Working with multi-agent-models – Pro and cons for daily routine

Florian Harder, Rapp Trans AG

Conference paper STRC 2012

STRC | 12th Swiss Transport Research Conference
Monte Verità / Ascona, May 02-04,2012

Working with multi-agent-models – Pro and cons for daily routine

Florian Harder, Phd
Rapp Trans AG
Basel

Phone: +41 61 335 78 67
Fax: + 41 61 335 77 33
email: florian.harder@rapp.ch

May 2012

Abstract

Because of its multidisciplinary character multi-agent-modelling opens a broad approach to analyze traffic based questions. Thus it is able to look at a wide spectrum of transportation phenomena. Depending on the layout of design and complexity travel behavior, theory of action and/or land use models can be implemented.

This Phd-research was made at the University of Wuerzburg and covers questions about the impact of road pricing on the reduction of negative external effects. Although there are many different forms of road pricing, nine scenarios have been developed.

Multi-agent modelling is one of the state-of-the-art techniques to evaluate the scenarios. The used multi-agent-model is created in SeSAm. This “Shell for Simulated Agent Systems“ is a tool for so-called agent based simulation and provides a generic environment for modelling and experimenting. It was designed at the University of Wuerzburg – Chair of Artificial Intelligence and Applied Informatics.

Within the model there are 250.251 agents, 203.128 of them are active. They act close to reality. This means the agents create their own individual day plan, have preferences and are subjects to restrictions, which are adapted from an empirical made population.

There are seven phases of activities. Most important is the identification of the mode used by the agent. Economic and psychological aspects provide a utility which is used to rate the available means of transport – car (cheapest and shortest route), public transport, bike and walking.

The simulated results are reaching the conclusion that each scenario generates a specific improvement of traffic and hence an increase in quality of everyday life. The results are located in quite a narrow range, even though each of them has its specific strength and weaknesses.

The results are also clarifying the possibilities of the work with such multi-agent-systems. Based upon this research I´ll like to make a case for using this approach more often in the field of non-research working, though not holding back negative aspects such as complex data mining and calculation times.

Keywords

multi agent modelling – product life cycle - SeSAm – virtual population – homogenous groups – mode choice model – time advantage – cost advantage – road pricing

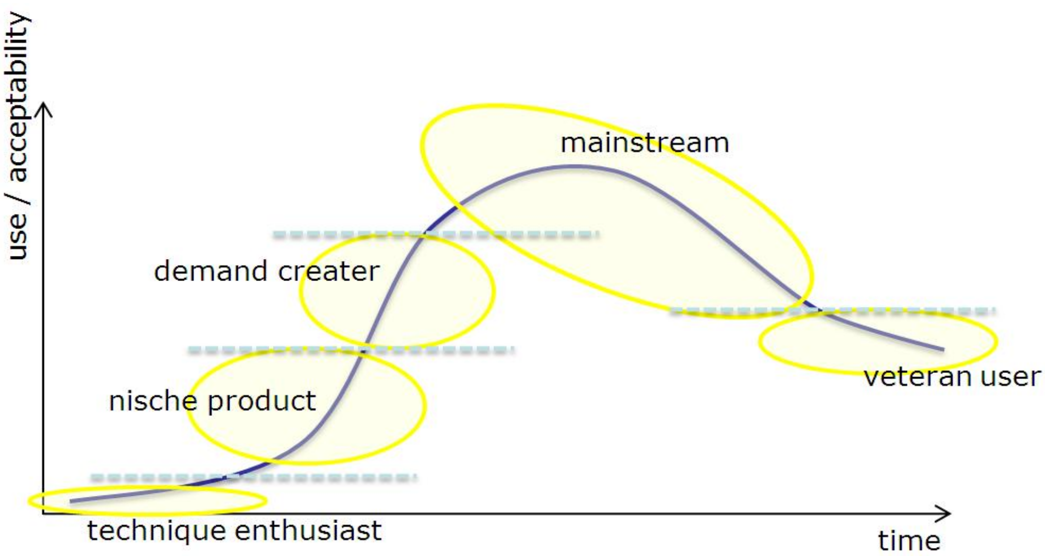
1. Introduction

Transportation modelling explains and forecasts the spatial behavior of people, goods and information in any kind of environment. These models are used in planning and science, e.g. to describe relationships between usage and position of places, range of offered means of transport, the behavior of people, companies and infrastructure as well as environment itself.

