PERGAMON

Transportation Research Part C 10 (2002) 351–371

TRANSPORTATION RESEARCH PART C

www.elsevier.com/locate/trc

Modelling lane changing and merging in microscopic traffic simulation

Peter Hidas *

School of Civil and Environmental Engineering, University of New South Wales, Sydney, NSW 2052, Australia

Abstract

This paper introduces Simulation of Intelligent TRAnsport Systems (SITRAS), a massive multi-agent simulation system in which driver-vehicle objects are modelled as autonomous agents. The simulation outputs can be used for the evaluation of Intelligent Transport Systems applications such as congestion and incident management, public transport priority and dynamic route guidance. The model concepts and specifications, and the first applications of the model in the area of incident modelling in urban arterial networks were described in previous publications. This paper presents the details of the lane changing and merging algorithms developed for the SITRAS model. These models incorporate procedures for 'forced' and 'co-operative' lane changing which are essential for lane changing under congested (and incident-affected) traffic conditions. The paper describes the algorithms and presents simulation examples to demonstrate the effects of the implemented models. The results indicate that only the forced and cooperative lane changing models can produce realistic flow-speed relationships during congested conditions.

Keywords: Microscopic traffic simulation; Lane changing; Merging; Multi-agent systems

1. Introduction

Microscopic traffic simulation models are becoming an increasingly important tool for transport system analysis and management. They permit the traffic engineer to study and evaluate the performance of transport network systems at the operational level, under various alternative management options. Controversial or new techniques can be tried and tested without any disruption to traffic in a real network.

One area where simulation models have an important role is the analysis of traffic incidents, such as accidents or vehicle breakdowns. These incidents create a temporary reduction of capacity thus

* Tel.: +61-2-9385-5988; fax: +61-2-9385-6139.
E-mail address: p.hidas@unsw.edu.au (P. Hidas).

PII: S0968-090X(02)00026-8

leading to severe congestion in the surrounding road network. Intelligent transport systems (ITS) applications, such as congestion and incident management systems and advanced traveller information systems, are increasingly developed today as potential solutions to such congestion problems that neither the construction of new transport infrastructure, nor the application of traditional traffic management measures have been able to solve. Such systems require detailed input data about the effects of potential incidents in order to develop appropriate control strategies that can be readily implemented when an incident occurs. While incidents are very difficult to study in real life, through simulation modelling, incidents with varying severity, time of occurrence, duration and location may be investigated and the effects of various proposed control strategies may be evaluated.

Simulation of Intelligent TRAnsport Systems (SITRAS) is a microscopic time-interval update simulation model being developed at the University of New South Wales since 1995. The model is implemented in an object-oriented structure under the Microsoft Windows operating system on IBM-compatible personal computer. The main aim of the model is to provide a general evaluation tool for ITS applications such as congestion and incident management, public transport priority and dynamic route guidance. The model concepts and specifications and the first application of the model in the area of incident modelling in urban arterial networks were described in previous papers (Hidas and Behabahanizadeh, 1995, 1997; Behbahanizadeh and Hidas, 1996).

This paper concentrates on the lane changing and merging (as a special case of lane changing) algorithms developed for the SITRAS model. Lane changing is a vital component of any traffic simulation model, and the rules describing drivers' lane changing manoeuvres are well described in the literature (e.g. Gipps, 1986). These rules assume that a lane changing manoeuvre takes place without interference with vehicles in the destination lane, i.e. a vehicle changes lane without forcing other vehicles in the destination lane to slow down or stop. However, lane changing could never occur in a congested situation according to these rules. In such situations vehicles have to 'force' their way into the destination lane and field observations show that most drivers are willing to accept this behaviour under congested conditions. The interactions between drivers involved in such a manoeuvre require complex behavioural decision-making processes which can be modelled by Autonomous Agent techniques. No information dealing with 'forced' or 'co-operative' lane changing situations was found in the literature. Thus, by expanding the capabilities of previous models, a model was developed and implemented in SITRAS which is able to handle lane changing under congested (and incident-affected) traffic conditions.

This paper is structured as follows. In the next section the relevant literature on lane changing is critically reviewed and summarised. The following section provides a brief overview of the main features of the SITRAS model, to the extent necessary for the understanding of the lane changing procedures. The next sections describe the lane changing and merging algorithms and procedures, and present simulation examples to demonstrate the practical application of the model. The final section summarises the findings and conclusions of the paper and suggestions for further model development.

2. Literature review on lane changing

Gipps (1986) proposed a framework for the structure of lane changing decisions in urban driving situations including the influence of traffic signals, obstructions and different vehicle types such as heavy vehicles. The model concentrates on the decision-making process considering the potentially conflicting goals and assuming a logical driver behaviour. The decision whether or not to change lane is based on the following factors in Gipps' model:

- whether it is physically possible and safe to change lanes without an unacceptable risk of collision,
- the location of permanent obstructions,
- the presence of special purpose lanes such as transit lanes,
- the driver's intended turning movement,
- the presence of heavy vehicles, and
- the possibility of gaining a speed advantage.

The model also considers the urgency of the lane changing manoeuvre in terms of the distance of the intended turn of the driver. The urgency of the manoeuvre is modelled through the drivers' gap acceptance and braking behaviour. However, the model assumes that a lane changing manoeuvre takes place only when it is 'safe', i.e. when a gap of sufficient size is available in the target lane.

While Gipps' model provided a convenient starting point for the implementation of the lane changing algorithms in SITRAS, this assumption was found to be a serious limitation in congested and incident-affected conditions which needed further consideration. Another arguable aspect of Gipps' model is that the check of the feasibility of lane changing is performed before actually checking whether the vehicle needs to change lane and thus this check needs to be done for every vehicle during the vehicle update process. This appears to be illogical, however, from a computational efficiency point of view it is beneficial to perform the fastest check first, in order to minimise the number of vehicles for which further checks are needed. In Gipps' model the feasibility of a lane change is based on relatively simple conditions, which may justify the selected order. However, in a model where more complex procedures are applied when a lane change is necessary but unfeasible, it is better to establish first the need for a lane change before dealing with the feasibility of the manoeuvre.

In the last decade a number of new microscopic and macroscopic traffic simulation models were developed which incorporated some form of a lane changing model. Unfortunately, very little detailed information about these lane changing models is published in the literature. Most publications mention that the implemented lane changing model is based on a set of rules, but the description of the rules is usually superficial and incomplete. For example, Fritzsche (1994) describes a microscopic traffic simulation model to be used for the analysis of bottleneck situations, e.g. when one lane of a multi-lane road is temporarily closed. This is a typical situation where vehicles trapped behind the lane closure during congested flow conditions cannot move into the unblocked lane without the active cooperation of drivers in the unblocked lane. The description of the lane changing rules is very brief in the paper, and cooperative or forced lane changing behaviour is not considered.

Yousif and Hunt (1995) developed a microscopic simulation model for the investigation of lane changing behaviour on multi-lane unidirectional roadways. The rules pertaining to the desire and the possibility to change lane are based on similar logic to that described by Gipps (1986). Again, the assumption of the model is that if the available gap in the target lane is smaller than a given acceptable limit, no lane changing will take place. The main concern of the study is the relationship between lane utilisation and traffic flow on dual-carriageway roads under normal flow conditions (i.e. without incidents) and the model is adequate for this purpose. However, it could not produce realistic results when incidents or lane closures affect the flow conditions.

