Effects of Co-Evolution in a Complex Traffic Network

Ana L. C. Bazzan and Denise de Oliveira
Instituto de Informática
UFRGS
Caixa Postal 15064
91.501-970 Porto Alegre, RS, Brazil
{bazzan,edenise}@inf.ufrgs.br

Franziska Klügl
Dep. of Artificial Intelligence
University of Würzburg
Am Hubland
97074 Würzburg, Germany
kluegl@informatik.uni-wuerzburg.de

Kai Nagel
Inst. for Land and Sea Transport Systems
TU Berlin Sek SG12
Salzufer 17–19
10587 Berlin, Germany
nagel@vsp.tu-berlin.de

ABSTRACT

One way to cope with the increasing demand in transportation networks is to integrate standard solutions with more intelligent measures. This paper discusses the effects of integrating co-evolving decision-making regarding route choices (by drivers) and control measures (by traffic lights) We use microscopic modeling and simulation, in opposition to the classical network analysis. General questions here are whether co-evolution pays off, and, if so, what kind of evolutionary approach shall be used. This is challenging for networks other than the two-route one due to the complexity of route-choice behavior, as well as control strategies by the traffic lights. Moreover, the more agents, the less effective learning strategies are, especially when the integration among them depicts complex interelationships. The approach was tested in different scenarios using centralized and decentralized methods.

1. INTRODUCTION

Urban mobility is one of the key topics in modern societies. Especially in medium to big cities, the urban space has to be adapted to cope with the increasing needs of the commuters. In transportation engineering the expression of the transport needs is called *demand*. This demand (in terms volume of vehicles, pedestrians, freight, etc.) is commonly used to quantify transport *supply*. This is the expression of the capacity of transportation infrastructures and modes. Supply is expressed in terms of infrastructures (capacity), services (frequency), and networks. The increasing demand of transport needs we observe nowadays has to be accommodated either with increasing supply (e.g. road capacity), or with a better use of the existing infrastructure. Since an expansion of the capacity is not always socially or economically attainable or feasible, transportation and traffic engineering now seek to optimize the management of both the supply and the demand using concepts and techniques from intelligent transportation systems (ITS). These refer to the application of modern technologies to the operation and control of transportation systems [11].

From the side of supply, several measures have been adopted in the last years, such as congestion charging in urban areas (London), restriction of traffic in the historical center (Rome, Paris, Amsterdam), alternace of vehicles allowed to circulate in a given day (São Paulo, Mexico City).

From the point of view of the demand, several attempts exist not only to divert trips both spatially as well as temporally, but also to distribute the demand within the available infrastructure. In this context, it is now commonly recognized that the human actor has to be brought into the loop. With the amount of information that we have nowadays, it is almost impossible to disregard the influence of real-time information systems over the decision-making process of the individuals.

Hence, within the project "Large Scale Agent-based Traffic Simulation for Predicting Traffic Conditions", our long term goal is to tackle a complex problem like traffic from the point of view of information science. This project is the result of an accumulated experience with microscopic models modeling tools for traffic and transportation management. These range from traffic signal optimization [1], binary route choice, and effect of information on commuters [4], to microscopic modeling of physical movement [7].

An important milestone in the project is to propose a methodology to integrate complex behavioral models of human travelers reacting to traffic patterns and control measures of these traffic patterns, focusing on distributed and decentralized methods. Classically, this is done via network analysis. There, it is assumed that individual road users seek to optimize their individual costs regarding the trips they make by selecting the "best" route. This is the basis of the well known traffic network analysis based on Wardrop's equilibrium principle [15]. This method predicts a long term *average* state of the network. However, assuming steady state network supply and demand conditions from day-to-day, this equilibrium based method cannot, in most cases, cope with the dynamics of the modern transportation systems. Moreover, it is definitely not adequate for answering questions related to what happens in the network *within* a given day, as the variability in the demand and the available capacity of the network tend to be high. Just think about changing weather conditions from day-to-day and within a single day!

In summary, as equilibria based concepts overlook this variability, it seems obvious that it is not adequate to be used in microscopic modeling and simulation. Therefore, the general aim of this paper is to investigate what happens when different actors adapt, each having its own goal. The objective of *local* traffic control is obviously to find a control scheme that minimizes queues in a spatially limited area (e.g. around a traffic light). The objective of drivers is normally to minimize their travel time – at least in commuting situations. Finally, from the point of view of the whole system, the goal is to assure reasonable travel times for *all* user, which can be highly conflicting with some individual utilities (a social dilemma). This is a well-known issue. Tumer and Wolpert [13] for instance shown that there is no general approach to deal with this complex question of collectives.