But planning agencies and research centres aim at different targets. Planning agencies mostly work on contract and hence have to stay within a limited budget of time and money. Though researchers also face budget strings they are able to work more open. Basic research is one major achievement planning agencies don´t have time and money for. Thus the result and more important the used methods differ.

Computer based innovations and new forms of data collecting and sampling made is possible to look at traffic even more detailed and interdependent and also improved the quality of modelling. But this happens mostly at software developers or research centres. The following figure shows why.

Figure 1 Product life cycle of software



Referring to the theory of product life cycles, products have an arc, just like human beings. From birth to death, human beings pass various stages e.g. birth, growth, maturity, decline and death. A similar life cycle can be seen in the case of products. The product life cycle goes through multiple phases, involves many professional disciplines and requires many skills, tools and processes. Product life cycle has to do with the life of a product in the market with respect to business/commercial costs and sales measures.

In case of transportation software this means: the more people use it the more people will do so, too. But it´s not the large number of planning agencies that start because of enthusiasm or filling a niche.

Why and when is a new modelling technique transferred from science to planning? And why does it often take so long? I´d like to find some answers to these question by looking at a new family of models, multi-agent-models, and why they aren´t used in everyday planning life yet.

2. Multi agent models

In the last years different ideas were generated to deal with the limited characteristics of traditional transportation models. Most of them can be subsumed under the category “multi agent models”. Center of this kind of modelling are homogeneous or diverse acting entities, so called agents.

Definition of agents is more or less clear. Depending on the reviewed issue and personal intention the definition terms drift apart. Close to a generally accepted terminology is FRANKLIN and GRAESSNER´s (1997) conception of an agent being “a system situated within and a part of an environment and acts on it, over time… and so as to effect what it senses in the future”.

Hence agents are physical or virtual entities situated in a natural or artificial environment. They have a (limited) perception of this environment and change it. Therefore agents have the ability to modify their individual acting and/or to act reactively; i.e. actions are not predefined, but can be adapted to new situations. This rational autonomy complies with the individual goal (partial in combination for common good as well). Agents are also public minded. This means they interact with other agents.

Another important factor of multi agent systems is the artificial environment, which contains the agents. It is the container where all perceptions, actions and movements of agents happen. In addition to that there are different objects within this system that are passive. They can be perceived by agents and agents can influence them or even change them. Objects, agents and the environment are connected by lots of relations.

In brief: Multi agent systems are computed-based systems containing agents who are in cooperation and/or completion to fulfill their individual and/or collective tasks.

This methodology is more and more used to establish models of human-environmental or social-environmental phenomena. A big advantage over traditional models is the ability to analyze the connection of interactions within a dynamic dimension

3. SeSAm model

3.1 Software

The model described in the paper is based on the shell SeSAm (Shell for Simulated Agent Systems). It supplies a generic environment for modelling and simulating of agent-based questions. Developed at the University of Würzburg (Institute of informatics – Artificial Intelligence and Applied Information) SeSAm allows an easy programming, experimenting and analyzing without asking for high-end programming skills.

The presented multi agent system covers more than 250'000 traffic participants. These agents react close to reality, i.e. (having) they have their individual day plan and preferences, and being subject to individual restrictions derived from empirical analysis.[1]

