a driver would adopt under the ideal conditions described previously and in the absence of other traffic. In Figure 7.1, mean desired speed is given by the y-axis intercepts of the graphs. For the model, it was 4 kph or 4.2 percent higher than for the HCM, accounting for the former's higher speeds at all traffic volumes. The desired speed distribution used in the model was from Australia (See Table 6.3) where speeds on two lane highways are generally higher than in the United States (the source of the HCM data). Other authors also consider the HCM to be conservative (eg. Morrall and Werner 1982).

The second series of model runs was made holding all the factors in equation 7.1 constant except the directional split factor, and the results are shown in Figure 7.2. The graph shows the same desired speed effects as shown in Figure 7.1 for ideal conditions. Otherwise, the model results were sufficiently parallel to those of the HCM to conclude that directional split effects were modeled correctly. These effects were partially caused by changes in both the demand and supply of passing opportunities with changes in directional split and therefore the results also give evidence that passing was modelled correctly.

The driver prototypes used in the model runs were not structured to account for variations in lane width and shoulder width (See section 6.2.1) because there exists little evidence

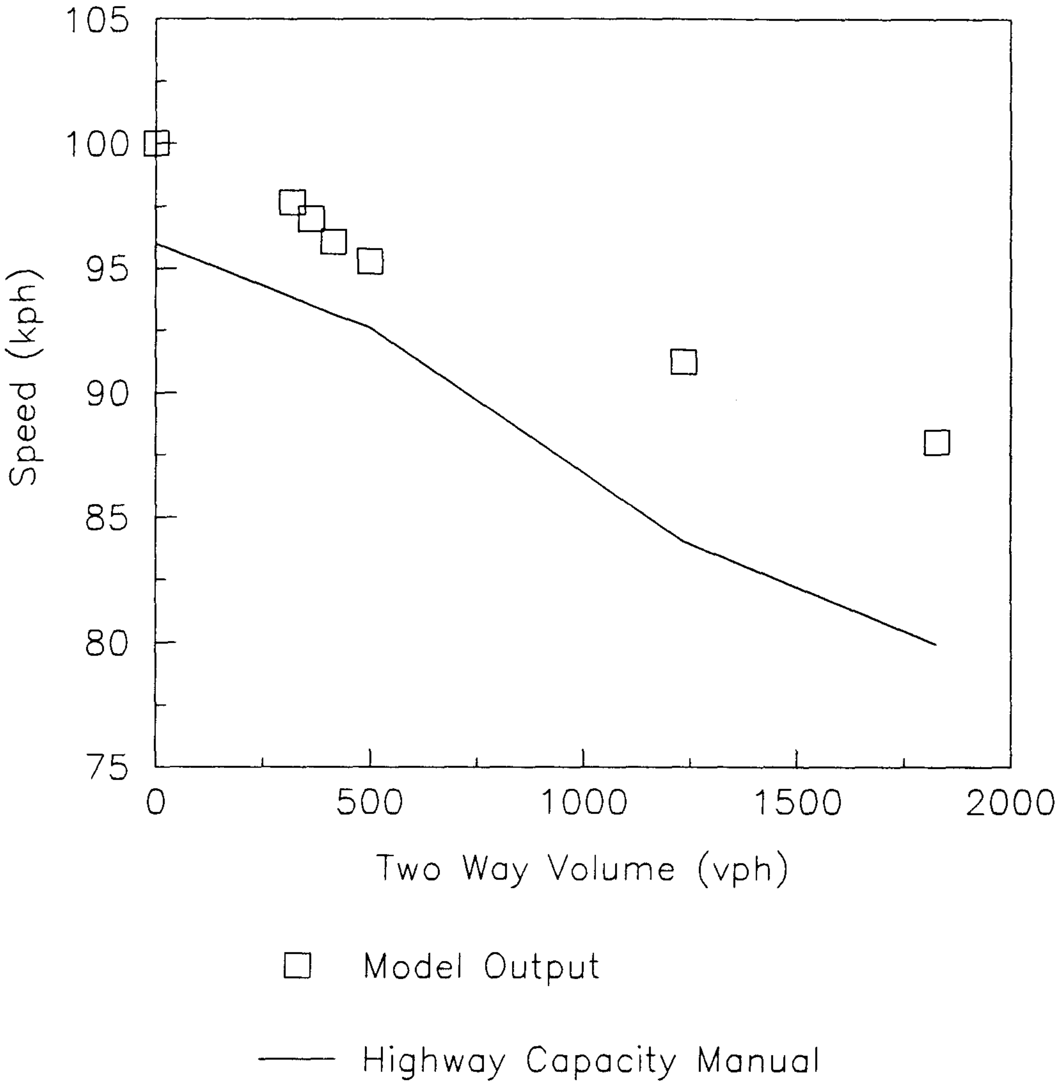


Figure 7 . 1 Space Mean Speed under Ideal Conditions

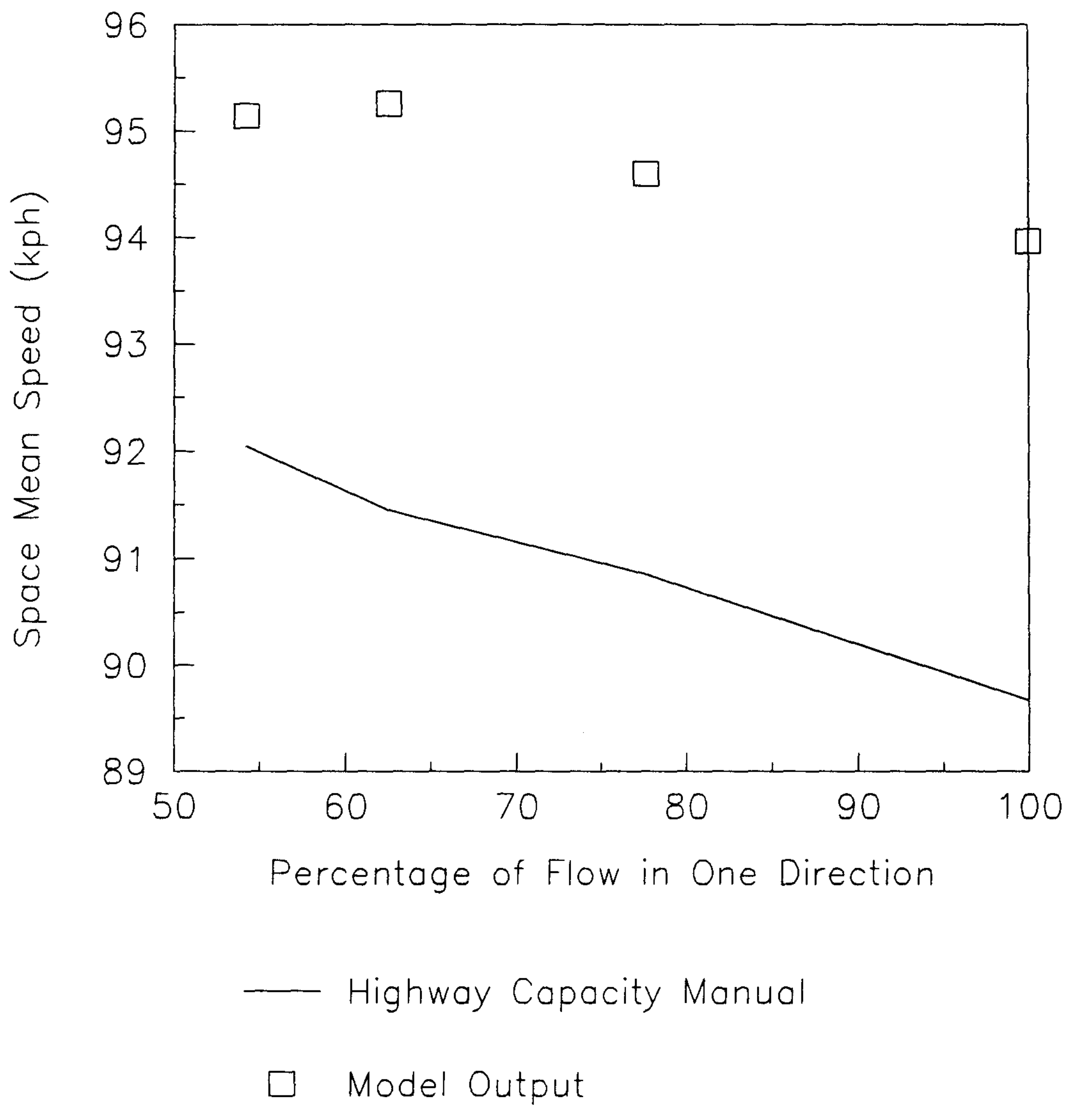


Figure 7.2 Directional Split Effects on Space Mean Speed

(apart from the continued use of the factor in the HCM) to support doing this. In fact, there is more evidence to the contrary (McLean 1989). Consequently, no comparisons were made varying shoulder or lane widths.

The third factor in equation 7.1 is that of heavy vehicles. The third series of model runs was made holding all but this factor constant with the results shown in Figure 7.3. This figure shows that over the range of 0 to 40 percent trucks, the model results did not differ from those of the HCM by more than 4.5 percent, suggesting that they were at least reasonable. The figure also shows that in the model, space mean speeds dropped at a higher rate than in the HCM for increases in truck traffic. This is perhaps explained by the desired speed distributions of the truck drivers used in the two. In the model, truck driver desired speeds were 80 percent of those of passenger car drivers, as noted in section 6.2.1. Perhaps, higher truck driver desired speeds underlie the HCM results giving the more shallow sloping curve. Another possible explanation is that there is a higher frequency of cars passing trucks underlying the HCM calculations. Without such passing and desired speed data, the exact cause of the effect cannot be determined. At low truck percentages, the passenger car desired speed effect described before was also evident.

7.1.2 Grades. The same approach to level of service for level sections is taken in the HCM for sections on grades except that the service flow calculation and those for deviation from ideal factors differed somewhat. Service flow is calculated as follows:

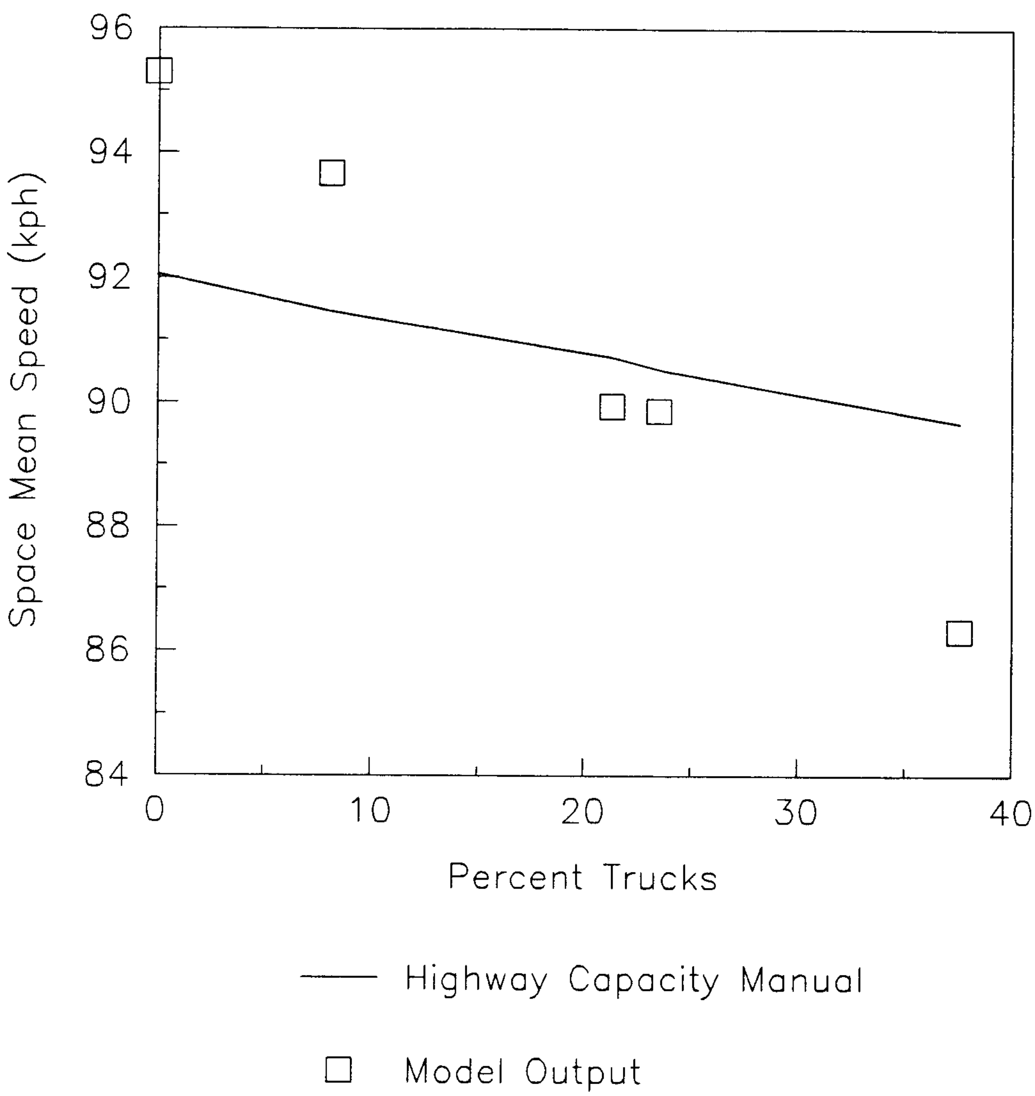


Figure 7.3 Truck Effects on Space Mean Speed

$$SF_i = 2800 (v/c)_i \; {}_d f_w f_g f_{HV} \tag{7.2}$$

where SF_i = service flow rate (two directional) for level of service i, or speed i, and prevailing conditions (vph)
v/c_i = ratio of service flow rate to ideal capacity for speed i
f_d = adjustment factor for directional split
f_w = adjustment factor for narrow lanes or shoulders
f_g = adjustment factor for the effects of grades on passenger cars
f_{HV} = adjustment factor for the presence of heavy vehicles in the upgrade stream.

Model runs were made on a four kilometre section similar to that described in the last section except that the grade of the middle two kilometre section was varied. The end sections were kept level. Several model runs were made varying only this grade and using a constant traffic stream consisting of 20 percent trucks and a total volume of about 500 vph.

Figure 7.4 shows grade effects on space mean speed for both the model and the HCM. It should be noted that the speeds did not differ in the range of 0 to 7 degrees slope by more than 5 kph or 6.6 percent. The graph shows that on level sections, the HCM and the model predicted the same space mean speed. This is because, at 20 percent truck traffic, the effect of the higher passenger car driver desired speeds in the model balanced that of the HCM's apparently higher truck driver desired speeds (See Figure 7.3). On grades, model speeds were lower than those of the HCM. This may be because of a difference in the truck population used in the model and that in the HCM. On grades, the acceleration capabilities of trucks often govern speeds rather than the desires of truck drivers. In the model, the truck traffic consisted of a fleet of heavily loaded Mercedes Benz 1113s (See Table 3.1 and

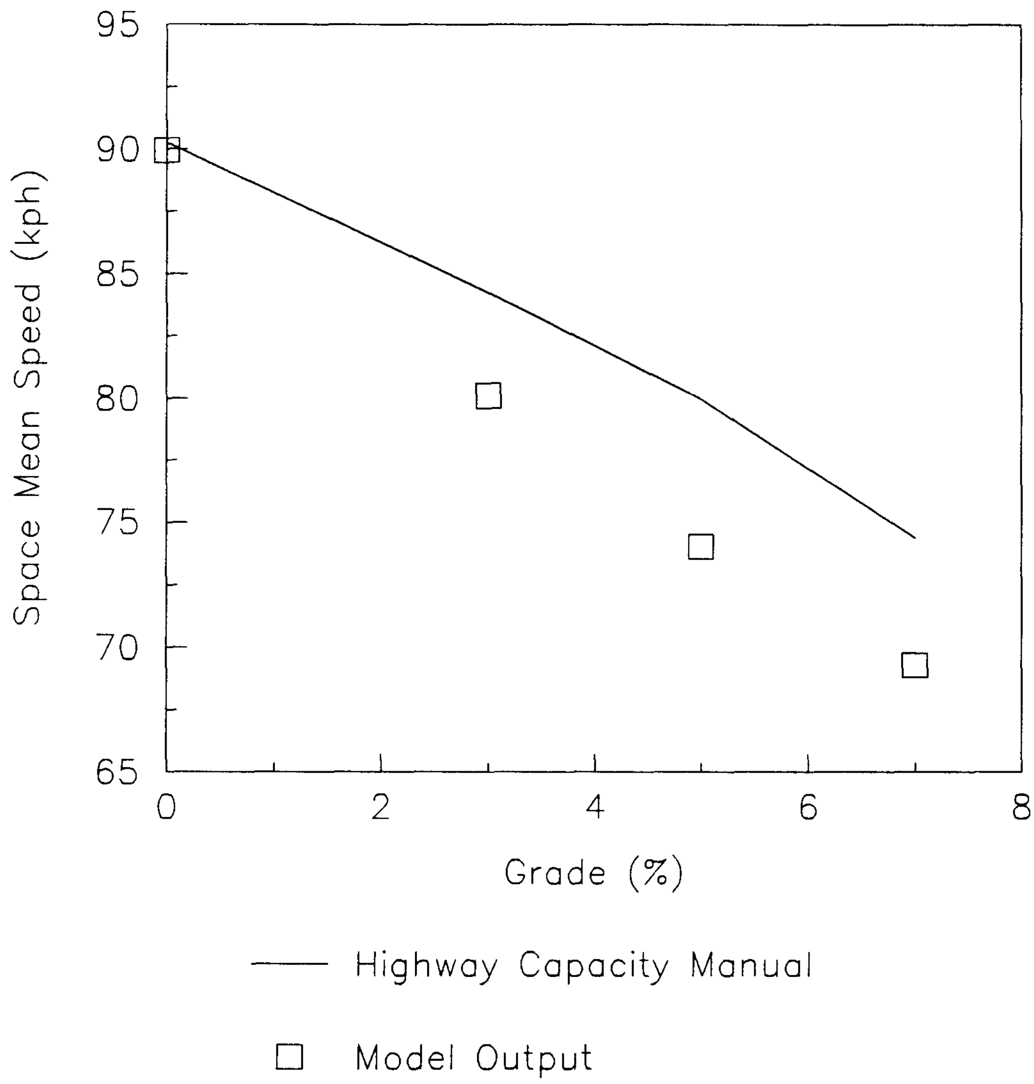


Figure 7.4 Grade Effects on Space Mean Speed

Figure 3.2). The acceleration capabilities of these trucks may have been lower than those of the trucks in the HCM traffic stream, consequently dropping the model speeds below those of the HCM.

