Simulation of Adaptive Agents: Learning Heuristics for Route Choice in a Commuter Scenario

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

Ana L. C. Bazzan
Institute of Informatics, UFRGS
Caixa Postal 15064
91501-970 Porto Alegre, RS, Brazil
bazzan@inf.ufrgs.br

ABSTRACT
An important prerequisite for traffic management is to find efficient ways to model and predict traffic flow. Here we are presenting a naïve model for the route choice adaptation of learning commuters with heuristics based behaviour. Our simulation results show that the heuristics learnt lead to a situation similar to that obtained in real experiments.

Categories and Subject Descriptors
I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence – *Coherence and Coordination, Multiagent Systems*

General Terms
Economics

Keywords
Adaptation and Learning, Traffic Simulation

1. INTRODUCTION
Adequate modelling and prediction of traffic flow is an interesting research task. Nowadays it becomes more and more important, as Advanced Travel Information Systems (ATIS), for instance dynamic route guidance systems, are deployed. Also systems that provide traffic forecast are planned. However, drivers decisions in reaction to these information may alter the traffic situation and potentially make the predictions of the ATIS obsolete. Thus a traffic forecast system has to incorporate drivers reactions, drivers decision making. However, it is not at all clear, how and under consideration of which information drivers select their routes.

Additionally, traffic system can be taken as a good example for flexible and emergent organisations. The inter-dependence of actions leads to a high frequency of implicit co-ordination decisions. The more reliable the information that a driver gets about the current and future state of the traffic network, the more his actions — e. g. his route choices — depend on what he beliefs about the decisions of the other road users. Thus studies about route decisions in a commuters scenario are interesting from a more general point of view.

AAMAS '02, July 15-19, 2002, Bologna, Italy.

2. MODELING DECISION MAKING
There are several areas that made contributions to modelling drivers decision making:

- Microscopic simulations derived from microeconomic theories already involve more travel alternatives, joint and dynamic decision-making, contingency planning under uncertainty (e.g. due to congestion), and an increasing frequency of co-ordination decisions. However, when choices are complex, utility maximisation seems no longer to be a tenable assumption. In a way similar to the utility maximisation theory, behavioural decision theory states that a good decision is a choice of actions that meet the decision-makers' objectives.
- Experimental economics deals with the acquisition of data related to scenarios, which are economically relevant. Thus, there is a close connection with the theory of decision as well as with game theory. Typical for such experiments is that one can observe the influence of available information. If the experiments are repeated by the same subjects (which is the standard practice), these are able to learn new patterns, which will in their turn, influence the further experiments.
- In evolutionary game theory the prerequisites of rationality of the participants are weakened: The anticipation of a solution by the players frequently leads to the assumption that they have observed past interactions; they do not need to know explicitly how their actions influence those of their opponents, since they may asymptotically converge to a steady state represented by a set of evolutionary stable strategies (ESS). For that they just have to know their own payoffs for applying a suitable learning rule [1].

3. MODEL FOR COMMUTING
The following model corresponds to the concrete scenarios set in the first round of experiments in the SURVIVE project. The scenario consists of two route. The agents have to select either the main route M or the side route S. If a significant number of commuters use the normally faster route M, the route side S might be faster. On the other hand, many drivers may think the same way and opt to select the side road. Their decisions depend on their beliefs about the environment and the behaviour of the other drivers. This may be seen as a game with incomplete information, since the basic information (what other participants are deciding), is not known. The game is iterated a certain number of rounds and behavioural tendencies evolve in the course of time. After deciding which route to take, the environment calculates the payoff of all agents according to the following formula:.

$$reward\ (i) = 40 - \begin{cases} 6 + 2 * number_M\text{, if } i \text{ has choosen } M \\ 12 + 3 * number_S\text{, if } i \text{ has choosen } S \end{cases}$$

The drivers do not know this reward function. The parameters $number_M$ and $number_S$ represent the number of commuters in the main and secondary road, respectively. We set n=18 and the total number of rounds to 200, because this was the scenario that was used as an experimental set-up in the SURVIVE-Project[5]. Nevertheless we tested also configurations with n=900 agents producing the same results.