3.2 Virtual population

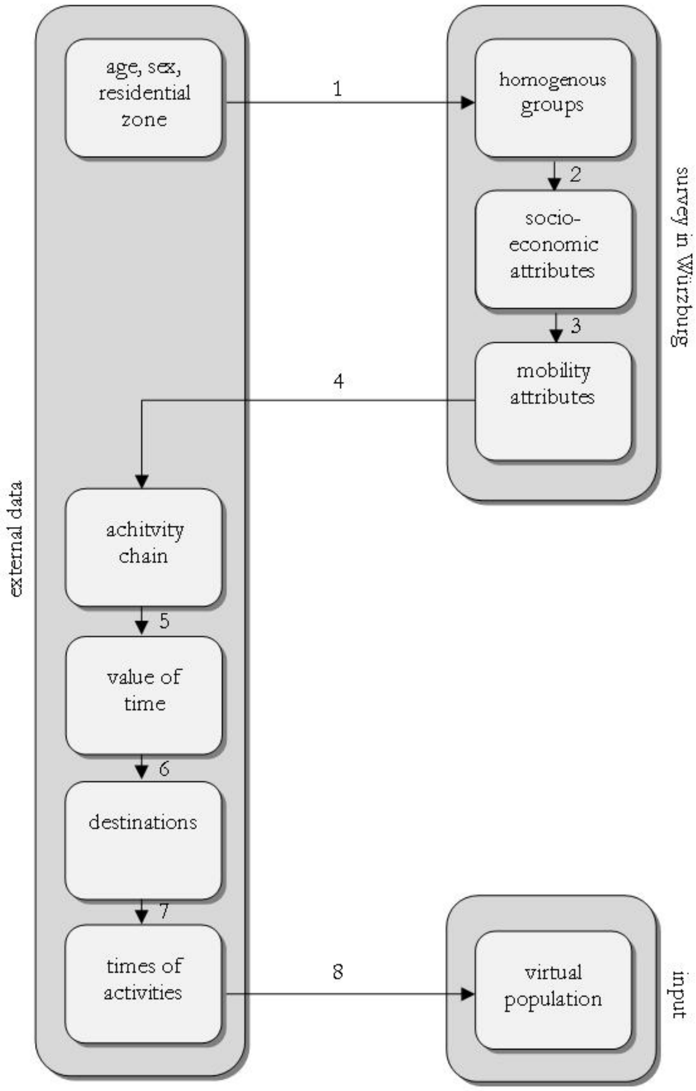
Starting point for generating a virtual population is a survey in Würzburg, carried out in February/March 2008. Almost 900 valid interviews were made. Each participant was older than 18 years, thus only potential car users were asked.

The survey included nearly 120 questions: 19 questions about the social structure, 17 about lifestyle and mobility as well as 90 stated-preferences relating to their behavior in traffic and their future acting in context of road pricing. The distribution between the sexes of all participants is well balanced (49.4 % male, 50.6 % female) and the mean age is 40.1. On average three people live within one household according to reality.

In order to model the mobility of the population of Würzburg they have to be represented as agents with all characteristics close to reality. Figure 2 shows the evolution of this virtual population.

[1] A complete description of the model can be found in HARDER (2011)

Figure 2 Evolution of virtual population



The detailed sub-steps of generating can be divided according to origin of data. In addition to external data relating to Würzburg (e.g. registration data) and general values of time survey information are used. Following each step of evolution an agent database is generated.

Starting with statistical information about age, sex and residential zone each agent has been assigned to one of the four mobility groups based on the likelihood of membership. A 26-year old man from Würzburg-Altstadt (zone 60110) is a "car driver open to public transport" by a chance of 35 percent (step 1). In 28 percent of cases he will be living in a two person household and will have graduated from high school. In this way all attributes of the agent (from profession to monthly ticket holding) are created and always derived from survey data (step 2 and 3).

Based on group specific activity chains of modified PTV-VISEM (2002) demand modelling the number and type of activities are defined. There are six different types of activity like work, shopping, education, university, leisure as well as home, which is the first and also last activity of every agent: i.e. they start and finish their daily activities at home.

In our case the agent pulls an activity chain according to profession and car ownership (step 4). It is not possible to change this daily plan due to internal or external influences. This makes sense so that there is a good reproducibility as well as comparability between scenarios.

In step 5 the agent calculates his value of time, i.e. the opportunity costs of his time. Though the value of time varies from person to person and also depends on the purpose of journey; there is just one general value because of a lack of valid data for Würzburg. The model uses CIRILLO & AXHAUSEN´s (2006) value of 13 cents per minute, but can be modified if proper data is at hand.