7.1.3 Time Delay. Though the HCM defines level of service as a function partially of time delay, it provides no means to estimate it (accounting for deviations from ideal conditions). Instead, it contains a single graph of time delay versus two-way volume for ideal conditions. A series of model runs was made under these ideal conditions obtaining percent time delays for various volumes, and the results are shown in Figure 7.5. Note that the HCM defines a delayed vehicle as one travelling with a headway of 5 seconds behind another and this same criterion was used in the model runs.

The graph shows very good agreement between the model output and the HCM. Because the HCM describes this parameter as the primary measure of level of service, it can be concluded that the model can be used to give a good indication of level of service, at least under ideal conditions.

7.2 Field Data Comparisons

Field data was obtained for comparison with model outputs on the bases of space mean speed versus volume, percent following versus volume and passing versus volume. The first

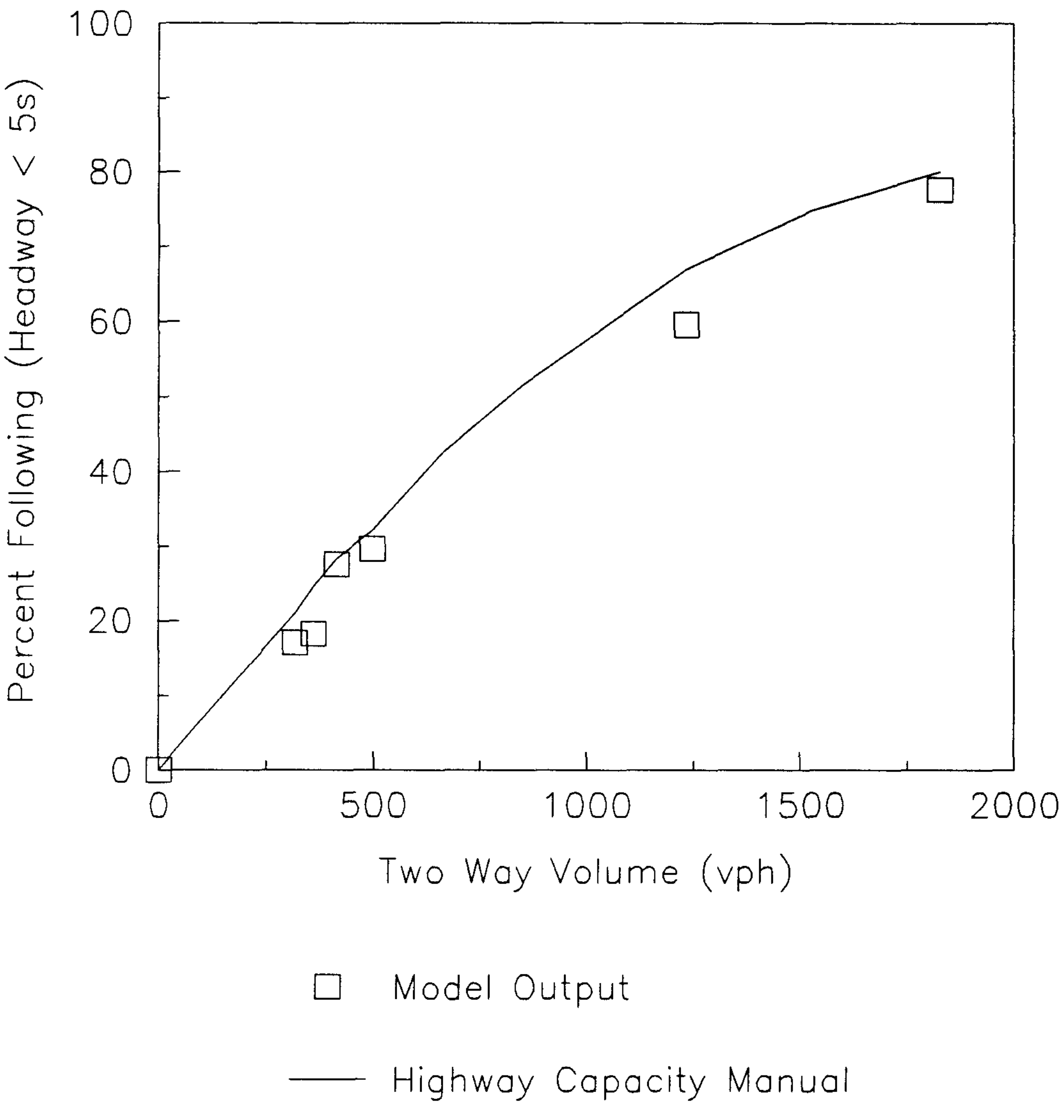


Figure 7.5 Percent Following under Ideal Conditions

two of these are keys in determining level of service on two lane rural highways.

7.2.1 Space Mean Speed. Speed and headway data were collected by Morrall and Werner (1982) on two sections of the Trans Canada Highway about 130 kilometres west of Calgary, Alberta during the spring and summer of 1980. Each of the road sections closely conformed with the ideal conditions of the Highway Capacity Manual, although the traffic composition did not. Typically, it consisted of 70 percent passenger cars, 5 percent trucks and 25 percent recreational vehicles.

Figures 7.6 and 7.7 show comparisons of space mean speed data from Morrall and Werner's two sites (called Anthracite and Sunshine) and model runs on the four kilometre ideal section described in section 7.1. Both graphs show that model speeds were consistently at the upper end of the range of speeds found in the field. This may have been because of the effects of the non-ideal traffic in the field, particularly the high proportion of recreational vehicles, compared with the 100 percent passenger car traffic in the model. Secondly, the sites on the Trans Canada Highway had posted speed limits of 100 kph. This may also have slowed faster drivers as noted by Morrall and Werner. In fact, these authors showed that the effect of a speed limit posting is to reduce the speed of a majority of drivers and reduce the variance at the same time (Figure 7.8).

7.2.2 Percent Following. Morrall and Werner obtained headway data in addition to speed data at Anthracite and Sunshine. From that they calculated percent following values for

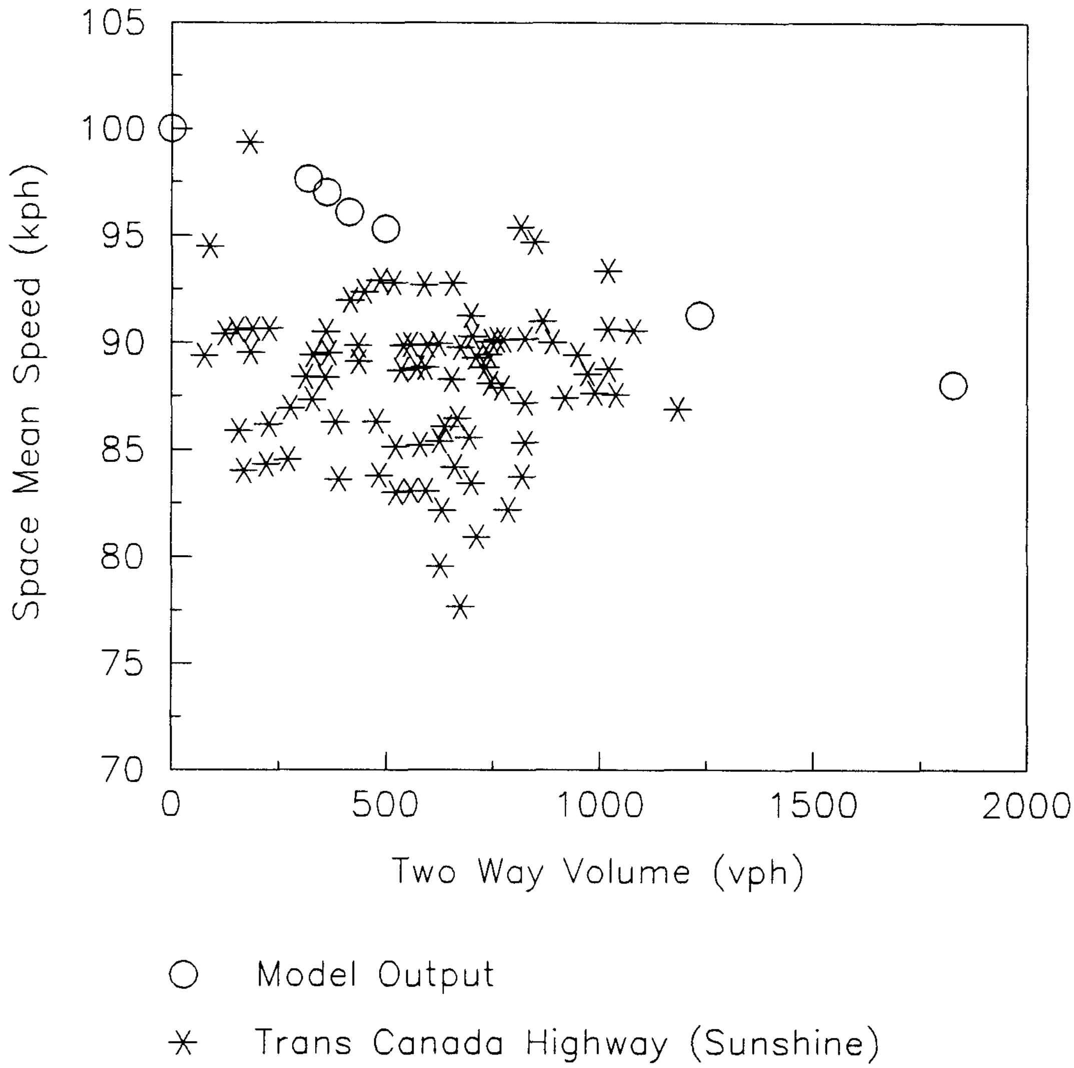


Figure 7.6 Space Mean Speed Comparison to Trans Canada Highway (Sunshine)

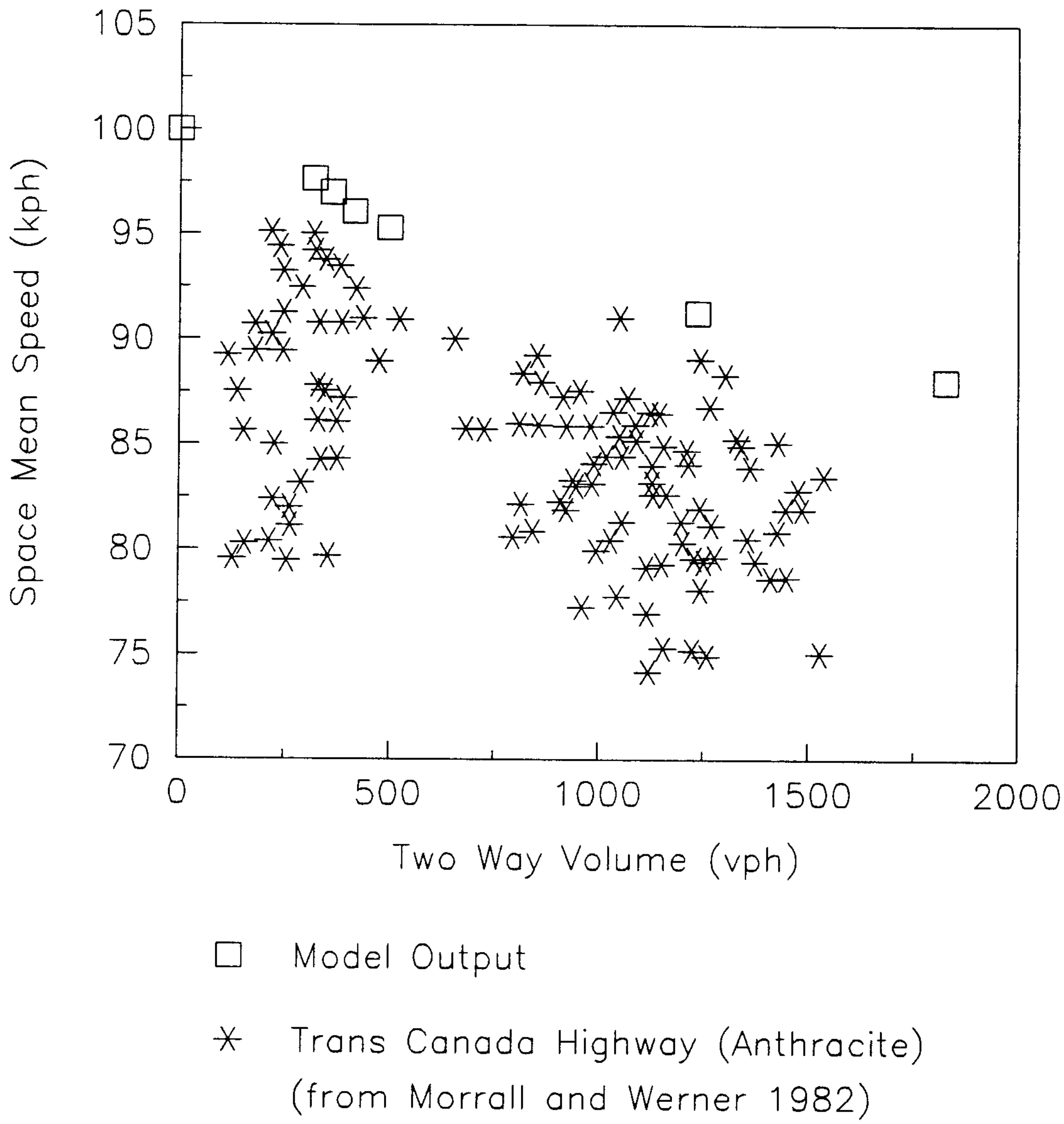


Figure 7.7 Space Mean Speed Comparison to Trans Canada Highway (Anthracite)