Wagner et al. (1997) describe a 'minimal microscopic' traffic model developed to reproduce macroscopic characteristics of traffic flow. The aim of their study is to define 'realistic' traffic rules for the modelling of lane usage on multi-lane roads. For lane changing, they define a set of rules that describe when a car wants to change lane, and a 'security constraint' rule which specifies that a car that wants to change lane is not allowed to hinder the car behind on the other lane. The model was found to be able to reproduce satisfactorily the lane usage characteristics on multi-lane roads over a wide range of flow levels under normal (incident-free) conditions.

Barceló et al. (1996) describe the AIMSUN2 microscopic traffic simulator developed for modelling real-time traffic management and information systems. The behaviour of each single vehicle on the network is continuously modelled throughout the simulation time period, according to several driver behaviour models (car following, lane changing, gap acceptance). The lane changing model is based on Gipps' model (1986). Although it is stated that AIMSUN2 can also model incidents, no information is given on how the model deals with lane changing under incident situations.

As opposed to the commonly used rule-based methods, a different approach is taken by Hunt and Lyons (1994) who developed a driver decision-making model for lane changing using neural networks. Their model works by assessing simple visual pattern-based input describing the driving environment around the vehicle about to change lane; it does not consider possible cooperation between drivers during lane changing.

One simulation model which explicitly addresses cooperative lane changing is MITSIM (Yang and Koutsopoulos, 1996), in which a courtesy yielding function is used to make space for a vehicle moving into the lane. Although details of the process are not described in the paper, the concept appears to be similar to that implemented in SITRAS.

While no information was found in the literature on forced lane changing, similar behaviour was described in a recent study of gap acceptance behaviour at roundabout entries by Troutbeck and Kako (1997). They found that 'gap-forcing' or 'priority-sharing' behaviour exists to a small extent even at low saturation levels, and at high saturation levels the ratio of forced gaps to all the merged gaps may be up to 6–12%. It seems that as the level of congestion increases drivers attempting to enter the roundabout are becoming more aggressive and/or drivers in the circulating carriageway are becoming more willing to share their right-of-way by slowing down to create an acceptable gap for the incoming vehicles. Common driving experience indicates that this priority-sharing behaviour is even more prevalent among vehicles trapped behind a lane blockage during incident situations, where the ratio of forced gaps is likely to be much higher than the ratios observed by Troutbeck and Kako (1997). The findings of this study were used to develop the forced and cooperative lane changing models in SITRAS.

2.1. The SITRAS model

The overall structure of the SITRAS model is presented in Fig. 1. The main modules of the model are: (i) route building; (ii) vehicle generation; (iii) route selection, based on individual driver characteristics; and (iv) vehicle-progression, based on car following and lane changing theory.

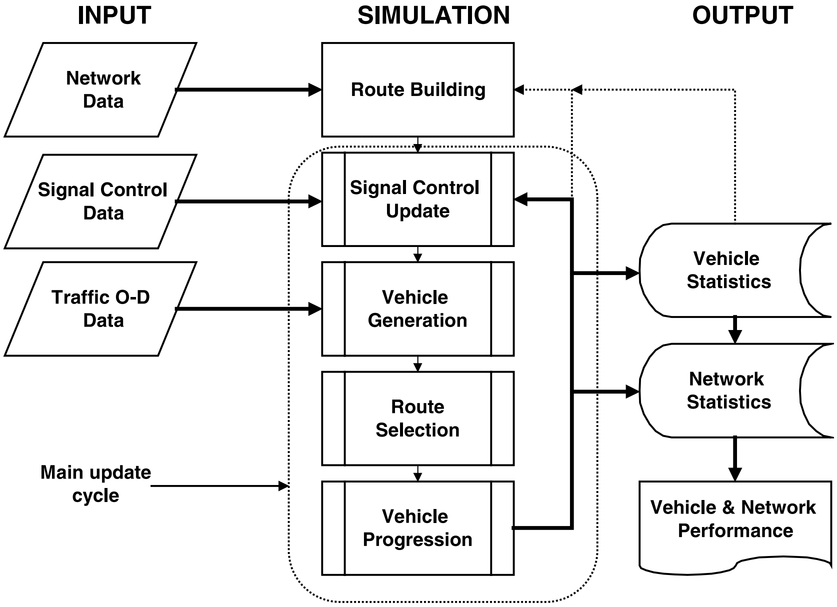

Fig. 1. The SITRAS model structure.

The driver-vehicle objects (DVOs) travel between their user-defined origin and destination, selecting their route according to the prevailing traffic conditions and their individual route choice characteristics. The model allows the simulation of two main DVO classes: (i) *Unguided* vehicle drivers have an imperfect knowledge of the prevailing network conditions. The route selection method for unguided vehicles uses Burrell's (1968) simulation method to calculate a *perceived* shortest route for each vehicle based on drivers' imperfect knowledge of network conditions. The perceived shortest route is calculated from the *fixed* minimum-path tree based on average travel costs representing the expected conditions according to the flow levels expected during the simulation period. The fixed minimum paths remain constant unless there is a significant change in the flow generation rates. This stochastic route choice method is combined in SITRAS with the drivers' familiarity with the network, which is linked to a given level of network hierarchy; (ii) *Guided* vehicles (i.e. vehicles fitted with an in-vehicle route guidance device). For guided vehicles, a deterministic route choice method is used to calculate the prevailing minimum cost route for each vehicle based on the average travel speeds of the last route update interval. As vehicles progress through the network, their average travel speed along each link is recorded by the link. The route building algorithm is called in regular intervals (currently every two minutes) to rebuild the *current* minimum paths to reflect the current flow conditions. SITRAS can then be used to evaluate the effects of dynamic route guidance systems (DRGS).

Fixed-time signal coordination and adaptive signal control strategies can be programmed into the network model. SITRAS can also model signalised intersections controlled by the SCATS system. Incidents of varying severity may be set to occur at any point and time on the network and for any duration. During the simulation run, an animation option allows the user to observe the current conditions at any part of the network. The model outputs provide various statistics of vehicle and network performance both in numerical and graphical formats.

In SITRAS, DVOs have individual characterstics, such as (i) vehicle type (passenger car, truck, bus, etc.); (ii) physical parameters related to vehicle type, such as vehicle size, maximum speed, maximum acceleration/deceleration; (iii) driver type (representing behavioural differences between less and more aggressive drivers in terms of desired speed, acceleration and gap acceptance); and (iv) drivers' level of knowledge of the network (which affects route choice behaviour).

DVOs have a set "goal" during the simulation: they want to get from their origin to their destination as quickly as possible. To achieve this goal, DVOs use their individual strategies based on their perception of network conditions and their individual characteristics. This involves a series of decisions made at periodic intervals during their journey: (i) route selection when entering a new link (i.e. road section), and (ii) calculation of acceleration (including deceleration) at each update interval time (one second) of the simulation. The speed and position of each DVO are then calculated from their selected acceleration.