The definition where these activities take place refers to PTV (2002) as well. Staring point of each route is the zone of last action, thus only the next zone has to be determined to create a chain of zones. This procedure is depending on depends on several factors: the distance between origin and destination by direct and traffic related routing; the potential of possible destinations; the sensitivity of restraints which are defined by separate parameters for each behaviorally homogeneous group and activity.

The selection of destinations results from a gravity model.

$$P_{i,j} = \frac{Z_j * f(w_{i,j})}{\sum_{k=1}^{B} (Z_k * f(w_{i,k}))}$$

According to this function each zone is having has activity related attractions. These structural characteristics are shown is table 1.

Table 1 Definiton of structural features

activity	structural feature	used data	source
H – home	population	(total) population	registration office Würzburg
W – work	jobs	employees at location	survey & projection
E – education	high school students	high school students	census bureau of Bavaria and Würzburg
S – shopping	retail industry	sales area	survey & projection (considering „Einzelhandels-monitoring der Stadt Würzburg“
U – university	students at universities and vocational schools	students	census bureau of Bavaria and Würzburg
L - leisure	leisure time facilities	-[2]	-

The restraint functions are defined as follows:

$$f(w_{i,j}) = e^{-\alpha * w_{i,j}} * w_{i,j}^{\beta}$$

wi,j represents the travel time (sec) between all zones creating a matrix. Usually equated with zero the function is reduced to WILSON´s LOGIT-formulation (PTV AG, 2002):

$$f(w_{i,j}) = e^{-\alpha * w_{i,j}}$$

α signifies the impact of restraining sensibility. The larger α the more important is the impact for selecting the specific zone. That is why large α values tend to create shorter routes, smaller α values admit longer routes. If the values equate zero the distances have no impact on trip distribution. Though different homogenous groups have different needs, α values vary partially, also depending on trip purpose.

[2] Assuming homogenous attractiveness due to missing data

Table 2 α parameters

group of people	activities				
	work	university	education	shopping	leisure
working with car	0.1	0.1	0.1	0.1	0.1
working without car	0.175	0.175	0.175	0.45	0.65
non-working with car	0.1	0.1	0.1	0.3	0.5
non-working without car	0.175	0.175	0.175	0.175	0.175
student	0.1	0.1	0.1	0.45	0.5

Source: PTV AG, 2002: 5-22; modified

Continuing with defining of zones of action, table 3 shows an exemplary case for an employee with access to a car on his way to work.

Table 3 Definition of zones - example

zone	60100	60120	60210	[...]	90026
60100	0.3347	0.0062	0.0280	…	0.0001
60120	0.0014	0.7979	0.0008	…	0.0002
60210	0.1966	0.0084	0.0814	…	0.0001
[…]	…	…	…	…	…
90026	0.0054	0.0005	0.0017	…	0.9482

Thus such an agent will choose zone 60100 for work with a chance of 19.66 percent if his actual position is at zone 60210. To complete the destination choice each agent has to repeat this process as often as activities are on his day plan. An exception is his last activity which has to be his home destination as mentioned before.

Once all zones are defined the actual node has to be selected. This is necessary because the routing does not work with zones but on nodes and some zones are connected to two or more nodes. Each connector node of a zone is weighted so that there is a likelihood to pick on. Because of the homogenous structure of zones each connector has an identical weight.

The final step of generating the virtual population is to set the time for activities. Table 4 shows the duration, earliest and latest starting time as well as the latest time of terminating the activity based on real times.