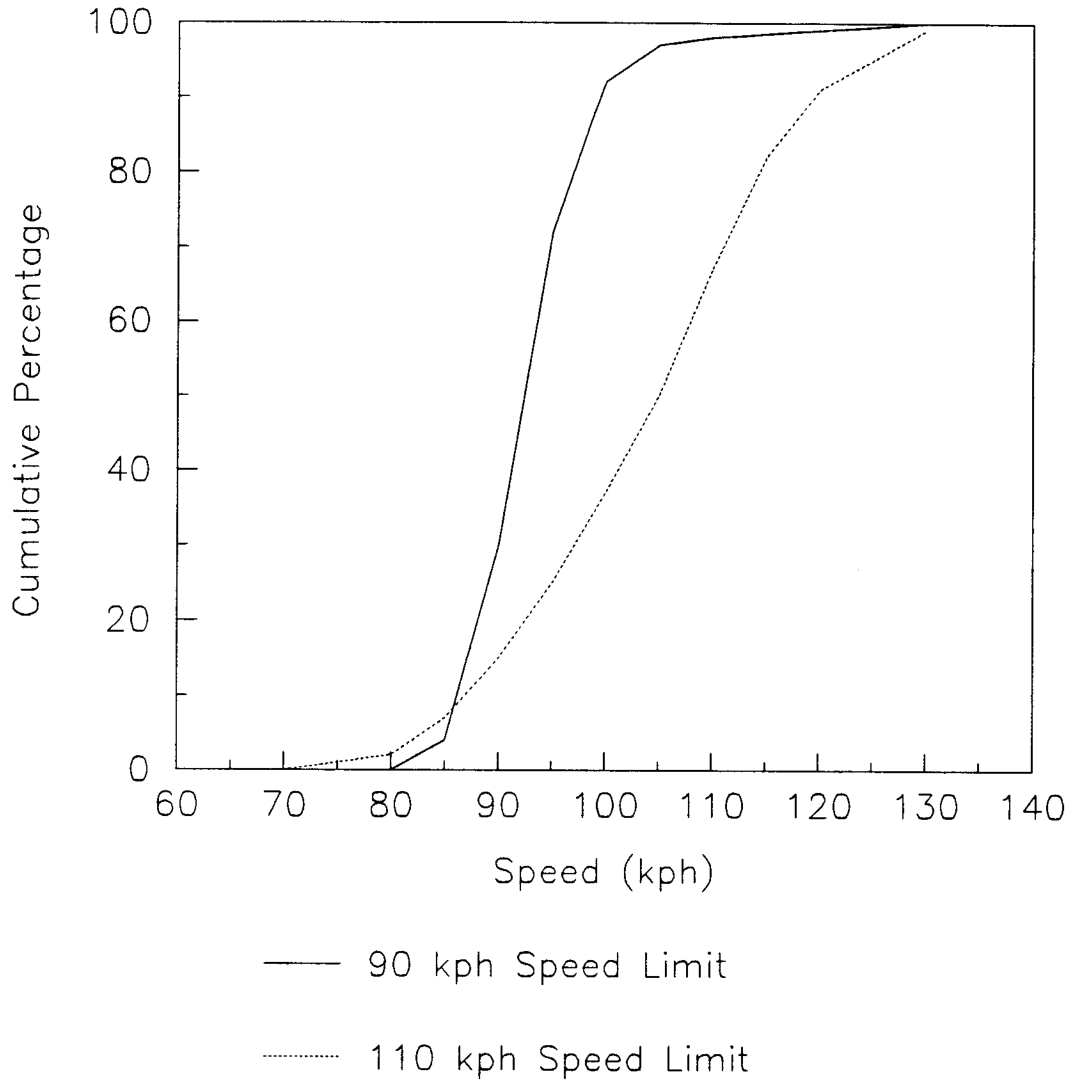


Figure 7.8 Impact of Speed Limits on Speed Distribution Curve

(Source: Morrall and Werner 1982)

various traffic volumes. These are shown in Figure 7.9, comparing them to model results. Note that the definitions of percent following used by Morrall and Werner and in the model differed, the former using a maximum following headway of 6 seconds. For this reason, it was expected that curves through the field data and through the model estimates would be of similar shape with the former lying just above the latter. This was in fact the case.

7.2.3 Passing. To calibrate a passing model, Dommerholt and Botma (1988) collected passing data on seven two lane rural highway sections in the Netherlands. This data and the results of several runs of the present model are shown in Figure 7.10. The sections from which the data were obtained had ideal road geometrics, but the traffic was not composed of 100 percent passenger cars, and was not balanced directionally. The traffic contained between 12 and 24 percent trucks, and the directional split mostly did not exceed 60/40.

The Netherlands data showed that above a two way flow of about 1000 vph, there is a considerable drop in passing frequency. Dommerholt and Botma suggest the following explanations:

> - drivers do not expect to gain much travel time by passing on these busy road sections and even stick to this tendency when volumes are not extreme;
> - drivers postpone their passing to nearby road sections where the opportunities are greater, e.g. at a doubling of the carriageway at an intersection with traffic lights.
>
> Both explanations refer only to drivers who know the road and its traffic conditions, which usually make up at least 80 percent of the population.

The driver prototypes of section 6 did not use such global conditions as road knowledge and traffic volume in their judgements and, consequently, while passing at low volumes was

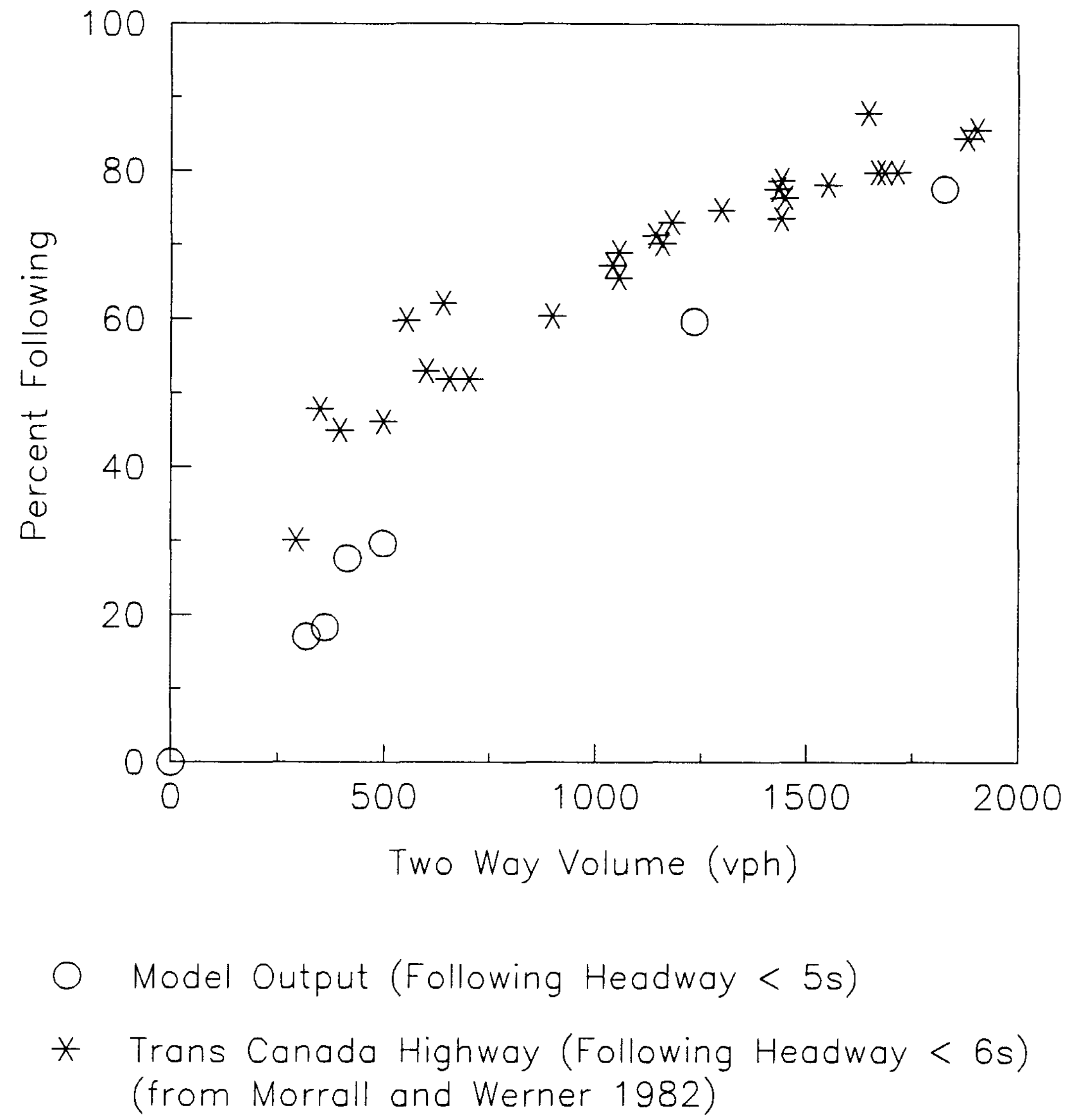


Figure 7.9 Percent Following Comparison to Trans Canada Highway

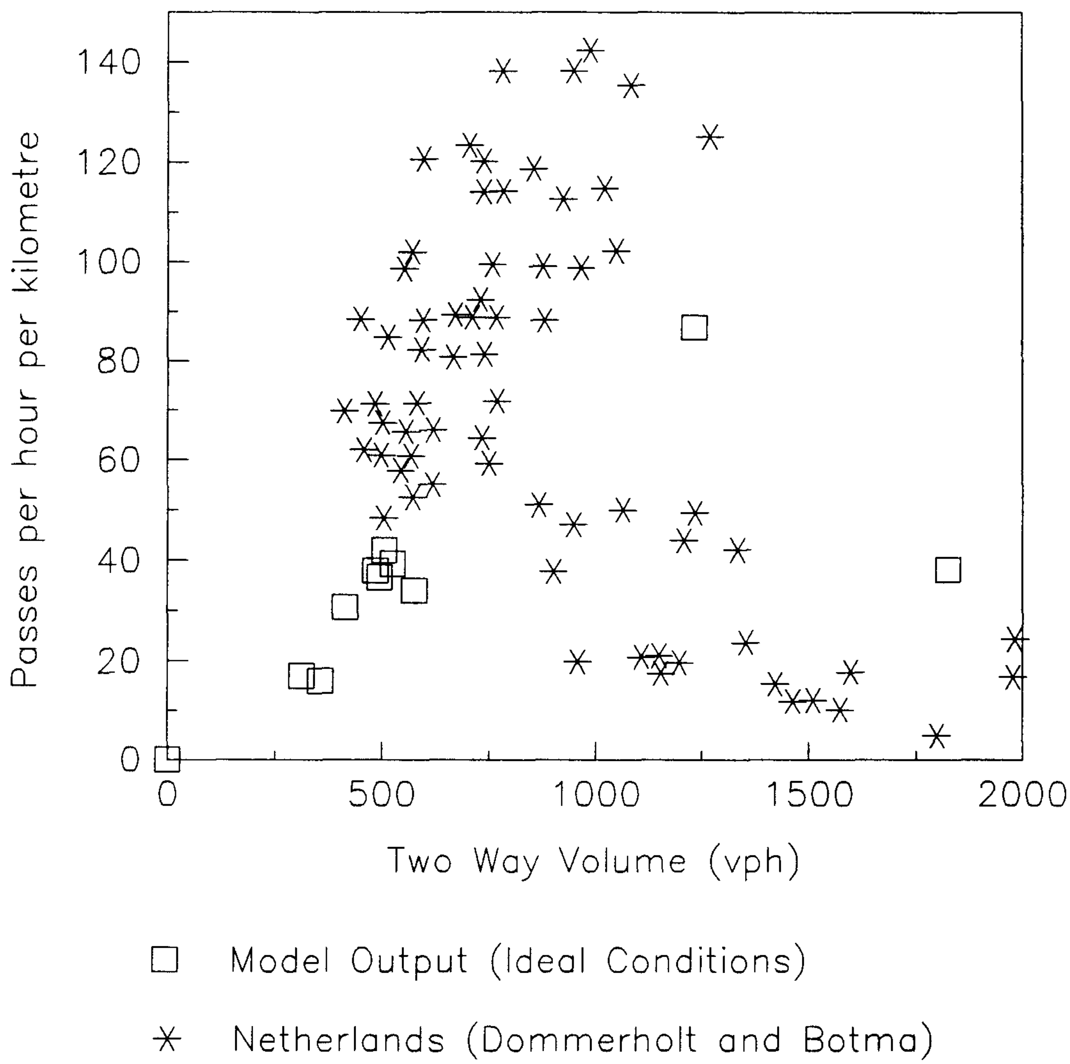


Figure 7 . 10 Passing Frequency Comparison to Field Data from Netherlands

similar in the model and the field, passing at high volumes was overestimated in the model. The inclusion of global conditions to rectify this is discussed in the recommendations section.

7.3 Other Outputs

In aggregately calibrating the model, it was better to test input-output mappings with field data or established standards than with examinations for reasonability. Generally, such data and standards were available for those measures most frequently used and considered most important. Tests involving these measures were discussed in sections 7.1 and 7.2. For completeness, however, the other input-output combinations of Table 7.1 were examined for reasonability, and the next three sections describe this, grouping the tests by the output measures in Table 7.1. All the runs in which volume was not a variable were completed at a two way traffic volume of about 500 vph.

7.3.1 Percent Following. As quoted at the beginning of section 7.1, the Highway Capacity Manual implies that percent following can be used interchangeably with 'time delay'. Strictly speaking, however, it the former which is measured or estimated. Time delay is a measure of the percentage of time one driver follows another *while travelling at a speed lower than his desired speed*. As a consequence, percent following, which is what can be observed in the field, is by definition equal to or greater than time delay, but not strictly the same thing. Because it is percent following that was estimated in the model, that term is used here.

Figure 7.5 showed the effects of volume on percent following. The other input factors, truck traffic, grades and directional split, also affected percent following in the model as shown in Figures 7.11 through 7.13. Figure 7.11 demonstrates that as the percentage of trucks in the traffic stream increased, so did the following percentage of passenger cars. Trucks' following percentages increased at a lower rate and were generally much lower than those of cars. This is because most car drivers move faster and tend to catch up to the trucks rather than vice versa. The effect of the low following percentages of the increasing numbers of trucks was to limit the rise of the overall following percentage to a rather low rate.

Figure 7.12 shows the effects of grades on percent following for trucks, passenger cars and the two combined. The effect of grades on the following percentages of trucks was to lower them. This was expected because of the increasing speed capability differential between cars and trucks on grades and the catching up effect noted previously. As grades rose, the following percentages of passenger cars eventually fell somewhat. This may be counter-intuitive but is perhaps caused by modelling moderate traffic flows on long (2 km) grades. On these grades, trucks were forced to slow down considerably and were therefore easier to pass. The long straight grades allowed for adequate sight distance before the crest to accommodate this passing. The results, in fact, showed that passing of trucks increased on the higher grades. Passing, however, was only accommodated because of the moderate traffic levels. Presumably, at higher traffic volumes, cars would not be able to pass the slow moving trucks because of opposing traffic and their following percentages would rise