The decision about acceleration is a complex process including: (i) reaching the DVOs maximum desired speed according to the prevailing network conditions; (ii) reaction to information received from the surrounding road environment (e.g. a blocked lane, a red traffic light or a STOP sign); and (iii) reaction to information received from other DVOs (e.g. car following, lane changing and merging).

Chaib-draa and Levesque (1996) proposed a framework for dealing with different types of situations (routine, familiar and unfamiliar situations) in multi-agent systems, and demonstrated its application in some scenarios of urban traffic simulation. Their model is based on a hierarchical structure defined by the skill-rule-knowledge levels of human behaviour and reasoning techniques proposed by Rasmussen (1986). Skill-based behaviour refers to fully automated activities (perception—execution) typically used in routine situations, rule-based behaviour corresponds to stereotyped actions (perception—situation recognition—planning—execution) mostly applicable in familiar situations, and knowledge-based behaviour refers to conscious activities involving problem solving and decision making (perception—situation recognition—decision making—planning execution) which are required in unfamiliar situations. This hierarchical structure is clearly recognisable in the different elements of the calculation of acceleration in SITRAS:

- reaching the DVOs maximum desired speed is a simple skill-based behaviour that does not involve interaction between several DVOs: if there is no other constraint, then accelerate to the maximum desired speed;
- reaction to information received from the road environment corresponds to rule-based behaviour: if the traffic light ahead is red then slow down to stop;
- reaction to information received from other DVOs usually requires more complex knowledge-based behaviour and decision making: for example if a DVO perceives that another DVO intends to move into its lane, it may or may not facilitate the manoeuvre depending on the conditions and the individual characteristics of the DVO.

This knowledge-based behaviour involves high level communication, coordination, negotiation and conflict resolution between the DVOs. The heuristic of setting the acceleration of the vehicle is illustrated in Fig. 2. This procedure is performed by each vehicle in each 1 s simulation update interval.

- Test if Speed is less than maximum desired speed on given link
  - if Yes: calculate Acc1required to reach maximum desired speed
- Test if End-Of-Link is close
  - if Yes: obtain Entry-Conditions from End-Node and calculate Acc2 according to the Entry Conditions
- Test if End-Of-Lane is close
  - if Yes: calculate Acc3 required to stop before end of lane
- Test if Leader Vehicle is close
  - if Yes: calculate Acc4 according to desired spacing behind leader vehicle
- Test if Vehicle provides courtesy to another vehicle that is about to move into this lane
  - if Yes: calculate Acc5 to create sufficient gap for entering vehicle
- Test if another vehicle is providing courtesy for moving into the required target lane
  - if Yes: calculate Acc6 according to desired spacing behind leader vehicle in the target lane
- Set Acceleration = Minimum of (Acc1... Acc6)

Fig. 2. Vehicle acceleration heuristic.

The above described characteristics of the driver-vehicle objects used in SITRAS correspond to those described for autonomous agents in multi-agent systems (Tokoro, 1996). Thus DVOs are, in effect, autonomous agents and SITRAS is a multi-agent system. This is despite the fact that SITRAS originally was developed without any specific knowledge of the autonomous agent concept. In fact, most microscopic simulation models—at least those developed in an object-oriented framework—would correspond to the specifications of multi-agent systems. This is an indication of how naturally suited the autonomous agent concept is for the modelling of vehicular traffic in road networks. Some differences among models can be identified in the level of 'intelligence' the DVOs possess in terms of reasoning abilities, communicating with other agents, and cooperating in solving common problems. In the last two years the author has started to explore systematically the possibilities of autonomous agent technologies in search of more efficient solutions to some weaknesses in the SITRAS model. This paper presents the first results of this on-going process.

It is also important to recognise the limitations of the DVOs as intelligent agents:

- In SITRAS, DVOs have no memory—they do not remember their past, and they can only plan ahead for the next 1 s. Thus DVOs cannot 'learn' from experience—all the knowledge they possess is procedural knowledge (Bigus and Bigus, 1998) encoded in the program. This in itself is not a problem if enough intelligence is programmed into the agents. The rules and conventions governing the movement of vehicles are well known and do not change during the simulation—they can therefore be clearly stated in the program code.
- DVOs have little direct contact with surrounding DVOs. Vehicles in the same lane of each road section are 'linked' to their immediate leader and follower vehicles, as shown in Fig. 3, to facilitate the most frequent decision making processes, but they do not 'know' about the vehicles in

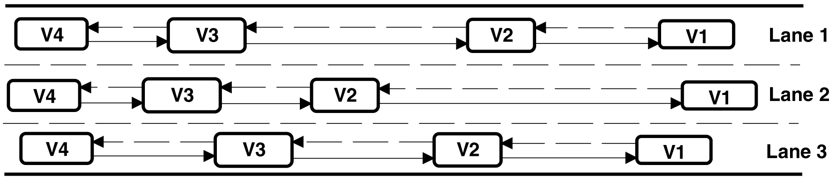

Fig. 3. Links between driver-vehicle objects in SITRAS.

adjacent lanes. A search process is used to identify vehicles in the adjacent lanes whenever this is necessary, e.g. when the vehicle needs to change lane.

These limitations are necessary in a massive multi-agent system such as SITRAS, where many thousands of agents may co-exist, in order to minimise the storage requirements of each DVO agent and maintain a reasonable running speed for the simulation. While the performance of DVOs is satisfactory in most situations, some weaknesses found in the model performance indicate that the current model is unable to reproduce realistic driver behaviour when vehicles merge into another lane. These weaknesses and the proposed improvements are described in the following sections.

2.2. The car-following model

The car-following model implemented in SITRAS is briefly summarised here because of its close relationships with the lane changing model. During the development of SITRAS it was found that none of the available car-following models was able to handle the full range of urban traffic situations without collisions. Therefore, a new car-following model was developed specifically for the microscopic simulation of urban interrupted traffic flow conditions. The model is based on the assumption that, when approaching and following a leader vehicle ($n-1$) at any time t the driver of the following vehicle (n) attempts to adjust its acceleration so as to reach a *desired spacing* after a time lag which takes T seconds. This condition can be described by the following equation:

$$x_{n-1}(t+T) - x_n(t+T) = \varepsilon D_n(t+T) \tag{1}$$

where $x_i(t)$ is the position and $D_i(t)$ is the desired spacing of vehicle i at time t and ε is the driver judgement error parameter, a random value with mean $= 1.0$ and standard deviation $= 0.11$ calibrated for typical urban flow conditions, variable in each time interval. Within the typical speed range of urban traffic (0 to 60 km/h) the desired spacing is assumed to be a linear function of the speed:

$$D_n = \alpha v_n + \beta \tag{2}$$

where v_i is the speed of vehicle i, and α and β are constants. From the condition in Eq. (1) the acceleration of the follower vehicle (a_n) can be calculated as

$$a_n = \frac{T}{\varepsilon\alpha_n T + 1/2T^2}(v_{n-1} - v_n) + \frac{1}{\varepsilon\alpha_n T + 1/2T^2}(x_{n-1} - x_n - \varepsilon\alpha_n v_n - \varepsilon\beta_n) + \frac{1/2T^2}{\varepsilon\alpha_n T + 1/2T^2}a_{n-1} \tag{3}$$

In Eq. (3) the desired spacing parameters (α and β) are used with the subscript n to emphasise that they are individual parameters of the follower vehicle n. The proposed model eliminates the problems associated with the reaction time inherent in previous models, while the use of distinct desired spacing characteristics for acceleration and deceleration reproduces the observed hysteresis loops due to the reaction delay in the acceleration phase. The model ensures a smooth transition from the unconstrained to the constrained situation as a follower vehicle gradually approaches a slower leader. The model has been validated for urban interrupted flow conditions. A full description of the car following model can be found in Hidas (1998).