Table 4 Times of activities

time	activities				
	work	university	education	shopping	leisure
minimum duration	7.0h	4.0h	5.5h	0.5h	1.5h
maximum duration	9.0h	8.0h	6.0h	2.0h	4.0h
early bird	06:00	08:00	08:00	09:00	09:00
later bird	14:00	14:00	08:30	19:30	22:30
lastest end	21:00	18:00	14:00	20:00	24:00

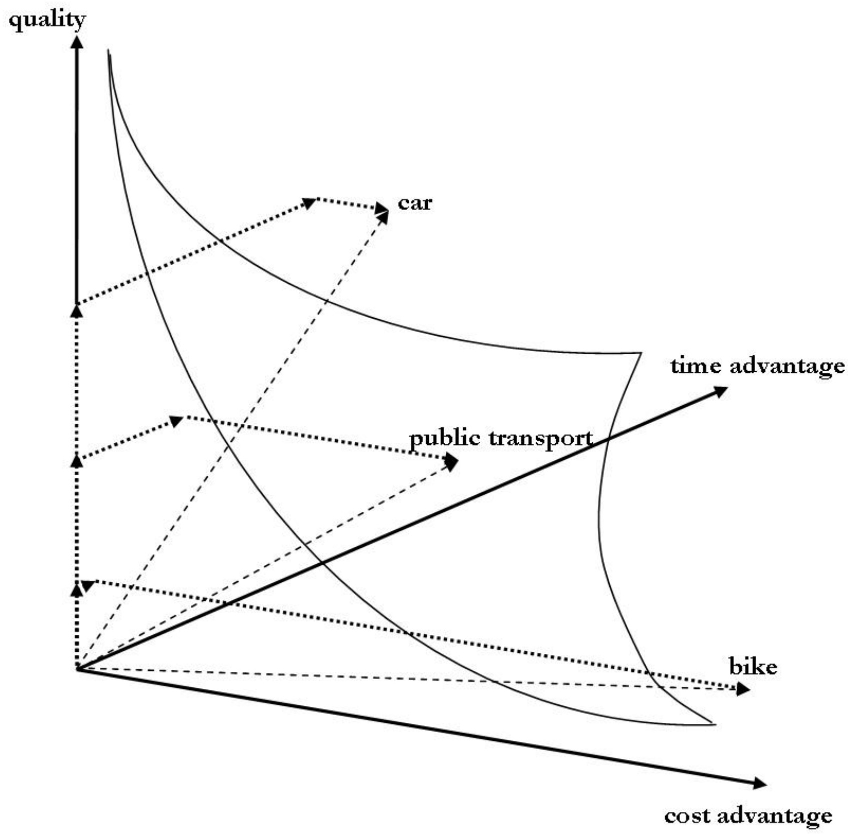
According to traditional four-step models the process is finished by mode choice and route assignment. As these two steps run dynamically and can change in different scenarios this is not part of generating the virtual population. The result is a database of agents featuring a complete set of activities concerning characteristics of each single agent as well as information about their detailed day plan.

3.3 Mode choice and trip distribution

Defining the first two steps (trip generation and trip distribution) within the process of creating virtual population, these results consistently stay the same. The third (mode choice) and fourth step (route assignment) are described in this section in detail.

GORR (1997) defines mode choice by assuming individual preferences; i.e. the indifference curves (all modes on one and the same curve are preferred equally) differ between different homogeneous groups. The sum of all characteristics results in a specific attraction of each mode and is crucial for choosing a mode together with individual preferences. The three-dimensional figure 3 demonstrates this context.

Figure 3 Optimal mode choice



Source: on the basis of GORR, H., 1997: 48

Upon the axis there are the three registered categories: quality, time and cost advantage. Though time and cost are characteristics that have a negative impact on mode choice when increasing, they are transferred into time and cost advantage by change of sign.

Modes with identical attractions are indifferent. Connecting these points a polygon is created and on the levels above and below this polygon several others are placed. Assuming that agents compare available modes they will choose that mode which reaches the highest preferences set (polygon).

If choosing rationally an agent will decide to use the car in the example at hand (figure 3), because that vector is ending ends at the preference set furthest above. The agent will gain maximum benefits.

As mentioned before travel time is a value to be minimized. The quicker the alternative, the better evaluation. That.inverse proportion is modelled by changing the sign. Central point is a fixed travel time.

$$T_i = t_{FIX} - t_i$$

Travel cost including the value of time, parking costs, tolls, costs for public transport and fuel are to be minimized as well. The calculated cost advantage is based on overhead costs (analogous to fixed travel time).

$$C_i = c_{FIX} - c_i$$

The evaluation of quality of transport as the third category is more difficult. A large amount of extra data is required (What is quality? How to measure? What is important? How do different groups value quality? etc.) which could not be covered by survey. But arguing that quality is nothing else than a form of choosing everyday anew, it can be equaled to some kind of routine; i.e. the more often someone uses a mode of transportation the better he will value the overall quality of it.