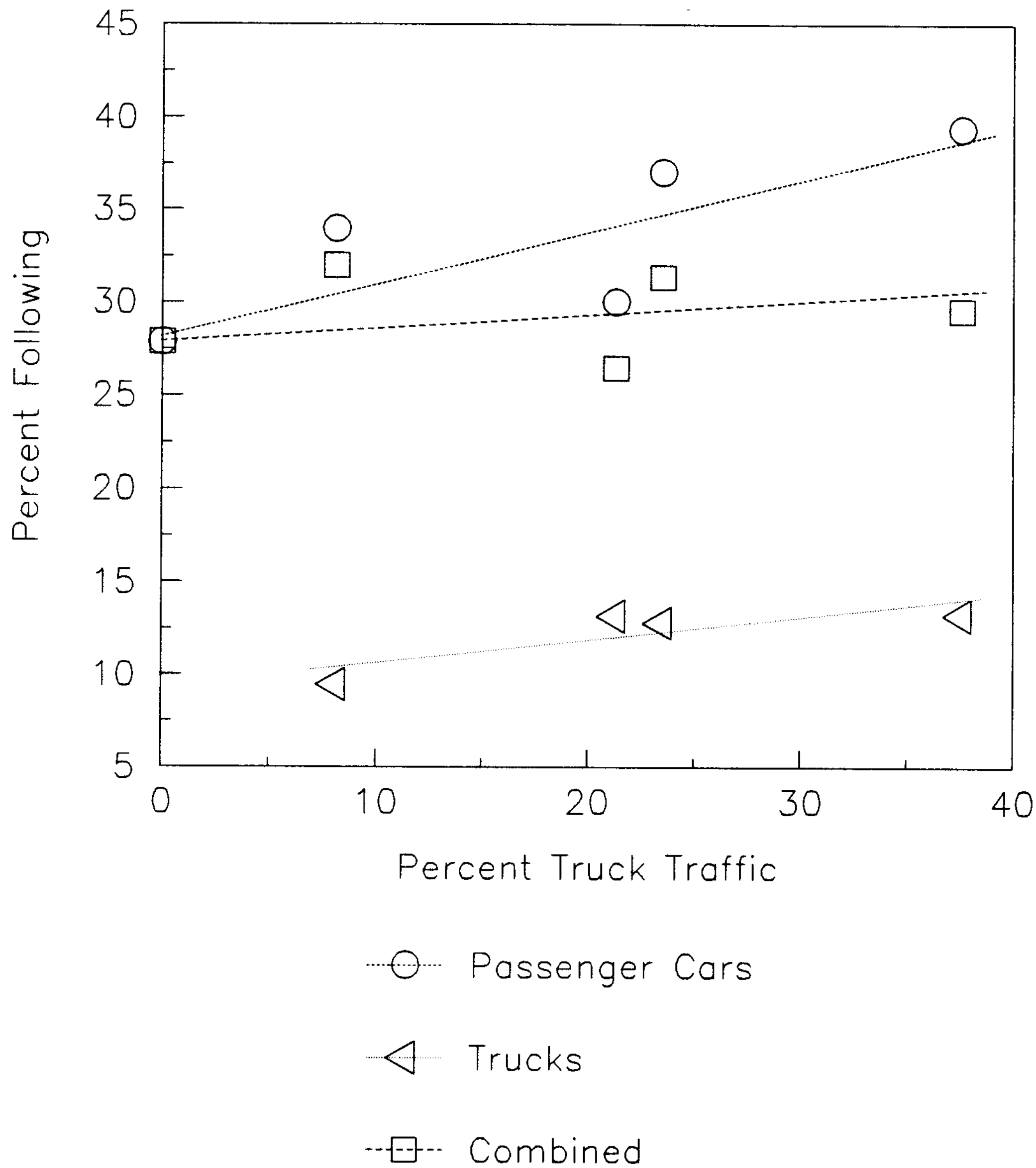


Figure 7.11 Truck Traffic Effects on Percent Following

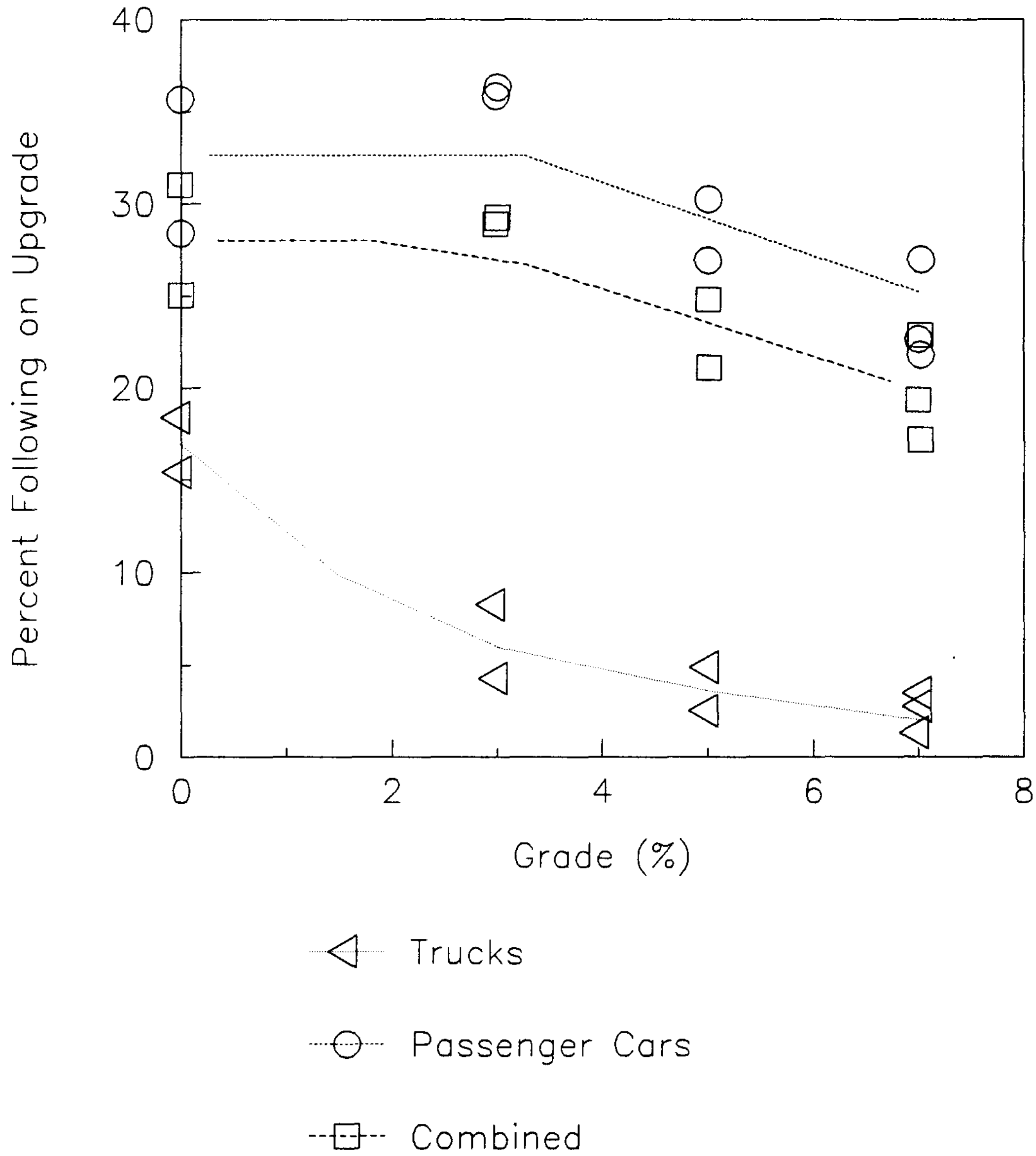


Figure 7 . 12 Grade Effects on Percent Following

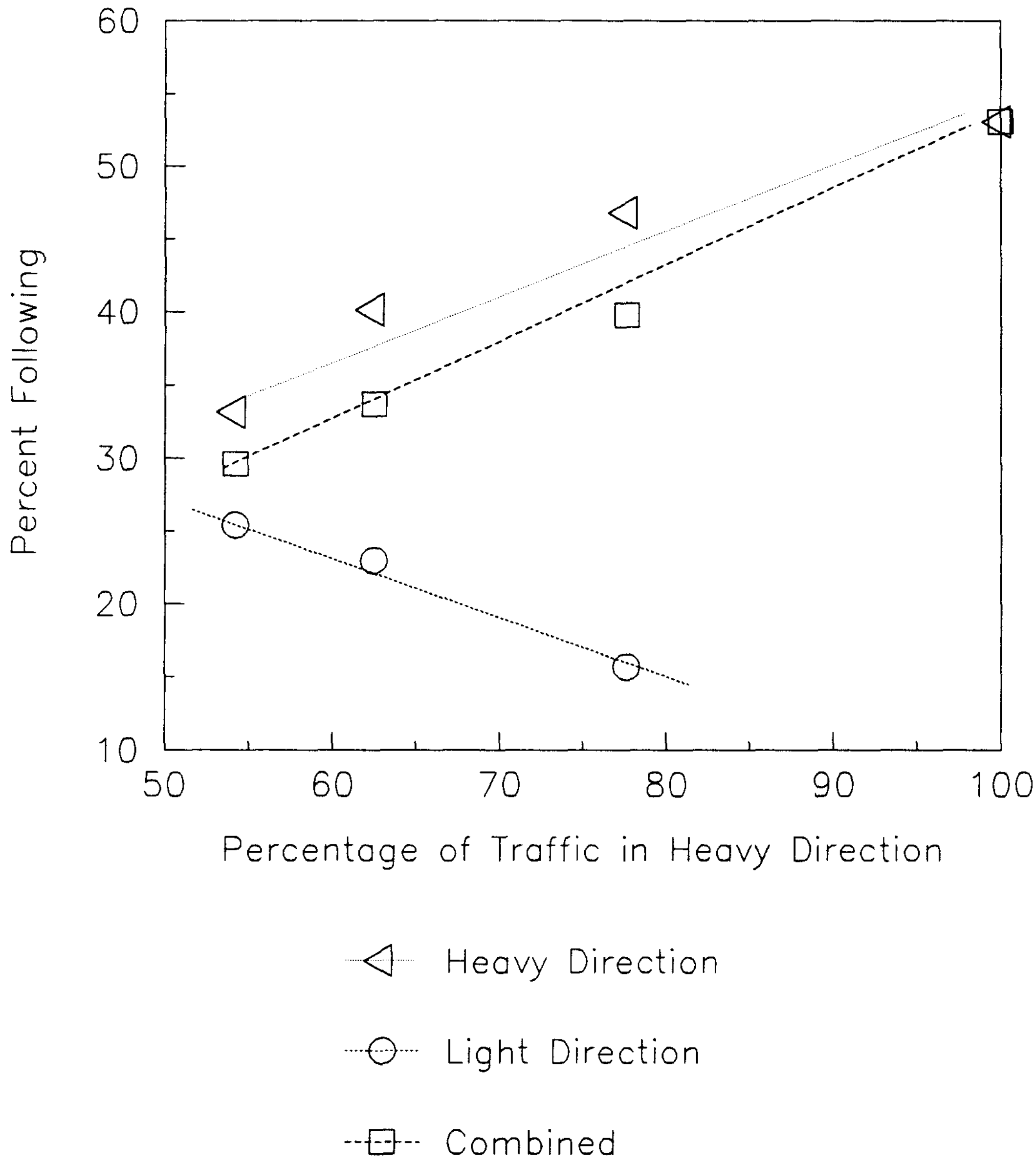


Figure 7.13 Directional Split Effects on Percent Following

accordingly.

Figure 7.13 shows that two direction (combined) following percentages were lowest for traffic balanced in both directions and increased as flows became more and more unbalanced. Note that, at 500 vph, the heavy direction and light direction following percentages intercepted the ordinate at the same point, approximately 28 percent. This was to be expected because at that point the flows were equal. This graph shows why it is important not to confuse percent following with time delay (as defined in the Highway Capacity Manual). Percent following was highest at the 0/100 directional split where passing was not inhibited by opposing traffic. Logically percent time delay should have been lowest at this point. In fact, it probably was, as evidenced by passing estimates (to be examined in section 7.3.3) which showed that passing frequencies were twice as high at the 0/100 split as at the 50/50 split. The higher following percentage at the 0/100 split was caused by routing the entire flow into one lane, rather than splitting it into two.

7.3.2 Speed Variance. Speed variance can be as important a measure in traffic analysis as mean speed. This is because speed differentials between vehicles can create hazardous situations and influence perceived levels of service as much as individual vehicle speeds.

In model runs under ideal conditions, speed variance was partially a function of traffic volume as shown in Figure 7.14. As expected, at higher volumes, speed variance estimates dropped relative to those at low volumes. The dropping speed variance was caused by

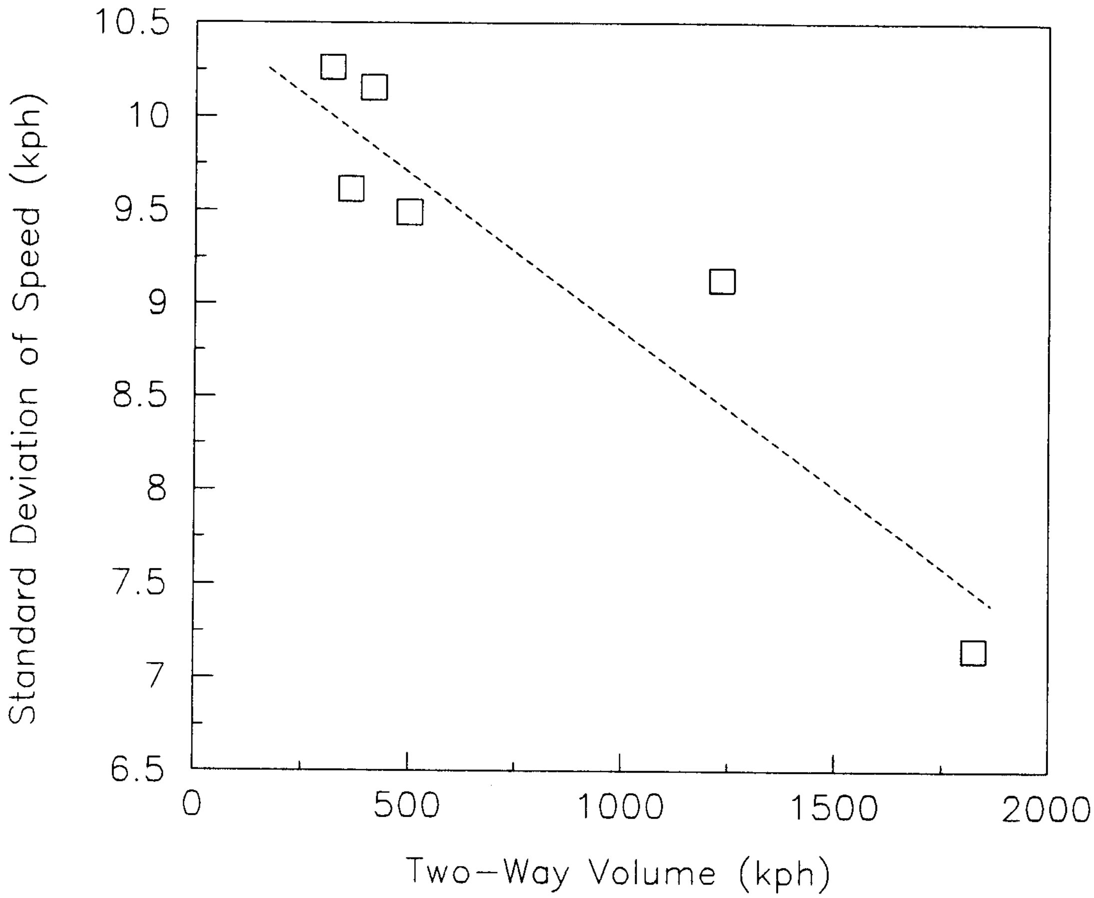


Figure 7 . 14 Volume Effects on Speed Variance

reduced passing opportunities for faster drivers with increasing opposing traffic. This meant that they were forced to travel at lower speeds in platoons behind slower drivers for longer periods of time.

Truck drivers tend to adopt lower speeds than drivers of passenger cars and consequently the addition of trucks to modelled traffic streams increased the speed variance at low to moderate traffic volumes. This is shown in Figure 7.15. Note that truck traffic had the effect of slowing down faster drivers so that the variance among passenger cars dropped with increasing truck percentage. At the same time, overall variance increased to a point and then levelled off as the speeds of cars and trucks approached each other. At higher traffic flows, levelling off and then decline of the overall variance would likely occur at lower truck percentages because of car drivers' inability to pass in the face of high opposing traffic. Platooning would increase and therefore a higher proportion of the traffic stream would travel at similar speeds.

The effect of trucks on speed variance is accentuated on grades as shown in Figure 7.16. This figure shows the results of a series of runs at 500 vph and 20 percent truck traffic. The overall speed variance rose with increasing grades as trucks were forced to use lower speeds compared to cars.

Model runs on level sections with 100 percent passenger cars and at various directional splits showed that directional split, at least at 500 vph, has little effect on speed variances.