Specifically, this paper investigates which strategy is better for drivers (e.g. adaptation or greedy actions). Also, what is better for traffic lights? Act greedily or just carry on a "well-designed" signal plan? After which volume of traffic does decentralized control of traffic lights starts to pay off? Does isolated, single-agent reinforcement learning make sense in traffic scenarios? What happens when drivers adapt concurrently? These are hot topics not only in traffic research, but also from a more general multiagent point of view as it refers to co-evolution.

The challenge of the present paper is to tackle more realistic scenarios, i.e. depart from binary route choice. To the best of our knowledge, the question on what happens when drivers and traffic lights adapt in a complex route scenario (e.g. a grid) has not been tackled so far.

In the next section we review these and other related issues. In Section 3 we describe the approach and the scenario. Section 4 discusses the results, while Section 5 presents the concluding remarks.

2. BACKGROUND: SUPPLY AND DEMAND IN TRAFFIC ENGINEERING

Learning and adaptation is an important issue in multiagent systems. Here we concentrate on pieces of related work which either deal with traffic scenarios directly or report close scenarios.

2.1 Management of Traffic Demand

Given its complexity, the area of traffic simulation and control has been tackled by many branches of applied and pure sciences. Therefore, several tools exist which target the problem isolately. Simulation tools in particular are quite old (1970s) and stable. On the side of demand forecasting, the arguably most used computational method is the so-called 4-step-process [10]. It consists of: trip generation, destination choice, mode choice, and route assignment. Route assignment includes route choice and a very basic traffic flow simulation, often, but not always, leading to a Nash Equilibrium. Over the years, the 4-step-process has been improved in many ways, most notably by (i) combining the first three steps into a single, traveler-oriented framework (*activity-based demand generation (ABDG)*) and by (ii) replacing traditional route assignment by so-called *dynamic traffic assignment (DTA)*, where the traffic flow simulation is much more realistic. Still, in the typical implementations, all traveler information gets lost in the connection between ABDG and DTA, making realistic agent-based modeling at the DTA-level difficult. An important distinction exists between day-to-day replanning and within-day (on-the-fly) replanning. Only the latter allows simulated travelers to react to ITS measures, although some level of ITS functionality can be successfully emulated with day-to-day replanning only.

Another related problem is the estimation of the state of the whole traffic network from partial sensor data. Although many schemes exist for incident detection, there are few deployments of large scale traffic state estimation. One exception is `www.autobahn.nrw.de`. It uses a traffic microsimulation to extrapolate between sensor locations, and intelligent methods combining the current state with historical data in order to make short-term predictions. However, the particles (vehicles) are very simple: They do not know their destinations, let alone the remainder of their daily plan. This was a necessary simplification to make the approach work, but it is necessary to overcome this simplification since the effects of the travelers' decisions are difficult if not impossible to estimate without these aspects.

What is missing it a true agent-based integration of these and other approaches. However, until now agent-based simulations with high-level agents on the scale required for traffic simulation of real-world networks have not been developed.

2.2 Real-Time Optimization of Traffic Lights

Signalized intersections are controlled by signal-timing plans (we use signal plan for short) which are implemented at traffic lights. A signal plan is a unique set of timing parameters comprising the cycle length L (the length of time for the complete sequence of the phase changes), and the split (the division of the cycle length among the various movements or phases). The criterion for obtaining the optimum signal timing *at a single intersection* is that it should lead to the minimum overall delay at the intersection. Several plans are normally required for an intersection to deal with changes in traffic volume, or, in an traffic-responsive system, that at least one plan exist and can be changed on the fly.

In [1], a MAS based approach is described in which each traffic light is modeled as an agent, each having a set of pre-defined signal plans to coordinate with neighbors. Different signal plans can be selected in order to coordinate in a given traffic direction. This approach uses techniques of evolutionary game theory. However, payoff matrices (or at least the utilities and preferences of the agents) are required, i.e. these figures have to be explicitly formalized by the designer of the system.