Assuming that the agents want to satisfy their preferences the best, a general conclusion can be made: One will use the mode that offers the largest time (Ti) and cost advantages (Ci) as well as is the favored mode (Rj). By introducing ft_j, fc_j, fr_j the different ratings by the homogenous groups are considered. α, β und γ are calibration factors. Set into phrase it is:

$$U_j = \sqrt{\alpha f t_j T^2 + \beta f c_j C^2 + \gamma f r_j R^2}$$

Computing five different utilities (public transport, bike, walk and two kinds of car driving – shortest in time and in cost) the agents rate the chances and take the best. This calculation of real time edge length is done by Dijkstra´s algorithm, which is implemented in SeSAm.

To shorten computing time there are some constraints: depending on prior decisions for example agents are not able to use their car if not at the same place. In addition on routes shorter than 500 meters they don´t use cars at all, on routes longer than 2´500 meters respectively 4´000 meters they do not walk or use the bike. But some agents are able to correct their mode choice or even the route. When driving to a new place they can react on changing conditions such as increasing toll rates or congestion. This action is called "on-the-fly-rerouting"

3.4 Additional agent classes

The multi agent system is not only made of manlike agents, in fact there are other kinds of agents: public transport carriers. Buses and tramways are also modelled with an engine to act.

Though being much simpler than manlike agents they only have to react in traffic condition and have people boarding or getting off. A total of 3´314 public transport agents operate within the simulated 24 hours each with an individual schedule representing the complete timetable of the city of Würzburg.

The network of this model was prepared with PTV-VISUM software and has been converted into a GIS format. By using a customized GIS- and graph plug-in, the network is imported as passive agents (objects) into SeSAm. Overall the network is built up of 2´105 nodes and 6´010 links. 482 nodes are origin or destination points and 266 nodes are operated by public transport.

3.5 Performance

The validity of the SeSAm model depends on the quality of matching reality. This is achieved by calibrating the model; i.e. by comparing real traffic densities with modelled results and adjusting parameters.

Basically there are three types of calibration parameters:

- T_{fix} and C_{fix}, which are used as variables for defining time and cost advantages,
- the α, β und γ value in benefit analysis,
- the road network.

For comparison data from a PTV-VISUM assignment is available. The Pearson correlation coefficient is a measure of the linear dependence between two variables X and Y (here: SeSAm and PTV-VISUM assignment), giving a value between +1 and −1 inclusive. The final run reaches a value of 0,786 which is a good level ($p<0{,}001$). To understand how the typical value of the dependent variable changes when any one of the independent variables is varied, while the other independent variables are held fixed, a regression analysis was made. The outcome of this is:

$$y = 1{,}0233x + 239{,}42$$

This variance is reasoned inter alia in the blocking of residential roads for through traffic and the multi-modal routing based on costs in SeSAm. The VISUM-assignment sorely calculates on time and does not allow a modal shift.

Computing requirements have a big impact on implementation and realization of models. The object-oriented programming is easy to use, but has negative effects on computer performances. That is why computing times are significant higher than expected.

Run on a computer with 4 GB RAM and a processor of 2.2 GHz (AMD 875 Opteron) simulation the system with a single agent is quite fast (approx. 10 minutes). But because of interaction between agents and (caused) rerouting processes the computing time increases as well as need of RAM.

Calculating all 86.400 (representing one day; without preparation of network and generating the agents) takes about 12 days (!).