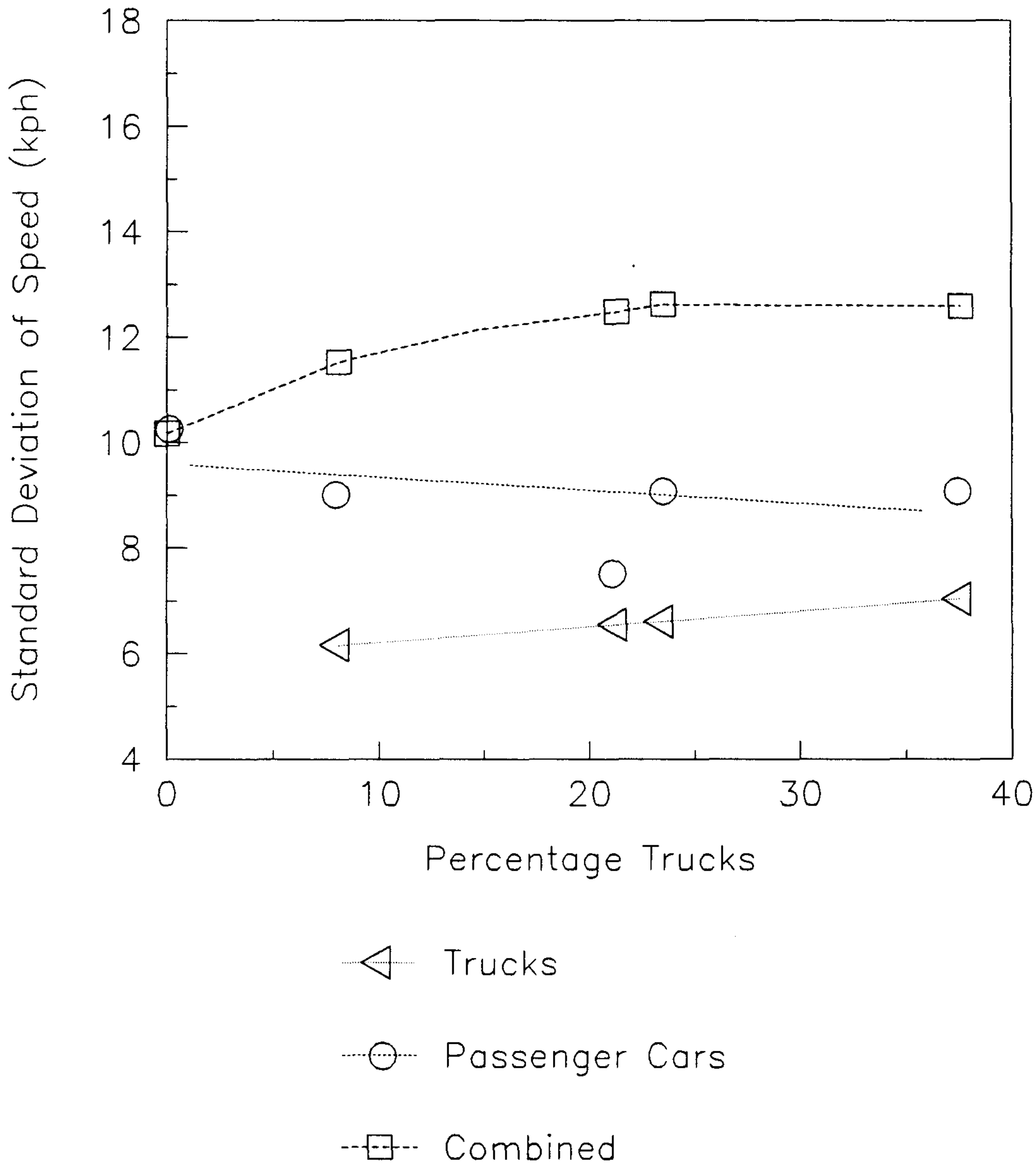


Figure 7 . 15 Effect of Trucks on Speed Variance

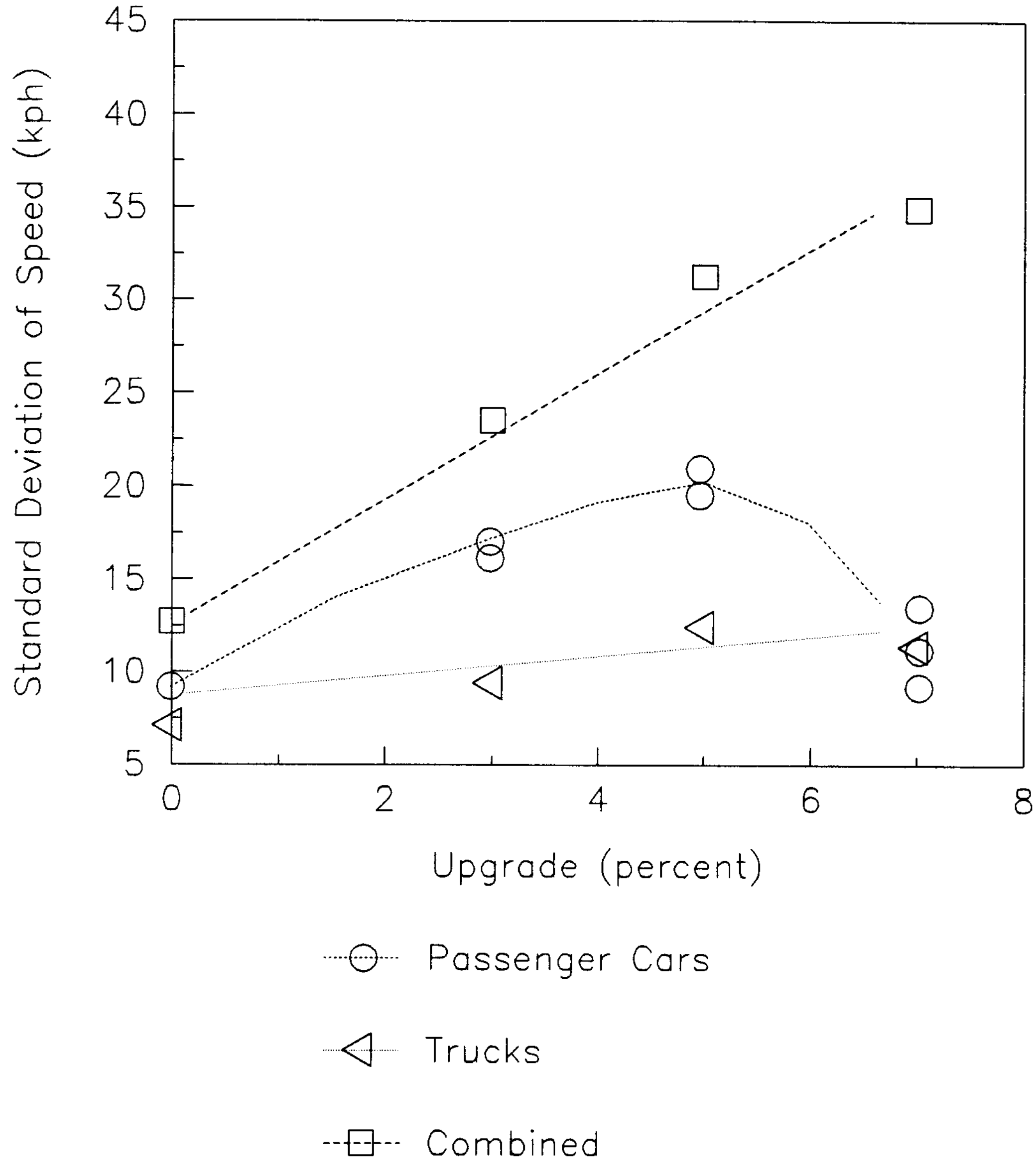


Figure 7 . 16 Grade Effects on Speed Variance

7.3.3Passing. Passing frequency estimates varied with traffic volume as demonstrated before, but they also varied with percentage truck traffic, grades and directional splits. The effect of the first of these, trucks, is shown in Figure 7.17. As expected, the model runs showed that passing increased with increasing truck traffic. This was because trucks tend to travel at speeds lower than those desired by most car drivers, prompting the latter to attempt more passing.

The effect of grades on passing at 500 vph and with 20 percent truck traffic is shown in Figure 7.18. As grades increased, passing on the upgrades decreased and that on the downgrades increased at about the same rate. The result was a relatively constant passing frequency for the roadway as a whole. It should be noted that the grades were 2 kilometres long so that sight distance limitations at the crests did not affect passing over much of the sections. As a consequence, passing on the upgrade did not decline very much with increasing grade. On shorter sections, it is expected that passing would decrease at a higher rate.

Model runs to demonstrate the effects of directional split on passing yielded results which were considered reasonable. Outputs are shown in Figure 7.19. The figure shows that as the percentage of traffic in the heavy direction increased, passing frequency rose in the heavy direction and dropped in the light direction. The increase in heavy direction passing had two causes. As the traffic flow in the heavy direction increased, there was more demand for passing (See Figure 7.10). Also, as the opposing traffic decreased, there were more gaps

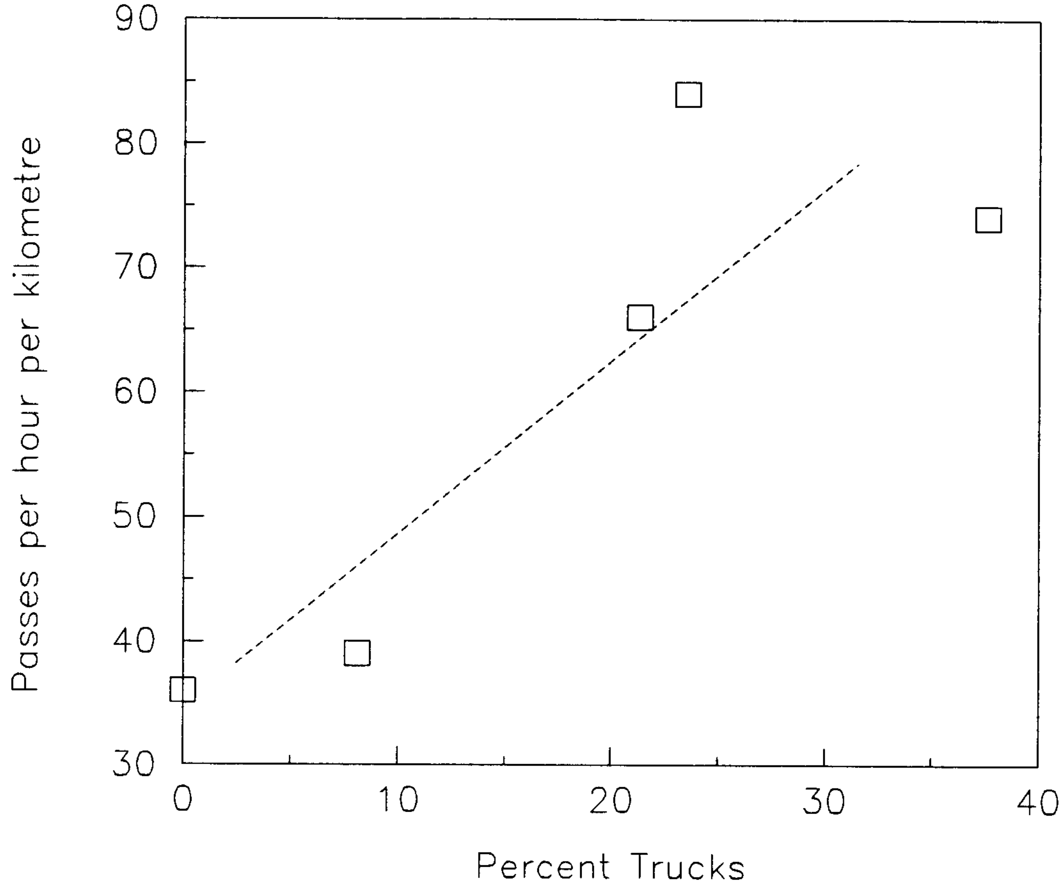


Figure 7 . 17 Effect of Trucks on Passing

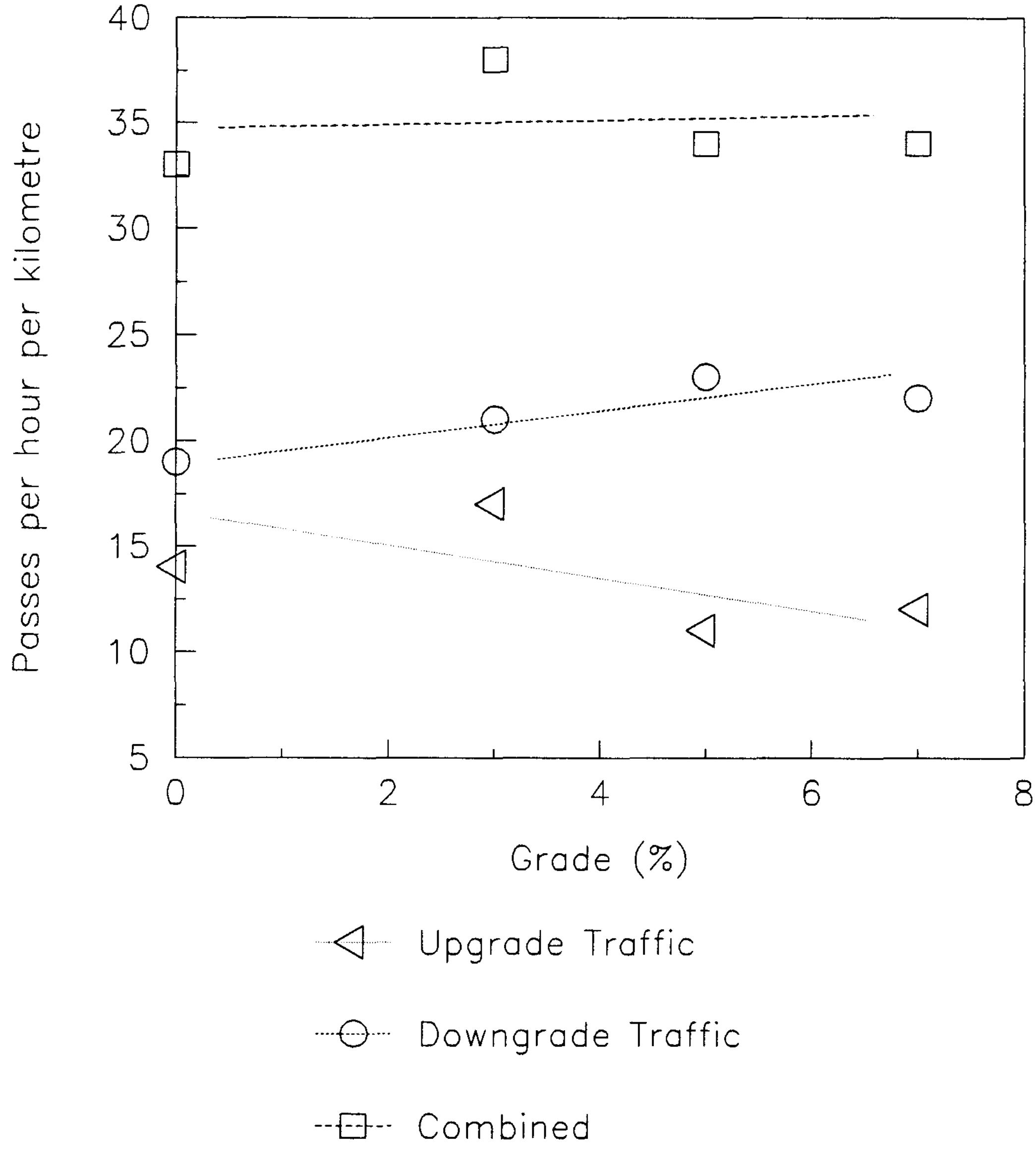


Figure 7 . 18 Grade Effects on Passing

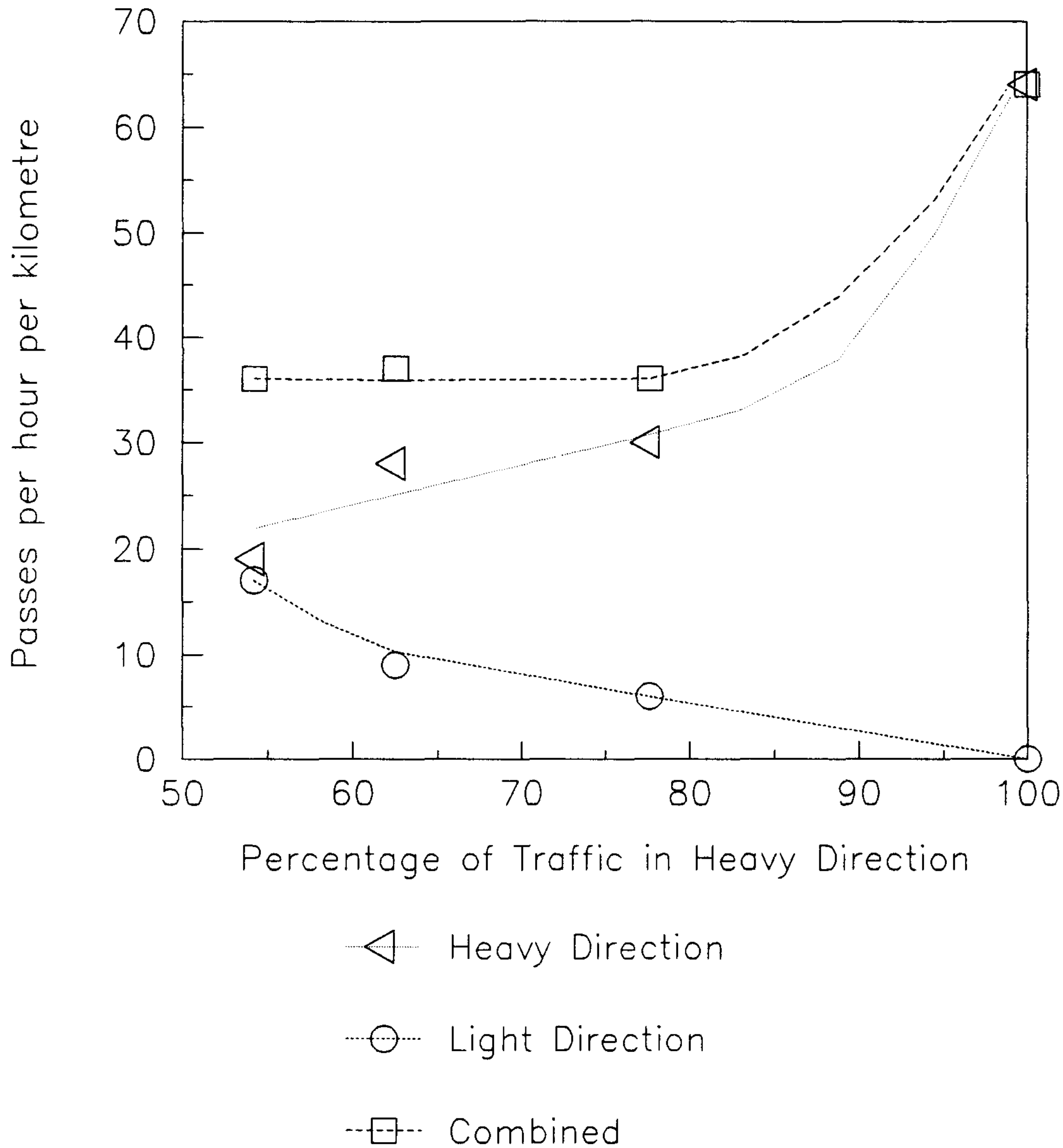


Figure 7 . 19 Directional Split Effects on Passing

to facilitate that passing. By the same token, the decrease in light direction passing can be explained by the dropping volume in that direction reducing demand and by the rising opposing volume reducing opportunity.