In [9] groups were considered and a technique from distributed constraint optimization was used, namely cooperative mediation. However, this mediation was not decentralized: group mediators communicate their decisions to the mediated agents in their groups and these agents just carry out the tasks.

Camponogara and Kraus [2] have studied a simple scenario with only two intersections, using stochastic game-theory and reinforcement learning. Their results with this approach were better than a *best-effort* (greedy), a random policy, and also better than Q-learning.

Also, in [8] a set of techniques were tried in order to improve the learning ability of the agents in a simple scenario.

Finally, a reservation-based system [3] is also reported but it is only slightly related here because it does not include conventional traffic lights.

2.3 The Need for Integration

Up to now, few attempts exist to integrate supply and demand in a single model. We review three of them here.

2.3.1 Learning based approach

The paper by [16] describes the use of reinforcement learning by the traffic light controllers (agents) in order to minimize the overall waiting time of vehicles in a small grid. Agents learn a value function which estimates the expected waiting times of single vehicles given different settings of traffic lights. One interesting issue tackled in this research is that a kind of co-learning is considered: the value functions are learned not only by the traffic lights, but also by the vehicles which can thus compute policies to select optimal routes to the respective destinations. The ideas and some of the results presented in that paper are important. However, strong assumptions turn difficult its use in the real world. First, the kind of communication and knowledge (or, more appropriate, communication *for* knowledge formation) has a high cost. Traffic light controllers are supposed to know vehicles destination in order to compute expected waiting times for each. Given the current technology, this is a strong assumption. Second, it seems that traffic lights can shift from red to green and opposite at each time step of the simulation. Third, there is no account of experience made by the drivers based on their local experiences only. What about if they just react to the (few) past experiences after the route is completed and the driver takes it again the next day (typical commuting scenario)? Finally, drivers being autonomous, it is not obvious to expect that all will use the best policy computed by the traffic light and not by themselves. Thus, in the present paper, we depart from the assumptions regarding communication and knowledge the actors have to have about each other.

2.3.2 Game theoretic approach

In [14] a two-level, three-player game is discussed which integrates traffic control and traffic assignment, i.e. both the control of traffic lights and the route choices by drivers are considered. Complete information is assumed, which means that all players (including the population of drivers) have to be aware of the movements of others. Although the paper reports interesting conclusions regarding e.g. the utility of cooperation among the players, this is probably valid only in that simple scenario. Besides, the assumptions that drivers always follow their shortest routes are difficult to justify in a real-world application.

In the present paper, we want to depart from both the two-route scenario and the assumption that traffic management centers are in charge of the control of traffic lights. Rather, we follow a trend of decentralization, in which each traffic light is able to sense its environment and react accordingly and autonomously, without having its actions computed by a central manager as it is the case in [14]. Moreover, the two-route scenario is a very didactic one and serves the purpose of the main aim of [14]. However, it is questionable whether the same mechanism can be used in more complex scenarios, as claimed. The reason for this is the fact that when the network is composed of tens of links, the number of routes increases and so the complexity of the route choice, given that now it is not trivial to compute the network and user equilibria.

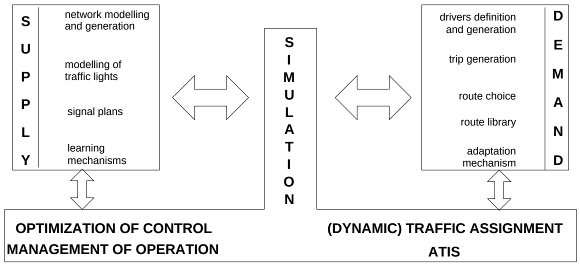

Figure 1: Schema of the Co-Evolution in an ITS Framework

2.3.3 Methodologies

Liu and colleagues [6] describe a modeling approach which integrates microsimulation of individual trip-makers' decisions and individual vehicle movements across the network. Moreover, in [6] the focus is on the description of the methodology which incorporate both demand and supply dynamics, so that the applications are only briefly described and not many options for the operation and control of traffic lights are reported. One scenario described deals with a simple network with four possible routes and two control policies. One can roughly be described as greedy, while the other is fixed signal plan based. In the present paper we do not explore the methodological issues as in [6] but, rather, investigate in more details particular issues of the integration and interaction between actors from the supply and demand sides.