2.3. The lane changing procedures in SITRAS

Lane changing is evaluated in each simulation update interval as part of the vehicle progression process. The overall structure of the lane changing model is depicted in Fig. 4. The structure shows only the main components of the process and the relationships among them. Each component is a complex process in itself. The components are described in more detail in the following sub-sections.

2.3.1. Is lane changing necessary?

A lane change may be necessary for a number of reasons. These reasons are evaluated one by one in the process in their order of importance. The result of the evaluation of a reason may be one of 'Essential', 'Desirable' and 'Unnecessary'. If lane changing is found to be 'Essential' for any reason, the rest of the reasons are not evaluated. This section of the model was developed from Gipps' (1986) decision structure with some additions. The following factors are considered.

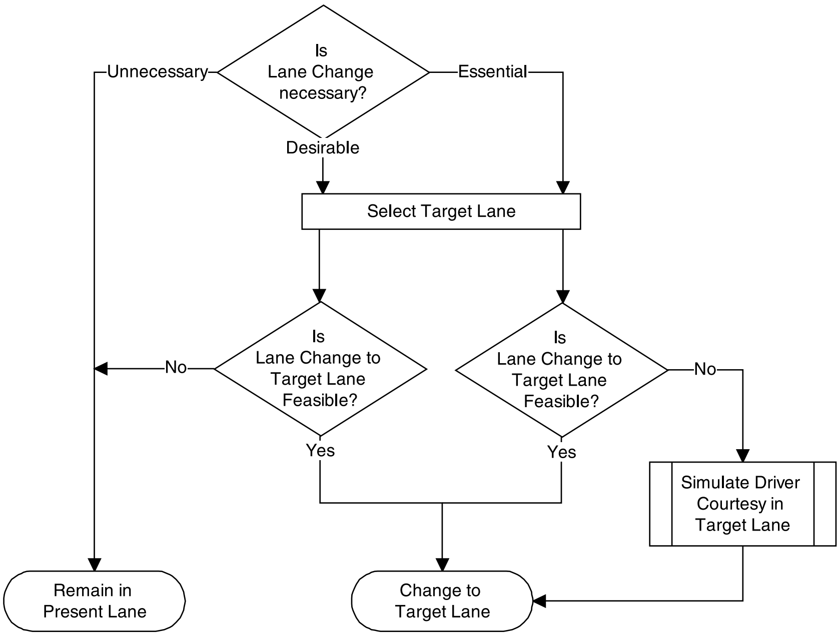

Fig. 4. Summary flowchart of the lane changing process in SITRAS.

1. *Turning movement*: If a vehicle is in a lane which allows its intended turning movement, the lane change is 'Unnecessary'. Otherwise, the necessity of the lane change depends on the distance from the turning point. Following Gipps' (1986) suggestions, the lane change is considered 'Essential' when the vehicle is less than 10 s away from the turn, it is 'Desirable' when the vehicle is between 10 and 50 s away, and 'Unnecessary' if the vehicle is even further from the turn. This is further influenced by the number of lane change manoeuvres needed to move into the appropriate lane for the required turning movement. Thus, a vehicle in lane 1 intending to move ultimately to lane 3 will start its manoeuvre earlier.
2. *End-of-lane*: If a vehicle is in a lane that ends (e.g. a merge acceleration lane), it has to move to a continuous lane. A lane change is 'Essential' when the vehicle is less than 8 s away from the end-of-lane, and it is 'Desirable' when the vehicle is further behind.
3. *Incident (lane blockage)*: If the lane ahead is blocked for any reason, such as a broken-down vehicle or a lane closure, a lane change is 'Essential' when the vehicle is less than 8 s away from the obstruction or the end-of-queue behind the obstruction, and it is 'Desirable' when the vehicle is further behind.
4. *Transit lane*: If a vehicle which is not entitled to travel in a transit lane is in a transit lane, e.g. after a recent left turn or passing over an incident in the adjacent lane, a lane change is 'Essential'.
5. *Speed advantage*: If the vehicle in front of the subject vehicle has a lower desired speed than that of the subject vehicle, and the two vehicles are relatively close (less than 200 m away) a lane change is 'Desirable', provided that the target lane allows the vehicle's intended turning movement or the vehicle is not 'close' to the turn (more than 10 s away).
6. *Queue advantage*: When approaching a stopped queue (e.g. at signalised intersections) a lane change is 'Desirable' if the queue in the target lane is at least 10 m shorter than the one in the current lane, provided that the target lane allows the vehicle's intended turning movement.

2.3.2. Select target lane

The target lane is the one which the vehicle intends to move into. It is the lane immediately to the left or to the right of the current lane. The selection of the target lane is combined with the evaluation of the necessity of lane changing. For some of the reasons, such as intended turning movement and transit lane, the direction of the lane change is determined by the reason, and only one direction can be considered, while for other reasons, such as incidents, speed or queue advantage, the lane on either side may be considered, depending on the feasibility of the conditions in those lanes. Thus, the target lane may be either the kerb-side or median-side lane of the current lane, or both lanes. In the latter case, the final target lane will be selected depending on the feasibility conditions in the next step of the process.

2.3.3. Is lane change to target lane feasible?

In the context of our model, a lane change is considered feasible if there is a gap of sufficient size in the target lane so that the vehicle can move into the target lane safely, without forcing other vehicles in the target lane to slow down significantly. To determine the feasibility of the manoeuvre, first a search, in the target lane, is conducted to find the vehicle just in front, and the vehicle just behind the subject vehicle. These would become the new leader and the new follower respectively of the subject vehicle should the lane change be feasible. The lane change is feasible if

(a) the deceleration (or acceleration) required for the subject vehicle to move behind the new leader vehicle is acceptable, and
(b) the deceleration required for the new follower vehicle to allow the subject vehicle to move into the lane is acceptable.

For both conditions the deceleration of the subject vehicle and the potential new follower vehicle are calculated using the car following model implemented in SITRAS. This acceleration is then compared with an 'acceptable deceleration', which depends on the urgency of the manoeuvre and it is calculated using a modified format of the formula suggested by Gipps (1986):

$$b_n = [2 - (D - x_n(t))/(10V_n)]b_{\mathrm{LC}}\theta \tag{4}$$

where b_n is the acceptable deceleration of vehicle n; D is the location of the intended turn or lane blockage; $x_n(t)$ is the location of vehicle n at time t; V_n is the desired (free) speed of vehicle n; b_{LC} is the average deceleration a vehicle is willing to accept in lane changing; θ is the 'driver aggressivity parameter'.