8 RESULTS AND APPLICATION

Traffic flow models are often designed to obtain parameter estimates in a particular real world problem, or for repeated use in solving a particular class of problems. The present model differs from these in that (1) the flexibility of network and traffic inputs allows for application in a wide variety of situations; and (2) the ability to not only calibrate drivers but define the structure of their behaviours allows for application in solving a wide variety of problems. This means that the same model can be used to investigate passing in a climbing lane section and to investigate turning movements at an intersection. It is just a matter of creating the appropriate road systems and programming driver prototypes with behaviour suited to the systems.

Because the model is not dedicated to solving a particular problem in a particular location, there was no predefined set of outputs for the model to generate. However, outputs one would potentially require can be classified into two categories: microscopic and macroscopic. Microscopic outputs give the actions of particular simulated drivers during, for example, passing manoeuvres or accidents. Macroscopic outputs are those dealing with traffic on a section of roadway or at a particular intersection and are measures of the aggregate behaviour of simulated drivers. Examples include speed distributions along a length of roadway and queue lengths at an intersection.

For applications which require microscopic outputs, the desire was to provide a facility for saving selected portions of a simulation run for later viewing and analysis. In this way, the particular behaviour of interest can be isolated from the rest of the simulation run in which the behaviour does not occur. Extracting incidents of particular behaviours is called importance sampling and section 8.1 discusses its implementation in the model.

For applications which require macroscopic outputs, there is no common vehicle of analysis such as importance sampling. For implementation, therefore, the desire was to provide a facility which allowed users to construct their own output measures using various statistical measures of driver, road and vehicle variables internal to the simulation. Section 8.2 discusses this.

8.1 Microscopic Outputs

Microscopic outputs, as noted previously, detail the actions of particular simulated drivers. Analysis of a particular behaviour in a simulation as in the real world requires observation of its occurrence. Therefore, the model was designed to extract incidents in which a behaviour in question occurred and save them for later viewing. To date, however, a generalized method of defining the desired behaviour has not been developed. Rather, the model has been set up to extract 'critical situations' which are defined as those in which the stress level in at least one driver rises above a user defined threshold.

The facility for viewing the incidents operates like a video cassette recorder, in that one can play and replay incidents, viewing a graphical recreation of them in plan view. One can step through incidents one second at a time in either the forward or backward direction, and pause to examine a single frame in more detail. Though the model uses a 0.2 second scanning time, it saves incidents using one second intervals because of the large amount of data involved. In addition, at each one second frame, a user can examine any of the following variables:

For any driver:

eye focus location;

short term memory records of the active environment; and,

conscious or unconscious program in progress, and its stage of completion.

For any vehicle:

speed;

acceleration;

steering angle;

state of turn signals;

state of headlights;

state of flashers; and,

maximum acceleration capability on present grade.

In this way, incidents can be recreated and the analyst can examine the events as they unfold not only externally as one would in the real world, but internally from the point of view of the various drivers.

8.2 Macroscopic Outputs

Unlike microscopic outputs which can be displayed via the common vehicle of importance sampling, macroscopic outputs on road sections or intersections can be as varied as the applications to which the model is put. Therefore, a facility which allows a user to construct his own outputs from a set of statistical measures and variables internal to the simulation was the ideal way to facilitate macroscopic outputs. Such a facility has not been developed to date. In its stead, a set of predefined outputs was included for using the model in a particular application, roadway design and improvement. This application was chosen because it suited the example driver prototypes developed in section 6 and because it is considered a valuable use of the model.

Assistance in the design of new roads and improvement of existing ones is a potentially useful application of the model because, in these processes, alternative designs are compared on three bases: (1) cost; (2) environmental impact; and (3) traffic operation (Wu and Heimbach 1981). This model can be used in the third of these.

A simulation of highway traffic provides particular advantages.

> Traffic operation studies for preliminary design alternatives include estimation of traffic performance resulting from vehicle-roadway interactions. Due to its complexity and its often random nature, traffic flow on highways cannot be characterized in a straightforward manner. The highway engineer resorts to empirical relations based on real-world observations. Even though these relations provide a general idea of the nature of traffic operations, they are not sensitive enough to detect either roadway traffic-flow interactions for any individual design alternative or the differences in these interactions between two or more alternative designs (Wu and Heimbach 1981).

On a particular project, the model would be used after formulating design alternatives. Traffic operations can be simulated on each alternative and the output used to provide statistics for evaluating each alternative and comparing among them. Desirable output statistics could include level of service, air pollution, road user costs as well as traffic movements at critical or problem locations (microscopic outputs).

In a study of the effects of improvements in passing manoeuvres on traffic flow on two lane roads, Cassel and Janoff (1968) incorporated the following output estimates:

- Volume;
- Average speed and standard deviation;
- Number of attempted and completed passes and aborts;
- Amount of delay (seconds) suffered by the vehicles which leave the road;
- Number of possible accident conditions termed emergency indicators, when some type of evasive (ie. acceleration or deceleration) action must be taken, during a passing manoeuvre;
- Number of projected accidents, when no evasive action can be taken in the model to deter a possible accident;
- Average safety margin (average time to meeting of oncoming car after completion of a pass); and,
- Number of speed change cycles.

Though this list provided the basis of the variables included as outputs in the present model, some changes were made because of the differing purposes of the Cassel and Janoff model and the application demonstrated here. While the former was concerned with passing, the latter dealt with traffic flow efficiency and levels of service. The output variables used included the following:

space mean speed;

time mean speed (and its standard deviation);

density;

traffic flow;

mean acceleration (and its standard deviation);

fuel consumption (and its standard deviation);

accident rate;

passing rate by vehicle type;

passing rate of vehicle type;

mean delay using a user defined following headway (and its standard deviation); and,

mean travel time (and its standard deviation).

After a simulation run, each of these output measures can be viewed in one of two ways, either as graphed profiles along user selected sections of road, or as totals or averages for those sections. For a particular section of road, presented outputs can be differentiated by direction of flow, by lane and by vehicle type, or aggregated across any or all of these.

It may be desired to run several simulations on each design alternative, each with a different mix or quantity of traffic. Outputs from these runs can then be incorporated into an objective function used to evaluate design alternatives, from which choices can be narrowed down or a final design selected.

9 CONCLUSION

The purpose of this work was to overcome the limitations of using conventional mathematics to model human automobile drivers. The result is a road simulation tool box in which there is a unique driver model suitable for many purposes and situations. Also developed was a framework within which driver behaviour can be structured into a model and used for analysis. The skeleton driver did not stray too far from its psychological underpinnings. It did not make any assumptions about the structure of driving behaviour and therefore does not bias the model.

The skeleton driver model was combined with an uncalibrated model of vehicle acceleration and steering capabilities and the building blocks of road systems to provide a tool box for use in traffic flow modelling. In order to model a particular traffic stream on a particular road network, a user first develops a driving population and a vehicle population. In building a driver prototype, one starts with the skeleton driver model and specifies environmental pattern recognition and anticipation abilities, and conscious and unconscious actions taken in various situations. These are specified using fuzzy if-then rules.

The vehicle population is developed by creating a set of vehicle prototypes, each of which is a copy of the acceleration and steering capabilities model calibrated to a different vehicle type. Once the driver and vehicle prototypes are developed, their frequencies of occurrence in the population must be specified, and then they can be used in simulations on user defined road networks.

Both microscopic outputs (dealing with the movements of particular vehicles) and macroscopic outputs (aggregated traffic stream measures) can be generated from simulation runs. Microscopic outputs, which are used to study individual behaviours, include driver eye focus locations, short term memories and actions in progress, and vehicle speeds, accelerations, steering angles, turn signals, headlights, flashers and present maximum acceleration capabilities. Macroscopic outputs, which are obtained on any section or sub-section of the network, include space mean speed, time mean speed, density, traffic flow, acceleration, fuel consumption, accident rate, passing rates, percent following, and travel time.

In order to demonstrate this process, driver and vehicle prototypes were developed for use in simulating traffic on two lane rural highways. Outputs of simulation runs with these prototypes were tested to show the effectiveness of the process in modelling a traffic flow situation. The important conclusion from these tests was not that the driver and vehicle prototypes effectively modeled traffic on two lane rural highways (which they did) but that the process of their development based on the skeleton driver model and vehicle capabilities model was validated.

Thus the application of fuzzy logic to traffic flow modelling proved an effective method of analyzing the operations of road networks, and the present model provided a good foundation for that application. The most important difference between the present model and other traffic simulations is the model's starting point: psychology. Lying in this difference is the most important benefit of using this model to analyze traffic flow. In starting from a psychological basis, one cannot make the implicit assumption that people operate like machines and according to conventional mathematics. This model forces one to explicitly lay out all abilities and behaviours in all situations so that the modeller is keenly aware of both the capabilities and limitations of the driver prototypes he creates. This will be a humbling experience for analysts who are used to specifying conventional equations and then pushing them to extremes as if they were testing the responses of an artificial system. Psychologists have long recognized the limitations of our knowledge of human behaviour. It is time that engineers did also.

10 RECOMMENDATIONS

Researchers and analysts can make improvements and extensions to this work in a number of areas. Firstly, and most obviously, they can develop driver and vehicle prototypes for application to different types of driving such as freeways, urban arterials and intersections. They can then calibrate them to local conditions and apply them to other analysis.

Secondly, there is a need for the psychological driver model itself to be updated to incorporate global variables. Such variables include traffic volumes and overall design standards. They play a role in the general strategies of drivers as opposed to manoeuvres in individual incidents. For example, a driver may be more reluctant to pass when he is part of a high volume of traffic because he would catch up to another vehicle shortly after passing.

Thirdly, as Figure 2.12 shows, the driver model lacks an integrator, a learning mechanism. There is a need to model short term learning in combination with the use of global variables. Short term learning temporarily affects driving. For example, a driver who runs over a pothole likely expects that there will be more and may adjust his overall driving strategy accordingly. By contrast, long term learning involves the permanent adjustment of driving strategy. It is a more gradual process and requires more reinforcement by repeated or dramatic incidents. Because individual simulated drivers make only one pass over the road network, long term learning is not of interest.

BIBLIOGRAPHY

Algea, Charles W. "A Development of a Conceptual Framework of the Driving Task." Human Factors, v 6 n 4, 1964.

Allen, T. M., H. Lunenfeld and G. J. Alexander. "Driver Information Needs." National Cooperative Highway Research Program Project 3-12, U. S. Federal Highway Administration, 1966.

Batz, Thomas M. "Evaluation of New Passing Zone Gore Design." Transportation Research Record, n 1239, 1989.

Bhise, Vivek D., James F. Meldrum, Lyman M. Forbes, Thomas H. Rockwell and Edward D. McDowell. "Predicting Driver Seeing Distance in Natural Rainfall." Human Factors, v 23 n 6, 1981.

Cassel, Arno and Michael S. Janoff. "A Simulation Model of a Two-Lane Rural Road." Highway Research Record, n 257, 1968.

Colbourn, Christopher J., Ivan D. Brown and Alan K. Copeman. "Drivers' Judgments of Safe Distances in Vehicle Following." Human Factors, v 20 n 1, 1978.

Cole, Barry L. and Philip K. Hughes. "A Field Trial of Attention and Search Conspicuity." Human Factors, v 26 p 299, 1984.

Council on Uniform Traffic Control Devices for Canada. Sign Pattern Manual. 1985.

Crawford, A. "The Overtaking Driver." Ergonomics, v 6 n 2, 1963.

Dommerholt, Wim and Hein Botma. "Model to Determine Operating Quality on Two-Lane Rural Roads." The ITE 58th Annual Meeting: Compendium of Technical Papers, Washington, D.C.: Institute of Transportation Engineers, 1988.

Donges, Edmund. "A Two-Level Model of Driver Steering Behavior." Human Factors, v 20 p 691, 1978.

Fanscher, P. S. and C. B. Winkler. "Downhill speed control and the use of retarders on heavy trucks." Braking of Road Vehicles, London: I Mech E Conference Publications, 1983.

Farber, Eugene. "Passing Behavior on Public Highways Under Daytime and Nighttime Conditions." Highway Research Record, n 292, 1969.

Fox, Phyllis and Frederick G. Lehman. Safety in Car Following: A Computer Simulation. Newark, N. J.: Newark College of Engineering, 1967.

Gerlough, Daniel L. and Matthew J. Huber. Traffic Flow Theory: A Monograph. Washington, D. C.: Transportation Research Board Special Report 165, National Research Council, 1975.

Glennon, John C. "New and Improved Model of Passing Sight Distance on Two-Lane Highways." Transportation Research Record, n 1195, 1989.

Godthelp, Hans, Gerard J. Blaauw and Jan Moraal. "Studies on Vehicle Guidance and Control." Transportation Research Record, n 1047.

Godthelp, Hans. "Vehicle Control During Curve Driving." Human Factors, v 28 n 2, 1986.

Godthelp, J. "Precognitive control: open- and closed-loop steering in a lane-change manoeuvre." Ergonomics, v 28 n 10, 1985.

Greenshields, B. D. "A Study of Traffic Capacity." Proceedings of the U.S. Highway Research Board, v 14 p 448, 1934.

Harwood, Douglas W., Andrew D. St. John and Davey L. Warren. "Operational and Safety Effectiveness of Passing Lanes on Two-Lane Highways." Transportation Research Record, n 1026, 1985.

Hay, William W. An Introduction to Transportation Engineering. Toronto: John Wiley and Sons, 1977.

Herrin, Gary D. and John B. Neuhardt. "An Empirical Model for Automobile Driver Horizontal CUrve Negotiation." Human Factors, v 16 n 2, 1974.

Highway Capacity Manual. Washington, D.C.: Transportation Research Board Special Report 209, National Research Council, 1985.

Hills, Brian L. "Vision, visibility and perception in driving." Perception, v 9 p 183, 1980.

Hoban, C. J., G. J. Fawcett and G. K. Robinson. A Model for Simulating Traffic on Two Lane Rural Roads: User Guide and Manual for TRARR Version 3.0. Victoria, Australia: Australian Road Research Board, 1985.

Homburger, Wolfgang S. "An Analysis of Safety at Upgrade Terminals of Climbing Lanes on Two-Lane Highways." Transportation Research Record, n 1122, 1987.

Hughes, P. K. and B. L. Cole. "What attracts attention when driving?" Ergonomics, v 29 n 3, 1986.

Hutchinson, B. G. and D. J. Parker. "Parametric analyses of large truck braking efficiencies." Canadian Journal of Civil Engineering, v 16 n 2, 1989.

Institute of Traffic Engineers. Transportation and Traffic Engineering Handbook. Ed. John E. Baerwald. Englewood Cliffs, N. J.: Prentice-Hall Inc., 1976.

Johansson, G. and F. Backlund. "Drivers and Road Signs." Ergonomics, v 13 n 6, 1970.

Johansson, Gunnar and Kare Rumar. "Drivers' Brake Reaction Times." Human Factors, v 13 n 1, 1971.

Kennedy, Norman, James H. Kell and Wolfgang S. Homburger. Fundamentals of Traffic Engineering. Berkeley, Cal.: Institute of Transportation and Traffic Engineering, University of California, 1966.

Koppa, Rodger J. and Gordon G. Hayes. "Driver Inputs during Emergency or Extreme Vehicle Maneuvers." Human Factors, v 18 n 4, 1976.

Kramer, U. and G. Rohr. "A model of driver behaviour." Ergonomics, v 25 n 10, 1982.

Kramer, Ulrich. "On the Application of Fuzzy Sets to the Analysis of the System Driver-Vehicle-Environment." Automatica, v 21 n 1, 1985.

Lee, Sang M. Introduction to Management Science. Toronto: The Dryden Press, 1983.

Lefeve, B. A. "Speed Characteristics on Vertical Curves." Proceedings of the U. S. Highway Research Board, v 32 p 395, 1953.

Lerner, Herman D., Henry E. Abbott Jr. and Robert B. Sleight. "Following Distance on the Highway Related to Speed, Trip Duration, Traffic, and Illumination." Human Factors, v 6 n 4, 1964.

Lindqvist, Torbjorn. "Observations of driving behaviour on a horizontal curve. Speed adaptation and vehicle alignment." Nordic Road and Transport Research, n 1, 1992.

Lindsay, Peter H. and Donald A. Norman. Human Information Processing: An Introduction to Psychology. San Francisco: Academic Press, 1977.

Lucas, G. G. Road Vehicle Performance: Methods of Measurement and Calculation. New York: Gordon and Breach Science Publishers, 1986.

Lyons, T. J., H. Rainford, J. R. Kenworthy and P. W. G. Newman. "Fuel and time implications of merging traffic at freeway entrances." Applied Mathematical Modelling, v 12, June 1988.