3. CO-EVOLUTION IN AN ITS FRAMEWORK

Figure 1 shows an scheme of the our approach, based on the interaction among supply, demand, and an ITS module. The latter is related to strategic decisions and is composed of a simulation sub-module, as well as sub-modules to implement optimization of control (e.g. traffic lights), management of operation, traffic assignment (static or dynamic), and an information system (ATIS).

We have developed a framework using the agent-based simulation environment SeSAm [5] for testing the effects of adaption of different elements of the supply and demand. The testbed consists of sub-modules for modeling and generation of the network and the agents, modeling of traffic lights and signal plans, learning and adaptation mechanisms, generation of the library of routes, the route choice algorithm, etc. The movement of vehicles is queue-based.

In the present paper we tackle the feedback loop between

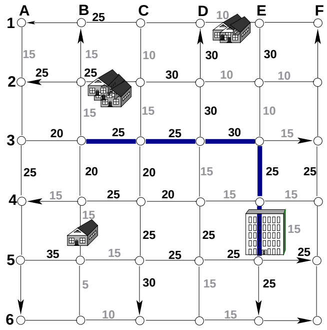

Figure 2: Grid 6x6 showing the main destination (E4E5), the three main origins (B5B4, E1D1, C2B2), and the "main street" (darker line).

supply and demand by means of studying scenarios in which drivers adapt (as in [4]) and traffic lights use learning mechanisms (e.g. in [1, 8]).

The scenario we use to exemplify the approach is a typical commuting scenario where drivers repeatedly select a route to go from an origin to a destination. It is not so simple as a two-route (binary decision) scenario. Rather, it captures desirable properties of real scenarios regarding the aim of this study, namely the co-evolution among drivers and traffic lights.

To model the supply, we use a grid with 36 nodes connected using one-way links, depicted in Figure 2 (values depicted at each link will be discussed later). All links are one-way and drivers can turn in each crossing. This kind of scenario is realistic and, in fact, from the point of view of route choice and equilibrium computation, it is also a very complex one as the number of possible routes is high.

Contrarily to simple two-route scenarios, in the grid it is possible to set arbitrary origins and destinations. Each driver has one particular origin (O) and destination (D) with a quite high number of optional route between them. To render the scenario more realistic, neither the distribution of O-D combinations, nor the capacity of links is homogeneous. On average, 60% of the road users have the same destination, namely the link labelled as E4E5. Other links have 1.7% probability of being a destination. Origins are nearly equally distributed in the grid, with three exceptions: links B5B4, E1D1, and C2B2 have, approximately, probabilities 3, 4, and 5% of being an origin respectively. The remaining links have each a probability of 1.5%. Regarding capacity, all links can hold up to 15 vehicles, except those located in the so called "main street" which can hold up to 45. This main street is formed by the links B3 to E3, E4, and E5.

The control is done via decentralized traffic lights. These are located in each node. Each has a signal plan which, by default, divides the cycle time 50-50% between the 2 phases. The actions of the traffic lights are to run the default plan or to priorize one phase. The strategies are: i) always keep the default signal plan; ii) greedy (run green time for the phase with the higher occupancy); iii) use single agent Q-learning.

Regarding the demand, the main actor is the driver. The simulation can generate any number of them, each knowing a given number of routes. Normally the simulations were done with drivers knowing one to five routes. These were generated via an algorithm that computes the shortest path (one route) and the shortest path via arbitrary detours (the other four). Drivers can use three strategies to select a route (before it departs): i) randomly; ii) greedily (always pick the one with best average travel time so far); iii) probabilistic way meaning that the average travel times so far are used to compute a probability to select the route to use.

4. RESULTS AND DISCUSSION

4.1 Metrics and Parameters

In order to evaluate the experiments, travel time (drivers) and occupation (links) were measured. Due to lack of space we discuss here only the mean travel time over the last 5 trips (henceforward *attl5t*) or travel time in a single trip. All experiments were repeated 20 times. The parameters used were: time out for the simulation (t_{out}) equal to 300 when the number of drivers is 400 or 500, 400 when there are 600 drivers, and 500 when there are 700 drivers; percentage of drivers who adapt is either 100 or zero (in this case all act greedily) but any combination can be used; percentage of traffic lights which act greedily is either zero or 100; a link is considered jammed if its occupancy is over 50%; cycle length for signal plans is 40 secs. For the Q-learning there is an experimentation phase of $10 \times t_{out}$, the learning rate is $\alpha = 0.1$ and the discount rate is $\lambda = 0.9$.