Thus, the closer the vehicle is to its intended turn or to the lane blockage, the larger the acceptable deceleration. The average deceleration a vehicle is willing to accept in lane changing, b_{LC}, is taken as half of the maximum deceleration of the vehicle (for passenger cars -2.1 m/s^2). The 'driver aggressivity parameter' is included to represent individual differences among drivers. In SITRAS each driver-vehicle object is assigned a driver type parameter, a number between 0 and 99. Larger numbers represent more aggressive drivers. The driver type is drawn from a normal distribution when the driver-vehicle object is created. The driver type is used as a modifying factor of the speed characteristics (desired free flow speed, maximum acceleration/deceleration) of the vehicle. In the formula of the acceptable deceleration the driver aggressivity parameter is calculated differently for the two conditions:

- For condition (a) it is the ratio of the subject vehicle driver type to the 'average' driver type (i.e. driver type = 50). Thus, the leader vehicle has no effect on the acceptable acceleration, it depends only on the aggressivity parameter of the subject driver.
- For condition (b) it is the ratio of the subject vehicle driver type to the new follower vehicle driver type. Thus, the more aggressive the subject vehicle driver compared to the new follower vehicle driver, the higher the acceptable deceleration.

If both the kerb-side and the median-side lanes are to be considered as the target lane, the feasibility conditions are first checked in the median-side lane, then, if it is found to be unfeasible, the kerb-side lane is checked.

If both conditions are satisfied in the target lane, the process continues with the execution of the lane changing. If the manoeuvre is not feasible and the manoeuvre is not essential, the lane changing process is abandoned. If the lane changing is essential, the process continues with the simulation of a forced lane changing manoeuvre.

2.3.4. Simulate driver courtesy in target lane

The forced lane changing algorithm developed in SITRAS is based on a 'driver courtesy' concept: the vehicle which wants to change lane sends a 'courtesy' request to subsequent vehicles

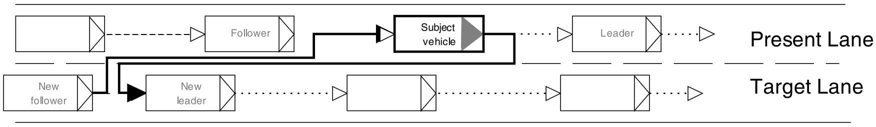

Fig. 5. Schematic representation of the forced lane changing concept.

in the target lane; the request is evaluated by each vehicle and depending on several factors such as the speed, position and driver type of the responding vehicle, it is either refused or accepted. When a vehicle 'provides courtesy' to another vehicle it reduces its acceleration to ensure that a free gap of sufficient length is created during the next few seconds for the lane changing vehicle.

The concept is illustrated in Fig. 5. Starting from the first vehicle in the target lane which is located behind the subject vehicle, the deceleration required for each potential new follower vehicle to allow the subject vehicle to move into the target lane is calculated from the car-following model. [1] This deceleration is compared with the 'acceptable' deceleration calculated by Eq. (4), with the only difference that the driver aggressivity parameter is taken as the ratio of the 'average' driver type (i.e. driver type = 50) to the driver type of the potential new follower vehicle; that is, it depends only on the aggressivity of the potential follower vehicle, and that the more aggressive the driver of the potential follower vehicle, the less deceleration it is prepared to accept. If the deceleration required for the vehicle is less than the maximum acceptable, the vehicle is selected to provide 'courtesy' to the subject vehicle, otherwise the evaluation continues with the next vehicle in the platoon.

Once the new follower vehicle is found in the target lane, its acceleration is calculated by the car following model with respect to the subject vehicle in the adjacent lane instead of its current leader in the same lane. At the same time, the acceleration of the subject vehicle is calculated with respect to its new leader vehicle in the target lane instead of its current leader in the same lane. Consequently, the new follower vehicle will gradually slow down, while the new leader vehicle will pass the subject vehicle and a gap of sufficient size will be created which will finally allow the subject vehicle to move into the target lane. This process may take several seconds (i.e. several vehicle update intervals), during which time there is a direct cooperation between these vehicles.

This cooperation between several vehicles in separate lanes presents difficult challenges for the implementation of the algorithm. In SITRAS the driver-vehicle-units (objects) are stored in their current lane in 'chains', i.e. they only 'know' about their immediate leader and follower (see Fig. 3). The vehicle update process is executed lane by lane and vehicle by vehicle, starting from the first vehicle in the lane. Therefore, under normal conditions, vehicles in separate lanes are not aware of each other, even when they are side by side. It would be possible to include in the vehicle objects pointers to the vehicles in the adjacent lanes, however, in a transport network simulation model where several thousand vehicles may be simulated at any one time, this would significantly and unnecessarily increase the size of the vehicle objects. In SITRAS a separate object type (called LaneChange object) was defined which holds the 3 vehicles participating in a forced lane changing

[1] The driver courtesy function did not consider the possibility of accelerating to merge in front of a vehicle ahead in the target lane. This was found to be one of the weaknesses of the lane changing model and the option was incorporated in the improved model which will be discussed later.

manoeuvre and which enables direct communication between the vehicles. The LaneChange object is created when a vehicle requires courtesy for its essential lane change and it is stored in a separate list in the network object until the manoeuvre is finally performed, then the object is destroyed. This provides a computationally more efficient solution.

2.3.5. Change to target lane

Although SITRAS is a microscopic simulation model, the aim of the model is to predict overall network performance under various demand and control conditions based on the cumulative evaluation of individual vehicle performance, such as total travel time, average speed, etc. This level of analysis does not require a realistic simulation of the lane changing manoeuvre itself. Therefore, once the lane changing manoeuvre is found to be feasible, it is executed as an instantaneous change in SITRAS. This includes changing the current lane parameter of the subject vehicle and resetting the leader-follower relationships according to the new situation both in the previous lane and the target lane.

2.3.6. Implementation details

The procedures presented above may appear complete in theory, but when first implemented in SITRAS, produced many situations where vehicles became stranded in both lanes or behaved unrealistically. A number of special cases had to be resolved to ensure the smooth and error-free operation of the model. These special cases include the following:

- If a vehicle just executed a lane change then it should not be considered for another lane change in the same update interval. This problem arises from the fact that the lane change update process is executed in each link lane-by-lane starting from the left-most lane. Therefore, if a vehicle moved from the left lane to the middle lane but its final aim is to move further to the right, it would jump two lanes without this additional condition.
- If a lane change to the target lane was found to be unfeasible but the lane change is essential, there is a need to check if the target lane is blocked: if it is not blocked the model will attempt to provide courtesy for the vehicle, while if the lane is blocked, the courtesy function is not activated.
- Sometimes it may happen that two vehicles travelling side-by-side try to move into each other's lane at the same time. In a case like that the courtesy function must ensure that they will not try to provide courtesy for each other (and thus become stranded).
- If a vehicle is already providing courtesy to a vehicle it will not be selected to provide courtesy for another vehicle.