Matson, T. M. and T. W. Forbes. "Overtaking and Passing Requirements as Determined from a Moving Vehicle." Proceedings of the U.S. Highway Research Board, v 18 n 1, 1938.

McLean, John. "Driver speed behaviour and rural road alignment design." Traffic Engineering and Control, v 22 n 4, 1981.

McLean, John R. Two Lane Highway Traffic Operations: Theory and Practice. New York: Gordon and Breach Science Publishers, 1989.

Morrall, John F. and Al Werner. "Measurement of level of service for two-lane rural highways." Canadian Journal of Civil Engineering, v 9 p 385, 1982.

Mourant, R. R., T. H. Rockwell and N. J. Rackoff. "Drivers' Eye Movements and Visual Workload." Highway Research Record, n 292, 1969.

Mourant, Ronald R. and Thomas H. Rockwell. "Mapping Eye-Movement Patterns to the Visual Scene in Driving: An Exploratory Study." Human Factors, v 12 n 1, 1970.

Mourant, Ronald R. and Thomas H. Rockwell. "Strategies of Visual Search by Novice and Experienced Drivers." Human Factors, v 14 n 4, 1972.

Navin, F.P.D. Positive Guidance. Victoria, B.C.: Highway Safety Branch, B.C. Ministry of Transportation and Highways, 1992.

Navin, F.P.D. "Truck braking distances and speed estimates." Canadian Journal of Civil Engineering, v 13 p 412, 1986.

Ohene, Frederick A. and Siamak A. Ardekani. "Minimum Passing Sight Distance for Completing or Aborting the Passing Maneuver." Institute of Transportation Engineers Journal, July 1988.

Olson, Paul L. and Michael Sivak. "Visibility Problems in Nighttime Driving." Accident Reconstruction Journal, November/December 1992.

Oppenlander, J. C. Variables Influencing Spot-Speed Characteristics. Washington, D. C.: National Research Council National Academy of Sciencies Special Report 89, 1966.

Papacostas, C. S. Fundamentals of Transportation Engineering. Englewood Cliffs, N. J.: Prentice-Hall, 1987.

Prisk, C. W. "Passing Practices on Rural Highways." Proceedings of the U.S. Highway Research Board, v 21, 1941.

Reid, L. D., E. N. Solowka and A. M. Billing. "A systematic study of driver steering behaviour." Ergonomics, v 24 n 6, 1981.

Riemersma, J. B. J. "Perceptual Factors in Driving Perception of Charactistics of Horizontal Curves." Proceedings of 14th Australian Road Research Board Conference, n 4.

Robinson, Gordon H., Donald J. Erickson, Gregory L. Thurston and Richard L. Clark. "Visual Search by Automobile Drivers." Human Factors, v 14 n 4, 1972.

St. John, A. D. and D. R. Kobett. Grade Effects on Traffic Flow Stability and Capacity. Washington, D. C.: National Research Council National Cooperative Highway Research Program Report 185, 1978.

Shelby, M. D. and P. R. Tutt. "Vehicle Speed and Placement Survey." U. S. Highway Research Board Bulletin, n 170, 1958.

Shinar, David. Psychology on the Road: The Human Factor in Traffic Safety. Toronto: John Wiley and Sons, 1978.

Shinar, David and Amos Drory. "Sign Registration in Daytime and Nighttime Driving." Human Factors, v 25 n 1, 1983.

Shulman, Harvey G. and Seth N. Greenberg. "Perceptual Deficit due to Division of Attention Between Memory and Perception." Journal of Experimental Psychology, v 88 n 2, 1971.

Shulman, Harvey G. and Ronald P. Fisher. "Expected Value as a Determinant of the Distribution of Attention." Journal of Experimental Psychology, v 93 n 2, 1972.

Sivak, Michael, David V. Post, Paul L. Olson and Robert J. Donohue. "Driver Responses to High-Mounted Brake Lights in Actual Traffic." Human Factors, v 23 n 2, 1981.

Stohner, Walter R. "Speeds of Passenger Cars on Wet and Dry Pavements." U. S. Highway Research Board Bulletin, n 139, 1956.

Summala, H. and J. Hietamaki. "Drivers' immediate responses to traffic signs." Ergonomics, v 27 n 2, 1984.

Sutton, Robert. Modelling Human Operators in Control System Design. Toronto: John Wiley and Sons Inc, 1990.

Taragin, A. "Effect of Roadway Width on Traffic Operations- Two-Lane Concrete Roads." Proceedings of the U. S. Highway Research Board, v 24, 1944.

Taragin, A. "Driver Performance on Horizontal Curves." Proceedings of the U. S. Highway Research Board, v 33, 1954.

Taylor, Donald H. "Accidents, Risks, and Models of Explanation." Human Factors, v 18 n 4, 1976.

Triggs, Thomas J. "The Influence of Oncoming Vehicles on Automobile Lateral Position." Human Factors, v 22 n 4, 1980.

U. S. Department of Transportation, Federal Highways Administration, Bureau of Public Roads. Overtaking and Passing on Two-Lane Rural Highways. Washington, D. C.: United States Government Printing Office, 1967.

U. S. Department of Transportation, Federal Highways Administration, Bureau of Public Roads. Overtaking and Passing Under Adverse Visibility Conditions: Summary Report. Washington, D. C.: Clearinghouse Publications, 1969.

U. S. Department of Transportation, Federal Highways Administration, Bureau of Public Roads. Overtaking and Passing Under Adverse Visibility Conditions: Driver Judgement and Decision Making in Accelerative Passing Situations on Two Lane Rural Highways. Washington, D. C.: Clearinghouse Publications, 1969.

U. S. Department of Transportation, Federal Highways Administration, Bureau of Public Roads. Overtaking and Passing Under Adverse Visibility Conditions: Driver Judgement and Decision Making in Overtaking and Flying Pass Situations. Washington, D. C.: Clearinghouse Publications, 1969.

U. S. Department of Transportation, Federal Highways Administration, Bureau of Public Roads. Overtaking and Passing Under Adverse Visibility Conditions: Driver Control During Overtaking and Passing. Washington, D. C.: Clearinghouse Publications, 1969.

Watanatada, Thawat, Ashok M. Dhareshwar and Paulo Roberto S. Rezende Lima. Vehicle Speeds and Operating Costs: Models for Road Planning and Management. Baltimore, Md.: The Johns Hopkins University Press, 1987.

Webster's New Collegiate Dictionary. Springfield, Massachusetts: G & C Merriam Company, 1980.

Wierda, Marcel. "The man and his wheel: cognitive and perceptual aspects." VTI Report 372A, 1991.

Willey, William E. "Survey of Downhill Speeds of Trucks on Mountain Grades." Proceedings of the U. S. Highway Research Board, v 30, 1951.

Wilson, Tay and W. Best. "Driving Strategies in Overtaking." Accident Analysis and Prevention, v 14 n 3, 1982.

Wright, Paul H. and Norman J. Ashford. Transportation Engineering. Toronto: John Wiley and Sons, 1989.

Wu, Shie-Shin and Clinton L. Heimbach. "Simulation of Highway Traffic on Two-Lane, Two-Way Rural Highways." Transportation Research Record, n 806, 1981.

Zwahlen, Helmut T. "STOP AHEAD and STOP Signs and Their Effect on Driver Eye Scanning and Driving Performance." Transportation Research Record, n 1168, 1988.

Zwahlen, Helmut T. "Eye Scanning Rules for Drivers - How do They Compare With Actual Observed Eye Scanning Behavior?" VTI Report 372A, 1991.

A1 INTRODUCTION TO FUZZY LOGIC

Fuzzy set theory is a superset of classical set theory rather than a sub-set. In classical set theory, elements in a domain either belong to a set entirely or do not belong at all. Fuzzy set theory includes both this possibility and the possibility that an element can belong to a set to a certain degree and at the same time not belong to that set to another degree. Examples of classical sets are the designations male and female. Leaving aside androgynous people, a person is either 100 percent male or 0 percent male. An example of a fuzzy set is the notion of a tall person. A given person can be moderately tall or somewhat tall, thus being considered partially a tall person and partially not a tall person.

Sets are defined by membership functions which, for every element in a domain, gives a degree of membership between 0 and 1. Membership functions for classical sets, also known as crisp sets, are integer functions whose range consists of only two values, 0 and 1. Fuzzy membership functions allow for fractional numbers in between. The two are shown graphically in Figure A1.1.

Sets can also be described mathematically. For discrete domains, a set is described as follows:

$$A = \mu_A(d_1)|d_1 \cup \ldots \cup \mu_A(d_n)|d_n = \sum_{i=1}^{n} \mu_A(d_i)|d_i \qquad \text{(A1.1)}$$

where A = fuzzy set
μ_A = membership function for set A

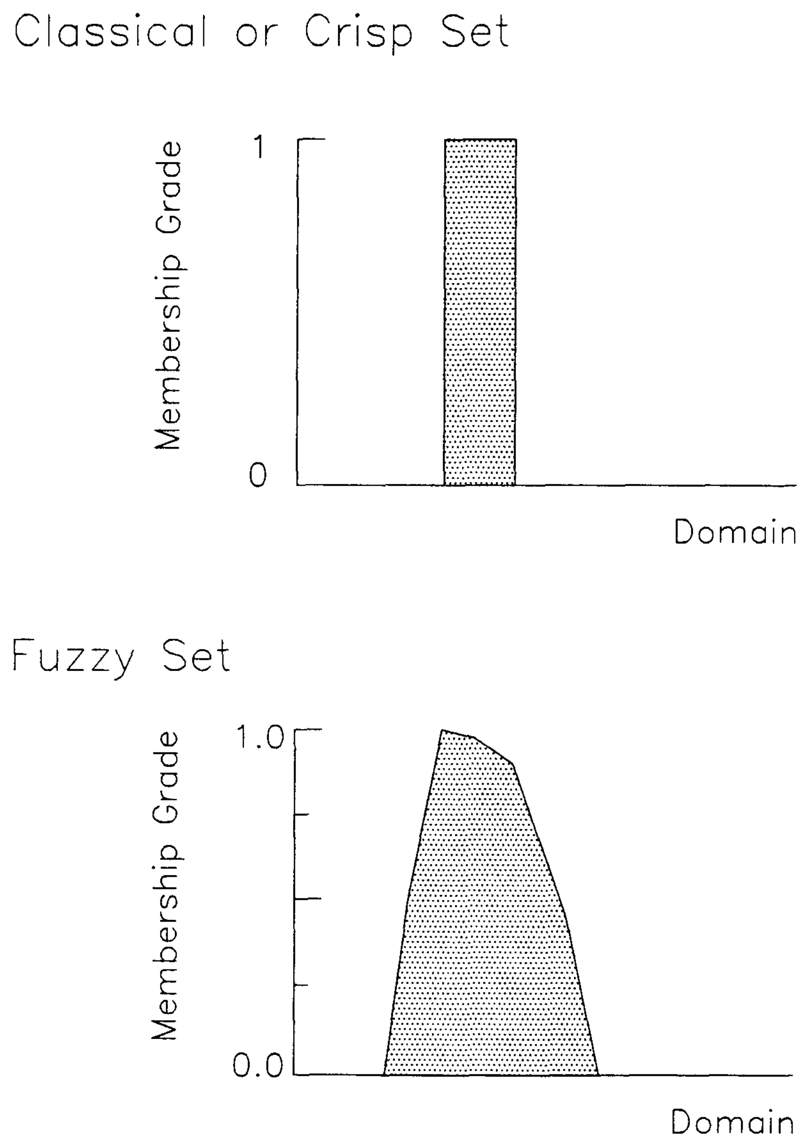


Figure A1. 1 Classical and Fuzzy Sets

d_i = element of domain
n = number of elements in domain

For continuous domains, the set is written as:

$$A = \int_D \mu_A(d)|d \tag{A1.2}$$

where A = fuzzy set
μ_A = membership function for set A
D = domain

Fuzzy sets are often used to represent words which have descriptive meanings such as small, medium, large, wide, narrow, close or far. For example, consider a parking lot which has 10 spaces. We can create fuzzy sets for the descriptive terms small, medium and large number of spaces available. The domain is discrete and includes the integers from 0 to 10. The sets may be defined as follows:

$$\text{Large} = 0.0|0 \cup 0.0|1 \cup 0.0|2 \cup 0.0|3 \cup 0.0|4 \cup 0.1|5 \cup 0.3|6 \cup 0.5|7 \cup 0.7|8 \cup 0.9|9 \cup 1.0|10 \tag{A1.3}$$

$$\text{Medium} = 0.0|0 \cup 0.1|1 \cup 0.3|2 \cup 0.5|3 \cup 0.7|4 \cup 1.0|5 \cup 0.7|6 \cup 0.5|7 \cup 0.3|8 \cup 0.1|9 \cup 0.0|10 \tag{A1.4}$$

$$\text{Small} = 1.0|0 \cup 0.9|1 \cup 0.7|2 \cup 0.5|3 \cup 0.3|4 \cup 0.1|5 \cup 0.0|6 \cup 0.0|7 \cup 0.0|8 \cup 0.0|9 \cup 0.0|10 \tag{A1.5}$$

These are shown in Figure A1.2 and will be used in subsequent examples.

Fuzzy Operations. Fuzzy sets can be related to each other using operators to produce another set. The most common operators are union and intersection. The union between

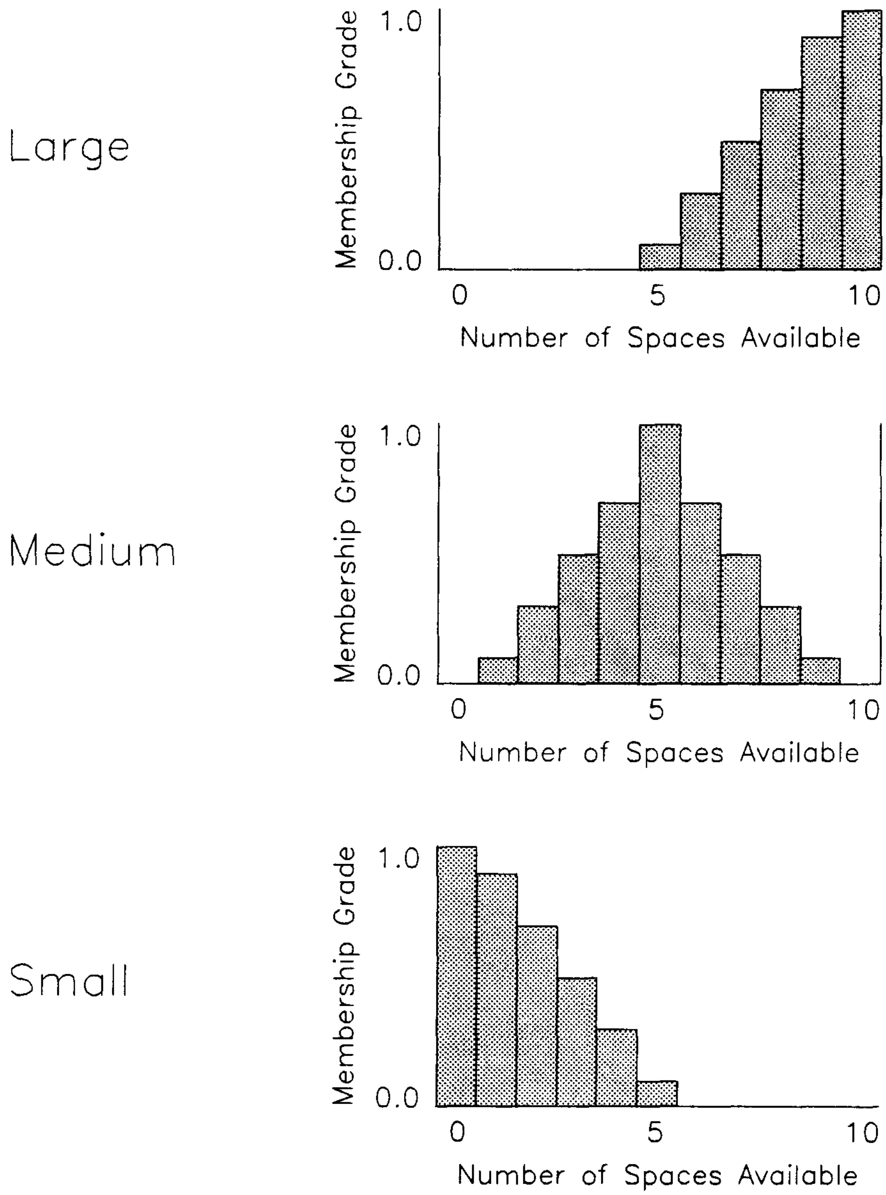


Figure A1.2 Example Fuzzy Sets

two fuzzy sets of the same domain produces another set in that domain and represents a connective 'OR' operation. The membership function of the resulting set is defined by:

$$\mu_{A \cup B}(d) = \max[\mu_A(d), \mu_B(d)] \tag{A1.6}$$

Using the parking lot example, the set of either a large or medium number of spaces available is calculated as shown:

Large or Medium = max [0.0,0.0]|0 U max [0.0,0.1]|1 U max [0.0,0.3]|2 U max [0.0,0.5]|3 U max [0.0,0.7]|4 U max [0.1,1.0]|5 U max [0.3,0.7]|6 U max [0.5,0.5]|7 U max [0.7,0.3]|8 U max [0.9,0.1]|9 U max [1.0,0.0]|10

= 0.0|0 U 0.1|1 U 0.3|2 U 0.5|3 U 0.7|4 U 1.0|5 U 0.7|6 U 0.5|7 U 0.7|8 U 0.9|9 U 1.0|10 (A1.7)

The intersection of two fuzzy sets represents the connective 'AND' operation and has a membership function as follows:

$$\mu_{A \cap B}(d) = \min[\mu_A(d), \mu_B(d)] \tag{A1.8}$$

The set of a medium to small number of spaces available in the parking lot example is determined here:

Small and Medium = min [1.0,0.0]|0 U min [0.9,0.1]|1 U min [0.7,0.3]|2 U min [0.5,0.5]|3 U min [0.3,0.7]|4 U min [0.1,1.0]|5 U min [0.0,0.7]|6 U min [0.0,0.5]|7 U min [0.0,0.3]|8 U min [0.0,0.1]|9 U min [0.0,0.0]|10

= 0.0|0 U 0.1|1 U 0.3|2 U 0.5|3 U 0.3|4 U 0.1|5 U

$$0.0|6 \cup 0.0|7 \cup 0.0|8 \cup 0.0|9 \cup 0.0|10 \qquad \text{(A1.9)}$$

The complement of a fuzzy set is sometimes used and represents the negation 'NOT'. The membership function of the complement is as shown:

$$\mu_{\neg A}(d) = 1 - \mu_A(d) \qquad \text{(A1.10)}$$

Again, using the parking lot example, the expression not a large number of spaces available is determined as follows:

$$\begin{aligned} \text{Not Large} &= [1 - 0.0]|0 \cup [1 - 0.0]|1 \cup [1 - 0.0]|2 \cup [1 - 0.0]|3 \cup \\ &\quad [1 - 0.0]|4 \cup [1 - 0.1]|5 \cup [1 - 0.3]|6 \cup [1 - 0.5]|7 \cup \\ &\quad [1 - 0.7]|8 \cup [1 - 0.9]|9 \cup [1 - 1.0]|10 \\ &= 1.0|0 \cup 1.0|1 \cup 1.0|2 \cup 1.0|3 \cup 1.0|4 \cup 0.9|5 \cup \\ &\quad 0.7|6 \cup 0.5|7 \cup 0.3|8 \cup 0.1|9 \cup 0.0|10 \qquad \text{(A1.11)} \end{aligned}$$

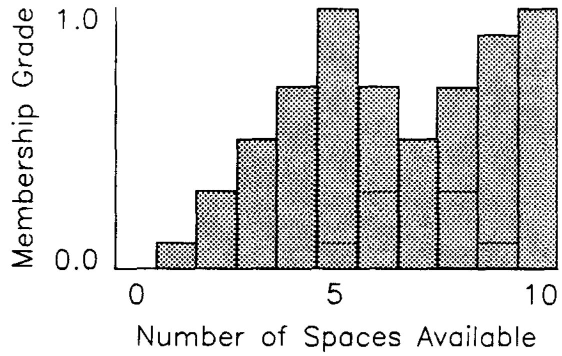
These sets are depicted in Figure A1.3.

Fuzzy relations. In a system in which there are fuzzy inputs and fuzzy outputs, the relationship between them is described using conditional statements known as fuzzy relations. For a fuzzy relation, the domains of the input and output variables are usually different and will be denoted as D and O respectively. The basic form of a fuzzy relation is as follows:

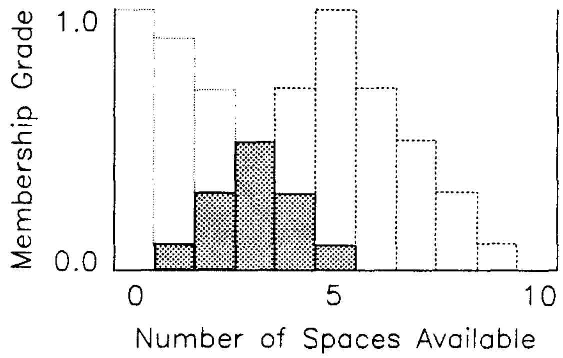
$$A \rightarrow B \qquad \text{(A1.12)}$$

A and B are fuzzy sets and the relation is read as 'If A then B'. A is known as the antecedent and B as the consequent. The relation is itself a fuzzy set in two dimensions and

Large or Medium



Small and Medium



Not Large

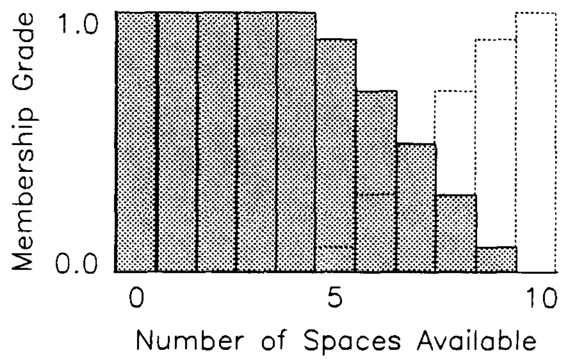


Figure A1.3 Fuzzy Operation Examples

is the Cartesian product of A and B. For finite and continuous domains, the membership functions of the relation R is defined by:

$$\mu_R(d,o) = \mu_{A\times B}(d,o) = \min[\mu_A(d),\mu_B(o)], \quad d\in D,\ o\in O \tag{A1.13}$$

$$\mu_R = \mu_{A\times B} = \int_{D\times O} \mu_A(d) \cap \mu_B(o) \,|(d,o) \tag{A1.14}$$

As an example, assume that there are two parking lots like the one described before and that both are used by people attending the same event. Suppose that a person at one of the lots can count the number of spaces available there and wishes to estimate the number available at the other lot based on this. What is necessary is to express the number of spaces available in one lot as a function of the spaces available in the other. There are no known exact relationships but some fuzzy ones can be expressed such as this:

> If there are a large number of spots available in one of the lots, there are a large or medium number of spots available at the other.

Recall that 'large number of spots available' was defined in equation A1.3 and 'large or medium number of spots available' was defined in equation A1.7. The relation between them is shown in matrix form and in graphical form in Figure A1.4.

A control strategy is defined as one or more fuzzy relations such as this which relate inputs to outputs in a system. In a control strategy, fuzzy relations are called rules. It is often necessary to form a combination of several rules, each of which is a relation R_i. If this is the case, individual relations are combined to give an overall one by calculating the union

$$R = LM \times L = \begin{bmatrix} \min[0.0,\ 0.0] & \min[0.0,\ 0.0] \ldots & \min[0.9,\ 0.0] & \min[1.0,\ 0.0] \\ \min[0.0,\ 0.1] & & & \min[1.0,\ 0.1] \\ \ldots & & & \ldots \\ \min[0.0,\ 0.9] & & & \min[1.0,\ 0.9] \\ \min[0.0,\ 1.0] & \min[0.0,\ 1.0] \ldots & \min[0.9,\ 1.0] & \min[1.0,\ 1.0] \end{bmatrix} = \begin{bmatrix} 0.0 & 0.0 \ldots & 0.0 & 0.0 \\ 0.0 & & & 0.1 \\ \ldots & \ldots & \ldots & \ldots \\ 0.0 & & & 0.9 \\ 0.0 & 0.0 \ldots & 0.9 & 1.0 \end{bmatrix}$$

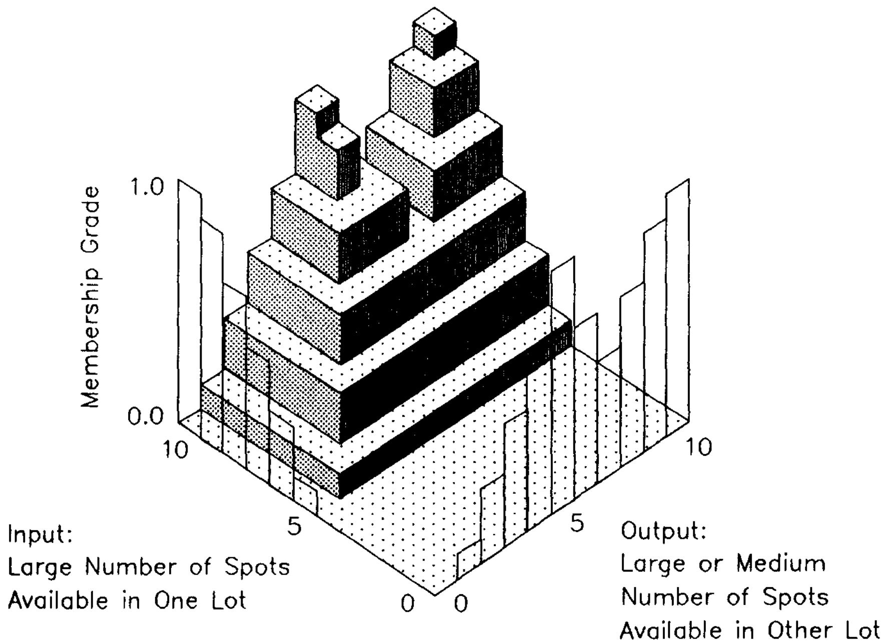


Figure A1.4 Fuzzy Relation Example

of them all.

$$R = R_1 \cup R_2 \cup \ldots \cup R_N \tag{A1.15}$$

where R_i = the fuzzy set produced by the ith rule
N = number of rules

Composition rule of inference. For a particular system, a control strategy is developed which relates a number of input sets to corresponding output sets. When the control strategy is applied, however, a given input set may not correspond exactly to any of those in the rules. The composition rule of inference is used to generate an output set in this situation. For an input set A' to a system with the rule R = A x B, the output set B' is determined as follows:

$$\mu_{B'}(o) = \max_{d}(\min[\mu_{A'}(d), \mu_R(d,o)]) \tag{A1.16}$$

Referring to the parking lot example and the relation in Figure A1.4, an input to this control strategy might be 'about 8', represented as follows:

About 8 = 0.0|0 U 0.0|1 U 0.0|2 U 0.0|3 U 0.0|4 U 0.0|5 U
0.1|6 U 0.5|7 U 1.0|8 U 0.5|9 U 0.1|10 (A1.17)

Using this input and the composition rule of inference, Figure A1.5 shows the generation of a corresponding fuzzy output. Note that the output set, also known as the control set, in this situation has membership function values less than 1. This represents the degree to which the input set differs from those in the rules and implies that knowledge of the particular situation is incomplete.

$$O = A8 \circ R = \begin{array}{l} \max (\\ \max (\\ \ldots \\ \max (\\ \max (\end{array} \begin{bmatrix} \min[0.0,\ 0.0] & \min[0.0,\ 0.0] \ldots & \min[0.5,\ 0.0] & \min[0.1,\ 0.0] \\ \min[0.0,\ 0.1] & & & \min[0.1,\ 0.1] \\ \ldots & & & \ldots \\ \min[0.0,\ 0.0] & & & \min[0.1,\ 0.9] \\ \min[0.0,\ 0.0] & \min[0.0,\ 0.0] \ldots & \min[0.5,\ 0.9] & \min[0.1,\ 1.0] \end{bmatrix} \begin{array}{l}) \\ ) \\ \ldots \\ ) \\ ) \end{array} = \begin{bmatrix} 0.0 \\ 0.1 \\ \ldots \\ 0.7 \\ 0.7 \end{bmatrix}$$

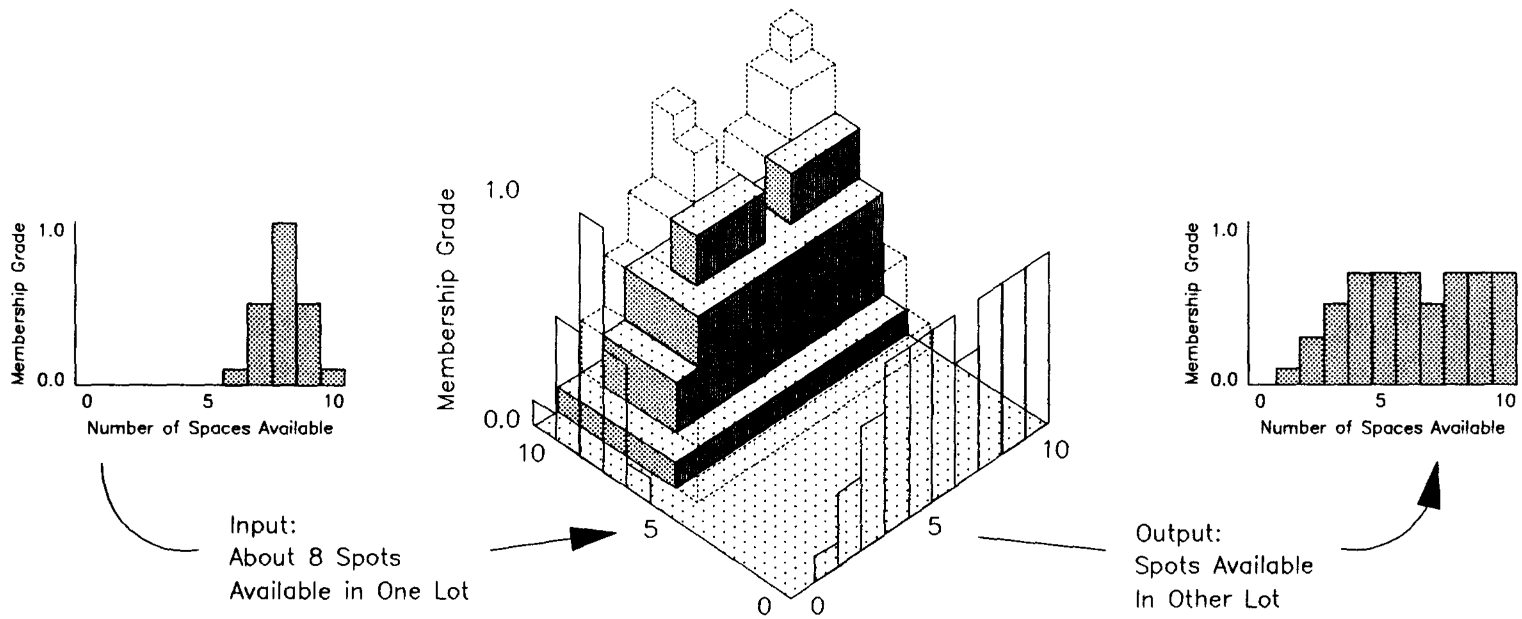


Figure A1.5 Composition Rule of Inference Example

The basic control strategy as described here formed the building block of the driver model in this thesis. In the model, the inputs to the strategies were subjective and consisted of fuzzy numbers (like the 'about 8' used in the parking lot example), but the outputs were crisp commands to drivers' vehicles. It remains, therefore, to transform the fuzzy control sets of the strategies to crisp values. In the model, this was done by taking the centroid of the fuzzy output as shown in Figure A1.6.

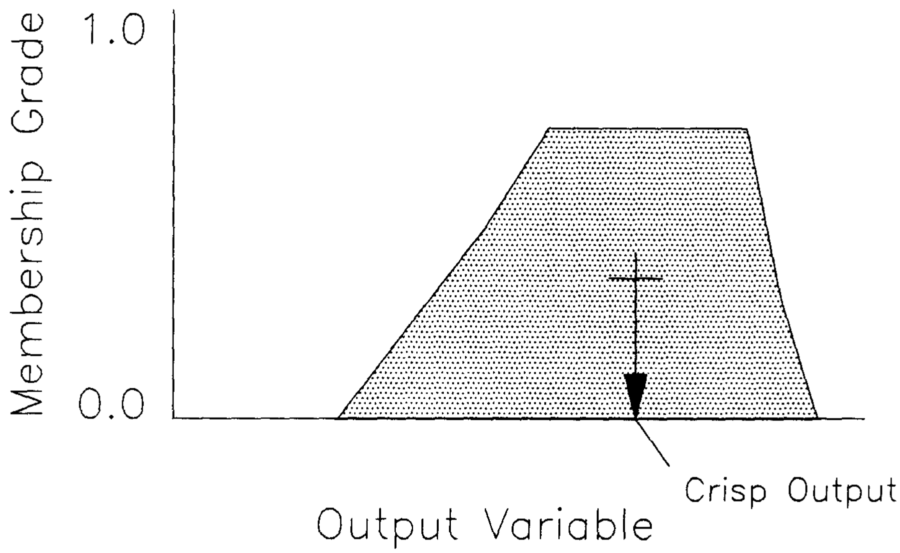


Figure A1.6 Crisp Output From Fuzzy