4.2 Global Optimization

Before we present the results of the simulations in which we use the co-evolution approach, i.e. agents learn in a decentralized way, we briefly discuss the results of simulations performed *without* this approach, for the sake of calibration and comparison. To this aim we use a centralized and heuristic optimization method in order to compute the optimal split of the cycle time between two traffic directions, in each intersection. This was done using the DAVINCI (Developing Agent-based simulations Via INtelligent CalIbration) Calibration Toolkit for SeSAm, which is a general purpose calibration and optimization tool. Although DAVINCI provides several global black box search strategies, like evolutionary strategies, simulated annealing or gradient based search, we have used the former only (specifically, GA were used).

DAVINCI allows specific values of given parameters of the simulation to be tested for optimality (the fitness). In our specific case, we have used it to find the best split time for each traffic light in the scenario depicted in Figure 2. The optimization objective is to minimize the average travel time over all drivers in a scenario with 400 drivers, where all drivers have only one route (the shortest path). Several combinations of splits are tested, in each intersection. At each time DAVINCI evolves a population of strings based on their fitness. For a cycle length of 40 s., several values for the split were tried in each intersection: 5/35, 10/30, 15/25, 20/20, ..., 35/5. One can see that with 36 intersections the combination of trials is huge. Thus the heuristic method must be used. At the end of the optimization, we are interested in knowing which was the best split for each intersection.

The resulting optimized splits, with a cycle time set to 40 s., can be seen in Figure 2: numbers depicted close to the respective links indicate how much green time the link receives. The optimization was made using GA with 100 generations, population size of 20 individuals and the convergence criterion set to 0.01. Using these optimized splits, the average travel time of drivers is 105. This value can

Type of Simulation	Average Travel Time Last 5 Trips
greedy drvs / fixed TLs	100
adapting drvs / fixed TLs	149
greedy drvs / greedy TLs	106
adapting drvs / greedy TLs	143
greedy drvs / Qlearning TLs	233

Table 1: Average Travel Time Last 5 Trips for 400 Drivers, under Different Conditions

be used to assess the utility of adapting drivers and traffic lights in a decentralized way.

4.3 Drivers and Traffic-Lights Learning in a Decentralized Way

In this section we discuss the simulations and results collected when drivers and traffic lights co-evolve using different strategies (explained next). As a measure of performance, we use the $attl5t$ defined previously. These are summarized in Table 1. For all scenarios described in this subsection, 400 drivers were used.

Greedy or adaptive drivers; fixed traffic lights. In the case of adapting drivers, the $attl5t$ is 149 time units, while this is 100 if drivers act greedily. The higher travel time is the price paid by the experimentation which the drivers continue to do, even though the optimal policy was achieved long before (remember that the $attl5t$ is computed only over the last 5 trips). The greedy action is of course much better after the optimal policy was learned. Notice that this travel time is slightly better than the one found by the heuristic optimization tool described before, which was 105. In summary, greedy actions by the drivers work because they tend to select the routes with the shortest path and this normally distributes drivers more evenly than longer routes.

Greedy or adaptive drivers; greedy traffic lights. When traffic lights also act greedily we can see that this does not automatically improve the outcome (in comparison with the case in which traffic lights are fixed): the $attl5t$ is 106. This happens because the degree of freedom of traffic lights' actions is low, as actions are highly constrained. For instance, acting greedily can be highly sub-optimal when for instance traffic light A serves direction D_1 (thus keeping D_2 with red light), when the downstream flow of D_1 is already jammed. In this case, the light might indeed provide green for vehicles on D_1 but these cannot move due to the downstream jam. Worse, jam may appear on the previously unjammed D_2 too due to the small share of green time. This explains why acting greedily at traffic lights is not necessarily a good policy. The travel time of 106, when compared to the travel time found by the centralized optimization tool (105), is of course similar. This is not surprising because the decentralized strategy does exactly the same as the centralized optimizer, namely drivers use their best route and traffic lights optimize greedily.