2.4. Example

Microscopic simulation models are not amenable to calibration and validation at the microscopic level, due to the random fluctuations in individual behavioural parameters. Instead, the objective of calibration and validation is to demonstrate that the overall/average performance of the model is in acceptable agreement with observed data and thus it is suitable for the intended purpose. Validation of a complex traffic simulation model is a complex process in which components of the model can be validated separately. The validation of the car-following model implemented in SITRAS has been described in a previous paper (Hidas, 1998). The overall model performance under various freeway incident conditions was tested in another experiment (Behbahanizadeh and Hidas, 1998). The simulation example presented here was devised to demonstrate the effects of forced lane changing on the overall model performance.

A 500 m section of a three-lane one-way urban arterial road was simulated for one hour (after 5 min of warm-up time) in SITRAS with a constant demand flow of 4000 veh/h containing only passenger cars. The simulation was first run without incidents, then with an incident blocking one lane close to the end of the link, and finally with an incident blocking two lanes at the same place. A snapshot of the simulation is depicted in Fig. 6. Each experiment was repeated twice, once with the driver courtesy function enabled and once the driver courtesy function disabled.

Flow and travel time output results from each simulation run are presented in Table 1. It can be seen from Table 1 that the driver courtesy function has an effect on the flow, although this effect is relatively minor, in the order of 5–10%. It is notable that while the flow decreases when the courtesy function is disabled in the no incident and one-lane incident cases, there is an increase in the flow in the two-lane incident case. This is likely due to the fact that in the two-lane incident

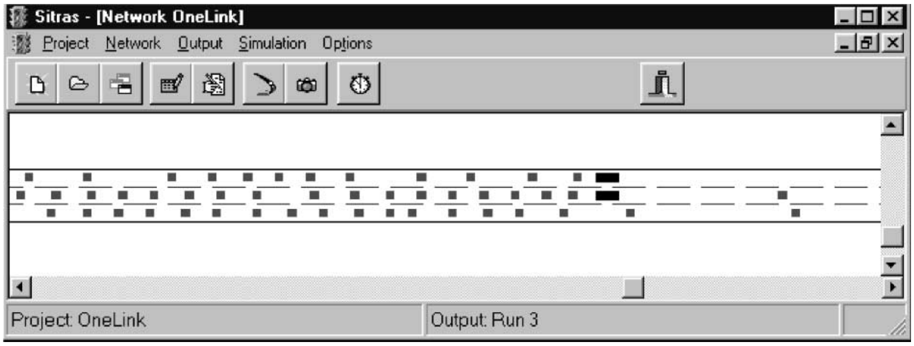

Fig. 6. A simulated road link in SITRAS with an incident blocking two lanes.

Table 1
Overall results from the simulation experiments

Simulation case	Vehicle count	Travel time (s)				
		Min	Max	Mean	SD	*t*-test
No incident						
With courtesy	4115	25	30	27	1.01	2.183
Without courtesy	3864	25	30	27	1.04	
One-lane incident						
With courtesy	1682	70	623	215	135	−7.961
Without courtesy	1610	85	1057	264	205	
Two-lane incident						
With courtesy	1308	66	744	258	141	−4.640
Without courtesy	1465	95	1152	292	240	

case a larger proportion of the vehicles have to slow down to allow vehicles from the blocked lanes to shift into the unblocked lane and this further reduces the capacity of the bottleneck compared to the without courtesy case. As expected, the courtesy function has a far more significant effect on the travel times during incident situations. While the mean travel times increase 15–25% when the courtesy function is disabled during incident situations, the maximum travel times go up by 70% at the same time. The standard deviations also show an increasing difference between the with and

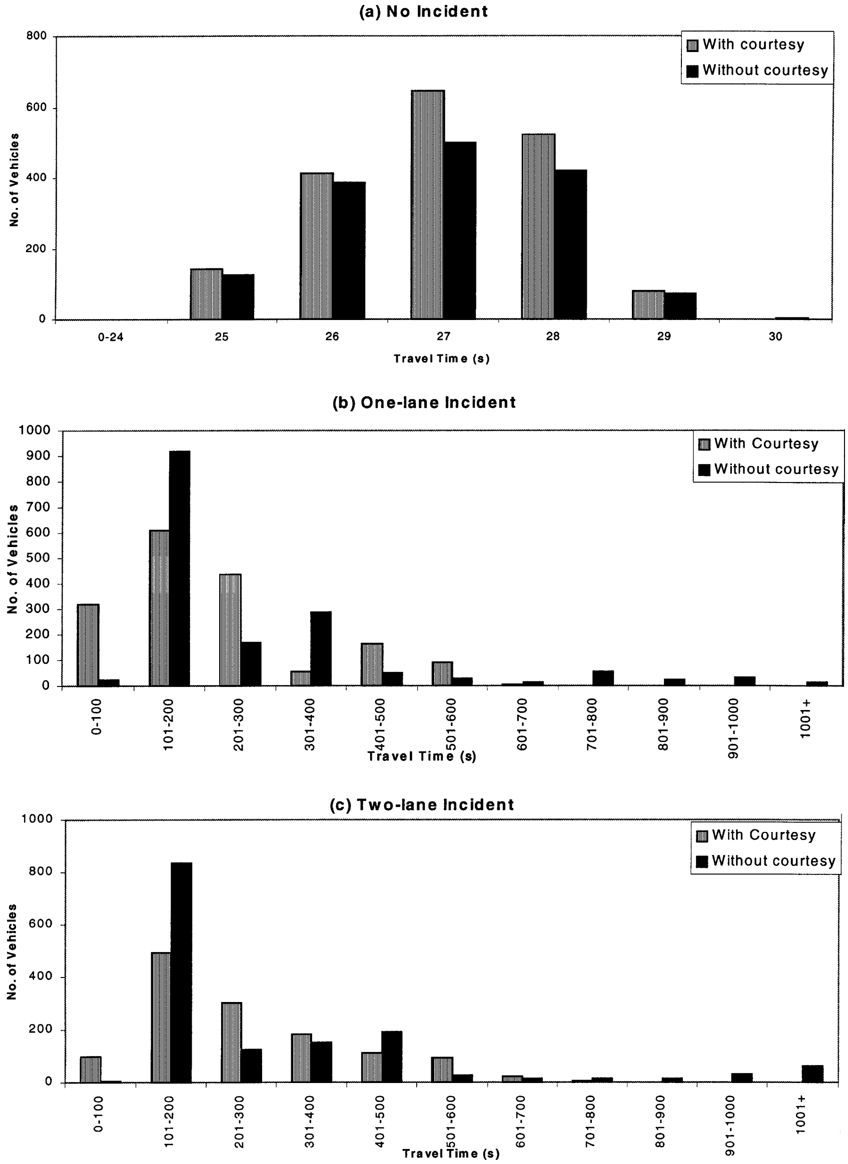

Fig. 7. Distribution of travel times.

without courtesy cases during incidents. The student *t*-test indicates that the difference between the travel time distributions of the with and without courtesy cases is statistically significant at the 0.05 level in all three simulation scenarios.