The most simple decision model for agents is based on a bias for choosing a certain route, called here "route choice heuristic". Practically, it is the probability according to which a driver selects the main route. With a certain not synchronised periodicity a driver updates this heuristic according to the rewards he has obtained on the routes he took until now. The update of the heuristic is done in a way similar to the one suggested by Harley in [2], namely according to the following formula:

$$heuristic\,(i) = \frac{\sum_i reward_M\,(i,t)}{\sum_i reward_M\,(i,t) + \sum_i reward_S\,(i,t)}$$

$Reward_M$ is reward the agent has gained so far on the main route *M*, while the $reward_S$ denotes his success on the side route *S*. The more a driver selects a route, the more feedback information he gets from this route. Therefore an important factor is *how often and in which intervals the heuristic is updated*. This is especially relevant because the reward depends on the other agents. When the agent is learning his individual heuristic, he is also implicitly adapting himself to the others. We considered different extensions to this very simple model (for further discussions see [3]).

The complete scenario was implemented using a development tool for multi-agent simulations called SeSAm [4]. We performed experiments varying especially such mean frequency of heuristic adaptation and observing the organisation of overall route choice.

4. RESULTS AND DISCUSSION

The values for the interval between two adaptation steps were set to: 0 (i.e. adaptation happens every round), or a mean of 5, 10, 20 and 50 (i.e. every n±n/2 rounds). Agents start with a heuristic value equal to 0.5 (i.e. equal probability to select both routes). Appropriate starting values were assigned when adaptation happens every round. We repeated every simulation run six times like in the real experiments.

4.1 Emergence of Route Stability

According to the reward-function, the *ideal final value* of heuristics should be 0,667. In none of our simulation experiments this value was learnt by drivers individually. However, with adaptation interval 0, 5, and 10, on average, the complete system learns this value as a mean, whereas the individual drivers "specialise" on selecting one route again and again. Using adaptation rates of 20 and 50, the experiment just seem to be too short for such slowly converging learning schemes. Figure 1 shows the distribution in a final situation for mean learning frequency equal to 5. Using this configuration the driver specialised most.

It is obvious that the better the complete system has learnt the equilibrium, the higher the overall and individual sum of reward is. The highest averaged sum of rewards can be found in the

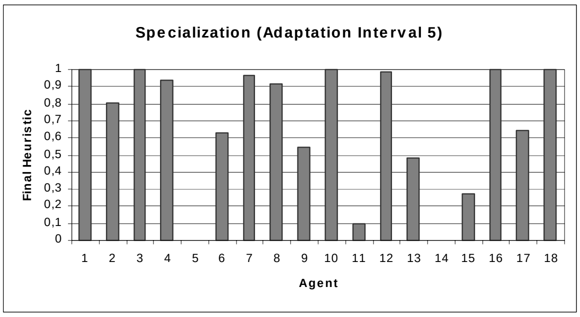

Figure 1: Final heuristics of all agents in an example experiment with learning about every 5 rounds.

experiments in which the inertia is small, i.e. equal to 5. There also the standard deviation of is lower than in the other configurations.

Although the scenario is simple, it is possible to compare the results of this simulation to real experimental data from the SURVIVE project. After a first glance one can say that in both, real experiments and our simulation an analogous form of specialisation or route stability is observable. However, it is too early to state general results regarding this validation. In future work we want to pursue the comparison between simulated and experimental data. Depending on the results we gain from this validation, we will extend the form of adaptation to consider more information, especially social one.

5. CONCLUSIONS

Our long-term work focuses on commuting scenarios in which drivers have to make decisions, which on their turn alter the traffic condition [6]. This paper continues on this track, by reverting to more basic simulations with agents that not only learn about the route, but also implicitly learn the "usual" route selections of the other road users. Our basic simulations have not only the advantage of being validate-able against experimental data, but can also be efficiently simulated thus promising an integration into large scale traffic simulations.

6. REFERENCES

[1] Fudenberg, D. and D. Levine. 1993. Steady State Learning and Nash Equilibrium. *Econometrica*, 61:547-573.

[2] Harley, C. B. 1981. Learning the Evolutionary Stable Strategy. *Journal of Theoretical Biology*, 89:611-631.

[3] Klügl, F. and A.L.C. Bazzan. 2002 Simulated Route Decision Behaviour: Simple Heuristics and Adaptation. In: Proceedings of the Workshop Survive (to appear), Springer-Verlag.

[4] Klügl, F., 2001. *Multi-Agent Simulation – Concepts, Tools, Application*, Addison Wesley, Munich, (in German)

[5] Selten, R., Schreckenberg, M., Pitz, T., Chmura, T. and Wahle, J. 2002. Experimental Investigation of Day-to-Day Route Choice-Behaviour. In: Proceedings of the Workshop Survive (to appear), Springer-Verlag.

[6] Wahle, J., Bazzan, A.L.C., Klügl, F., Schreckenberg, M. 2000. "Decision Dynamics in a Traffic Scenario". *Physica A*, v.287, n.3-4, p.669 - 681