Q-learning traffic lights. We have expected Q-learning to perform bad because it is already known that it does not have a good performance in noisy and non-stationary traffic scenarios [12]. In order to test this, we have implemented a Q-learning mechanism in the traffic lights. Available actions are: to open the phase serving either one direction (e.g. D_1), or the other (D_2). The states are the combination of states in both approaching links, i.e. $\{D_1\_jammed, D_1\_not\_jammed\} \times \{D_2\_jammed, D_2\_not\_jammed\}$. The low performance of Q-learning in traffic scenarios is due basically to the fact that the environment is non-stationary, not due to the poor discretization of states. Convergence is never achieved and so traffic lights keep experimenting.

Average Travel Time Last 5 Trips				
Type of Simulation	Nb. of Drivers			
	400	500	600	700
greedy drvs / fix TLs	100	136	227	411
greedy drvs / greedy TLs	106	139	215	380

Table 2: Average Travel Time Last 5 Trips for Different Number of Drivers, under Different Conditions

4.4 Scenarios With More Drivers

For more than 400 drives we only investigate the cases of greedy drivers / fixed traffic lights versus the case in which both drivers and traffic lights act greedily in order to test whether or not increasing volume of traffic (due to increasing number of drivers in the network) would cause greedy traffic lights to perform better. This is expected to be the case since once the number of drivers increases, greedy actions alone do not bring much gain; some kind of control in the traffic lights is expect to be helpful. In fact, 400, 500, 600 and 700 drivers mean an average occupancy of $\approx$ 40%, 47%, 59%, and 72% per link respectively. In Table 2 the $attl5t$ for these numbers of drivers are shown. The case for 400 drivers was discussed above. With more than 500 drivers, the $attl5t$ is lower when traffic lights also act greedily. In the case of 700 drivers, the improvement in travel time (411 vs. 380) is about 8%. Thus, the greedy traffic lights are successful in keeping the occupancy of links lower resulting in reduced travel times.

4.5 Overall Discussion

In the experiments presented we could see that different strategies by the drivers, as well by the traffic lights have distinct results in different settings. We give here the main conclusions. For the network depicted, increasing the links capacity from 15 to 20 would lead to much less jam (this was tested but is not shown here due lack of space). However, increasing network capacity is not always possible so that other measures must be taken. Diverting people and/or given them information both have limited performances. Thus the idea is to better use the control infrastructure. Therefore we have explored the capability of the traffic lights to cope with the increasing demand.

Regarding travel time, it was shown that the strategies implemented in the traffic lights pay off in several cases, especially when the demand increases. We have also measured the number of drivers who arrive before time t_{out}. Just to give a flavor of the figures, bad performance (around 75% arrived) was seen only when the drivers adapt probabilistically. This of course is a consequence of the high travel times (see Table 1). The general trend is that when the traffic lights also adapt or learn, the performance increases, by all metrics used.

Regarding the use of Q-learning, as said, single-agent learning (i.e. each agent isolately using Q-learning) is far from optimum here due to the non-stationarity nature of the scenario. This is true especially for those links located close to the main destination and the main street as they tend to be part of each driver's trip so that the pattern of volume of vehicles changes dramatically. A possible solution is to use collaborative traffic lights. In this case, traffic light A would at least ask/sense traffic light B downstream whether or not it shall act greedily. This however leads to a cascade of dependence among the traffic lights. In the worst case everybody has to consider everybody's state. Even if this is done in a centralized way (which is far from desirable), the number of state-action pairs prevents the use of multiagent Q-learning in its standard formulation.

5. CONCLUSION

Several studies and approaches exist for modeling travelers' decision-making. In commuting scenarios in particular, probabilistic adaptation in order to maximize private utilities is one of those approaches. However, there is hardly any attempt to study what happens when *both* the driver and the traffic light use some evolutionary mechanism in the same scenario or environment, especially if *no central control exist*. Then, the co-evolution happens in a decentralized fashion, in which case some forms of auto-organization or chaotic situations may arise. This is an important issue because, although ITS have reached a high technical standard, the reaction of drivers to these systems is fairly unknown. In the present paper we have investigated this loop. The results show an improvement regarding travel time and occupancy (thus, both the demand and supply side) when all actors co-evolve, especially in large-scale situations e.g. involving hundreds of drivers. This was compared with situations in which either only drivers or only traffic lights evolve, in different scenarios, and with a centralized optimization method.