The effect of the driver courtesy function can best be demonstrated through the distribution of travel times. Fig. 7 shows the distribution of travel times for each simulation case. It can be seen that in the no incident case (Fig. 7a), the two distributions corresponding to the with and without courtesy cases are very similar and they both have a shape similar to a normal distribution. In the one-lane and two-lane incident cases (Fig. 7b and c), when the courtesy function is enabled the distribution still has a normal-like shape although it is significantly skewed to the right. However, when the courtesy function is disabled the distribution appears to be like a combination of two populations: one including the majority of the vehicles which were able to shift into the unblocked lane without the cooperation of other vehicles, and the other including those vehicles which remained stranded behind the lane blockage for a long time. Field observations during traffic incidents show that the queue behind the lane blockage is moving at about the same speed as the one in the unblocked lane, therefore the travel time distribution seems clearly unrealistic without the courtesy function.

3. Merging problems

Merging is a special case of lane changing, where vehicles must move into the target lane because the lane they are currently travelling in, ends or blocked. A typical example is a freeway on-ramp acceleration lane. It is at such merge points where shortcomings of the first lane changing model were found: the absorption capacity of the merge lane was somewhat below the expected value, and the mean vehicle speed in the merge section was lower than observed on typical freeway merge sections (see Fig. 9). These shortcomings are caused by the following features of the current model:

- In the first lane changing model, only the gap directly beside the subject vehicle (the one that has to change lane) is checked for the feasibility of the lane change. No attempt is made to investigate other gaps and to devise a strategy to enter the most suitable gap.
- The decision to change lane is based on a gap acceptance criterion: after the lane change the available gaps between the new leader and follower vehicles must be longer than a certain 'critical gap'. In the current model the critical gap was assumed to be equal to the desired spacing used in the car following model. However, field observations show that many lane changing occur at much shorter gaps.
- If a vehicle cannot change lane right away, it tends to slow down and even stop. Consequently, the speed difference compared with the mean vehicle speed in the target lane increases which, in turn, makes lane changing even more difficult. While it may be necessary to slow down and stop before reaching the end of the lane, field observations show that vehicles on the merge lane tend to accelerate up to the mean speed in the target lane and only slow down and stop very rarely as a last resort.
- The 'driver courtesy' algorithm was developed with congested urban arterial road conditions in mind: when a DVO in the target lane decides to provide courtesy to the subject vehicle, it slows

down, or even stops, to allow the subject vehicle to move into the target lane. While this is a realistic (though not very frequent) behaviour on congested urban arterials, it is practically non-existent on motorways under high speed conditions.

- In real life, when approaching a merge point on the main carriageway, vehicles often move into the middle lane to make way for merging vehicles. This behaviour was not considered in the first lane changing model in SITRAS.

4. The improved model

In short, the above problems reveal the lack of intelligence of the DVO agents, in terms of perception and recognition of the traffic situation, decision-making, conflict resolution and co-operation between agents. A new concept has been developed to handle merging in a more intelligent and realistic manner. It includes the following components:

- Gap acceptance is evaluated against a shorter critical gap criterion based on the relative speed and acceleration of the subject vehicle with respect to its potential leader/follower in the target lane. Field observations show that when the relative speed is close to zero, drivers accept very short gaps, sometimes no more than 1–2 m.
- If a vehicle cannot change lane in the current time interval, it evaluates the flow conditions in the target lane, and by predicting the position and speed of adjacent vehicles in the target lane up to a few seconds ahead, it attempts to set an acceleration which may lead to a more favourable situation for lane changing. The heuristic of the getMergeAcceleration process is summarised in Fig. 8.
- A new procedure has been implemented to model realistic driver behaviour for the vehicles travelling in the main freeway lanes: as vehicles in the kerb lane approach the merge point, they check if there is a vehicle about to merge into their lane. If there is such a vehicle close-by, the vehicles in the kerb lane attempt to move into the middle or median lane so as to make way for the merging vehicle.

- examine each subsequent time interval t (t = 1, 2, 3)
  - calculate minimum/maximum position of the subject vehicle based on current speed and minimum/maximum acceleration
  - identify the first and subsequent vehicles in the target lane within the minimum/maximum position of the subject vehicle
  - examine each subsequent gap in the target lane:
    - if the gap is longer than the critical gap:
      - calculate the position where the subject vehicle should be for merging
      - calculate the acceleration needed for the subject vehicle to reach the merge position
      - stop if a suitable MergeAcceleration is found, otherwise continue
    - if the gap is shorter, take the next gap until the last one

Fig. 8. Merge-acceleration heuristic.

- If a vehicle changes lane with very short gaps (i.e. shorter than the desired spacing at the current speed) it creates a disturbance in the platoon of vehicles in the target lane. In such cases the car following model used in SITRAS would produce a large deceleration to redress the potentially dangerous situation. As a consequence, the mean speed in the target lane decreases significantly. However, field observations show that drivers are willing to accept the risk associated with such short gaps, and instead of braking hard, they tend to just decelerate very slightly, in order to gradually increase the gap and regain the normal desired spacing corresponding to the speed. This risk-taking behaviour is more prevalent on motorways under uninterrupted flow conditions, where drivers have a clear view far ahead and hence they can anticipate that the vehicles in front will not break suddenly in the next few seconds. The new algorithm allows vehicles to recognise such situations and to set a lower acceleration accordingly. This 'risk-acceleration' ensures that the gap between subsequent vehicles will gradually increase and reach the desired spacing within 5 s.
- The 'driver courtesy' function is not used at merge points under normal (uncongested) flow conditions. If a vehicle is unable to change lane before reaching the end of the lane, it has to stop and wait until it can merge without significantly slowing down any vehicle in the target lane. However this situation normally should not occur; if it does, it is a sign of weakness of the simulation model that needs to be corrected in the other components of the model.

The new concept has been implemented and tested with a wide range of flow levels. The model has to satisfy both macroscopic and microscopic traffic flow criteria:

- the simulated saturation flow must be reasonably close to the expected capacity of motorway merge sections known from the traffic engineering literature (TRB, 1994);
- the mean speed in the merge lanes must remain close to the uninterrupted flow speed expected on motorway sections;

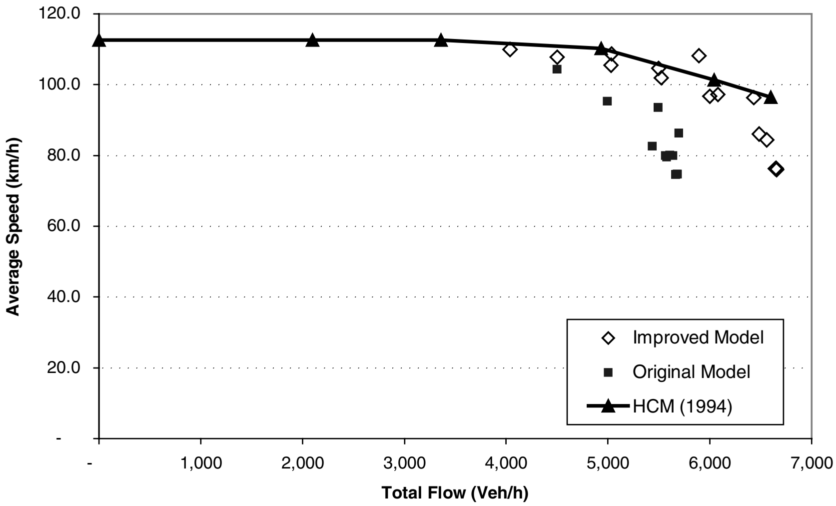

Fig. 9. Expected vs modelled speed/flow relationships at a merge section of a three-lane freeway.