This work can be extended in two main directions. First, we plan to integrate the tools developed by the authors indepently for supply and demand which are simulators with far more user-friendly capabilities and permit the modeling of even more realistic scenarios as trips can be richer etc. The results are not expect to differ in the general trends, though. The second extension relates to the use of heuristics for a MAS reinforcement learning in order to improve its performance. This is not a trivial extension as it is known that reinforcement learning for non-stationary environments is a hard problem, especially when several agents are involved.

Acknowledgments

The authors would like to thank CAPES (Brazil) and DAAD (Germany) for their support to the joint, bilateral project "Large Scale Agent-based Traffic Simulation for Predicting Traffic Conditions". Ana Bazzan is partially supported by CNPq and Alexander von Humboldt Stiftung; Denise de Oliveira is supported by CAPES.

6. REFERENCES

[1] Bazzan, A. L. C. A distributed approach for coordination of traffic signal agents. *Autonomous Agents and Multiagent Systems 10*, 1 (March 2005), 131–164.

[2] Camponogara, E., and Kraus Jr., W. Distributed learning agents in urban traffic control. In *EPIA* (2003), F. Moura-Pires and S. Abreu, Eds., pp. 324–335.

[3] Dresner, K., and Stone, P. Multiagent traffic management: A reservation-based intersection control mechanism. In *The Third International Joint Conference on Autonomous Agents and Multiagent Systems* (July 2004), pp. 530–537.

[4] Klügl, F., and Bazzan, A. L. C. Route decision behaviour in a commuting scenario. *Journal of Artificial Societies and Social Simulation 7*, 1 (2004).

[5] Klügl, F., Herrler, R., and Oechslein, C. From simulated to real environments: How to use SeSAm for software development. In *Proceedings of the 1st German Conferences MATES – Multiagent System Technologies* (2003), S. Berlin, Ed., no. 2831 in Lecture Notes in Artificial Intelligence, pp. 13–24.

[6] Liu, R., Van Vliet, D., and Watling, D. Microsimulation models incorporating both demand and supply dynamics. *Transportation Research Part A: Policy and Practice 40*, 2 (February 2006), 125–150.

[7] Nagel, K., and Schreckenberg, M. A cellular automaton model for freeway traffic. *Journal de Physique I 2* (1992), 2221.

[8] Nunes, L., and Oliveira, E. C. Learning from multiple sources. In *Proceedings of the 3rd International Joint Conference on Autonomous Agents and Multi Agent Systems, AAMAS* (New York, USA, July 2004), N. Jennings, C. Sierra, L. Sonenberg, and M. Tambe, Eds., vol. 3, New York, IEEE Computer Society, pp. 1106–1113.

[9] Oliveira, D., Bazzan, A. L. C., and Lesser, V. Using cooperative mediation to coordinate traffic lights: a case study. In *Proceedings of the 4th International Joint Conference on Autonomous Agents and Multi Agent Systems (AAMAS)* (July 2005), New York, IEEE Computer Society, pp. 463–470.

[10] Ortúzar, J., and Willumsen, L. G. *Modelling Transport*, 3rd ed. John Wiley & Sons, 2001.

[11] Roess, R. P., Prassas, E. S., and McShane, W. R. *Traffic Engineering*. Prentice Hall, 2004.

[12] Silva, B. C. d., Basso, E. W., Bazzan, A. L. C., and Engel, P. M. Dealing with non-stationary environments using context detection. In *Proceedings of the 23rd International Conference on Machine Learning (ICML 2006)* (June 2006), W. W. Cohen and A. Moore, Eds., ACM Press, pp. 217–224.

[13] Tumer, K., and Wolpert, D. A survey of collectives. In *Collectives and the Design of Complex Systems*, K. Tumer and D. Wolpert, Eds. Springer, 2004, pp. 1–42.

[14] van Zuylen, H. J., and Taale, H. Urban networks with ring roads: a two-level, three player game. In *Proc. of the 83rd Annual Meeting of the Transportation Research Board* (January 2004), TRB.

[15] Wardrop, J. G. Some theoretical aspects of road traffic research. In *Proceedings of the Institute of Civil Engineers* (1952), vol. 2, pp. 325–378.

[16] Wiering, M. Multi-agent reinforcement learning for traffic light control. In *Proceedings of the Seventeenth International Conference on Machine Learning (ICML 2000)* (2000), pp. 1151–1158.