- lane changing must take place without collisions between vehicles;
- merging vehicles should not remain stuck at the end of the acceleration lane.

The test runs made so far with the new model show a definitive improvement in terms of speed-flow relationships. Fig. 9 shows that the results from the improved model closely match the expected curve based on the US Highway Capacity Manual (TRB, 1994). The gap acceptance function was successfully refined until all collisions were eliminated. However, under congested conditions sometimes vehicles still remain stuck at the end of the acceleration lane. Work on further refinements and validation of the model is in progress.

5. Summary and conclusions

This paper has introduced SITRAS, a microscopic traffic network simulation model and presented the details of the lane changing models including both normal (unforced) and forced or co-operative lane changing. Preliminary results show that incident situations cannot be simulated realistically without the forced lane changing procedures.

Analysis of the model performance revealed some weaknesses at freeway merge sections caused by the limited perception and recognition of traffic situations, and over-simplified decision-making of the DVOs. A new, more intelligent concept has been developed for lane changing and merging behaviour. Testing and validation of the new model is on-going but first test results clearly show the superiority of the new model.

This experiment demonstrates that the autonomous agent concept is naturally suited to the microscopic modelling of road traffic. Because such a model includes a very large number of agents (up to many thousands), it is necessary to limit the 'intelligence' of individual agents in terms of their direct communication links with surrounding vehicles and in their learning abilities. It is argued that driver-vehicle agents (DVAs) do not need learning abilities in microscopic traffic simulation, because traffic rules are well defined and within the duration of microscopic traffic simulation studies (typically less than 24 h) the rules governing driver behaviour should also be stable, therefore they can be encoded in the program as procedural knowledge. However, the micromechanisms of these behavioural rules are not well known and learning and adaptation-cooperation techniques of autonomous agents may be used to develop these rules and 'train' the DVAs as part of the model calibration process. Although DVAs do not have a memory in SITRAS, the decision-making processes can include anticipation of adjacent agents actions for several seconds ahead; this allows DVAs to develop a more successful and longer term strategy for merging as illustrated in Fig. 8. In order for the model to be useful for the intended purpose, it must satisfy both macroscopic and microscopic traffic engineering criteria at the same time; autonomous agent techniques proved to be helpful in achieving this goal.

SITRAS is currently being further developed as a test-bed for the evaluation of incident management schemes in urban arterial networks. Part of this work is the calibration and validation of the model using data on traffic conditions and driver behaviour under incident situations collected in Sydney, Australia. A separate project funded by the University of New South Wales is under way to explore further possibilities offered by intelligent agent technologies in microscopic modelling of road traffic systems.

Acknowledgements

Parts of the model development and the simulation experiments were prepared by Kamran Behbahanizadeh. The author would like to acknowledge the positive role of the New South Wales Roads and Traffic Authority (RTA) in the current work on SITRAS. Recent developments of the model have greatly benefited from the support of the RTA through the grant provided for the incident modelling project. The author also thanks Ralf Schleiffer for his initiative and encouragement as editor of this special issue and the anonymous referees for their constructive comments.

References

Barceló, J., Casas, J., Codina, E., Fernández, A., Ferrer, J.L., García, D., Grau, R., 1996. PETRI: A parallel environment for a real-time traffic management and information system. In: Proceedings of the 3rd World Congress on ITS (published on CD-ROM), Orlando, Florida.

Behbahanizadeh, K., Hidas, P., 1996. Modelling traffic incidents in urban arterial networks. In: Proceedings of the 3rd World Congress on ITS (published on CD-ROM), Orlando, Florida, USA.

Behbahanizadeh, K., Hidas, P., 1998. Quantifying the Impacts of Incidents In Urban Freeways, presented at the 8th World Conference on Transport Research, July 1998, Antwerp, Belgium.

Bigus, J.P., Bigus, J., 1998. Constructing Intelligent Agents with Java—A programmer's Guide to Smarter Applications. Wiley Computer Publishing, New York.

Burrell, J.E., 1968. Multipath route assignment and its application to capacity restraint. In: Proceedings of the Fourth International Symposium on the Theory of Traffic Flow, Karlsruhe, Germany, pp. 210–219.

Chaib-draa, B., Levesque, P., 1996. Hierarchical model and communication by signs, signals and symbols in multi-agent environments, distributed software agents and applications. In: Perram, J.W., Müller, J.-P. (Eds.), Proceedings of the 6th European Workshop on Modelling Autonomous Agents in a Multi-Agent World, MAAMAW'94, Springer, pp. 151–165.

Fritzsche, H.-T., 1994. A model for traffic simulation. Traffic Eng. Control 35 (5), 317–321.

Gipps, P.G., 1986. A Model for the Structure of Lane-Changing Decisions. Transport. Res. 20B, 107–120.

Hidas, P., 1998. A car following model for urban traffic simulation. Traffic Eng. Control 39 (5), 300–309.

Hidas, P., Behabahanizadeh, K., 1995. Urban transport network simulation: an object-oriented approach. In: Wyatt, R., Hossein, H. (Eds.), Proceedings of the 4th International Conference on Computers in Urban Planning and Urban Management, vol. 2, Melbourne, Australia, pp. 229–243.

Hidas, P., Behbahanizadeh, K., 1997. Route guidance as an incident management tool in urban arterial networks. In: Proceedings of the Third International Conference of ITS Australia (published on CD-ROM), Brisbane, Queensland, Australia.

Hunt, J.G., Lyons, G.D., 1994. Modelling of dual-carriageway lane-changing using neural networks. Transport. Res. Part C 2 (4), 231–245, Elsevier.

Rasmussen, J., 1986. Information processing and Human-Machine Interaction: An Approach to Cognitive Engineering. North Holland.

Tokoro, M., 1996. Agents: towards a society in which humans and computers cohabitate, distributed software agents and applications. In: Perram, J.W., Müller, J.-P. (Eds.), Proceedings of the 6th European Workshop on Modelling Autonomous Agents in a Multi-Agent World, MAAMAW'94, Springer, pp. 1–10.

Transportation Research Board 1994. Highway Capacity Manual: Special Report 209. Third Edition, National Research Council Washington, DC.

Troutbeck, R.J., Kako, S., 1997. Limited priority merge at unsignalised intersections. In: Kyte, M. (Ed.), Proceedings of the Third International Symposium on Intersections Without Traffic Signals University of Idaho, Moscow, Idaho, USA, pp. 294–302.

Wagner, P., Nagel, K., Wolf, D.E., 1997. Realistic multi-lane traffic rules for cellular automata. Physica A 234, 687–698.

Yang, Q., Koutsopoulos, H.N., 1996. A Microscopic Traffic Simulator for Evaluation of Dynamic Traffic Management Systems. Transport. Res. Part C 4 (3), 113–129, Elsevier.

Yousif, S., Hunt, J., 1995. Modelling lane utilisation on British dual-carriageway roads: effects on lane-changing. Traffic Eng. Control 36 (12), 680–687.