4. Test Case

Test case of the application was a study about decreasing external effects by implementing road pricing actions. A total of ten scenarios were calculated:

- starting conditions (A): serves as comparison state and shows present-day traffic status.
- scenario K: without toll, but assuming an increase to fuel prices
- scenario KA: constant toll on roads crossing a cordon covering the city center (in- and outbound); price is 150 cent; just a few exceptions (local residents on their home-based ways)
- scenario KB: larger cordon, but variable toll (between 0 and 250 cent; depending on traffic densities; daily maximum 600 cent)
- scenario KAB: combination of KA and KB (variable toll; daily maximum 900 cent)
- scenario V: speed-based pricing; i.e. links are priced if speed fell under a certain level (maximum 1 cent per 100 meters; on all roads within city limits)
- scenario CO_2: emission-based pricing; i.e. analogous to scenario V, but with focus on emissions and differentiated by class of vehicle.
- scenario U: car-scrap bonus; according to 2009 German policy replacement of older and high-emission cars is assumed; analogous to scenario CO_2
- scenario D: distance-based pricing; cost per meter is 0.05 cent
- scenario ÖV: analogous to scenario D but public transport free of charge to cause to bigger modal shift

Table 5 shows nine evaluation values (pollution, average speed, public transport delays, external costs, revenues, social fairness, fairness between homogenous groups, traffic density and number of passengers.

Table 5 Scenario evaluations

	K	KA	KB	KAB	V	CO2	U	D	ÖV
pollutant emissions	+	+	+	+	++	+	++	+	+
average speed	+	+	+	+	++	++	++	++	++
public transport delays	+	++	+	+	++	++	++	++	++
external costs	+	O	++	++	++	++	+	+	++
revenues	++	+	+	+	O	+	+	+	--
social fairness	+	O	O	O	-	O	-	O	+
fairness between homogenous groups	+	+	+	+	++	--	-	--	--
traffic density	O	+	+	+	++	++	+	++	++
number of passengers	+	+	+	+	++	++	++	+	++

Every scenario shows a small improvement in pollutant emissions (CO_2). Especially scenario V und U generate markedly reduction compared to start conditions. Looking at average speed it is particularly noticeable that cordon-based scenarios as well as scenario K increase average speed but not as much as the other scenarios. A detailed view on public transport delays reveals a similar picture; except KA which also has significant changes, but doesn't have any advancement in external costs.

Generating revenues is unequally achieved by the different scenarios. On the one hand scenario K generated a high value by increasing fuel prices and on the other hand scenario ÖV gives away a lot of money by introducing a free public transport.

Over all, social fairness alters on the level of starting conditions. Scenario K and ÖV preserve equality of opportunities best, V and U worst. Evaluating fairness of homogenous groups (mobility style groups) $CO_{2,}$ D, ÖV and U have negative effects, favouring some groups and discriminating others groups.

Since they are an important part of creating acceptability traffic densities are crucial in evaluation. The results look quite similar to the outcome of pollutant emissions (*unsurprisingly*). Finally all scenarios increase the number of passengers. Scenario V, $CO_{2,}$ U and especially ÖV succeed very well at this.

In conclusion the nine scenarios are close together, each one having individual strengths and individual weaknesses. In fact only two scenarios can be highlighted at all: scenario V and CO_2.

5. Multi agent models in daily routine

Even though most of multi agent models deal with questions about – in the broadest sense – game theory it turns out that multi agent systems can also model traffic related topics. There is no need to limit such applications to shopping patterns (RAUH, 2008) or analysing non-dynamic phenomena such as the choice of doctors in rural areas (NEFF, 2011).

But why are there just a handful of reasonable multi agent projects? Why are these auspicious methods not used by consultants and transport authorities in everyday planning life?

When you ask people these questions answers like the following are heard:

- I do not see suitability for daily use!
- That does not work! I've never seen something like that.
- Interesting! But the market wants mainstream methods.
- Those are ideas from ivory towers of universities.
- We cannot afford this.
- This method is not well-engineering. The results will be wrong.
- Calculation times are too long.
- We do not have the know-how.

Some of the aspects are just simple exaggeration, others are serious. It is true: no one wants to wait nearly two weeks to get results, when most of the time findings should be at hand "preferably yesterday". And yes, maybe know-how has to be set-up. But on the other hand there are lots of opportunities connected with detailed modelling of individual transportation.

To some this is just academic game playing but others see in multi agent modelling the final step to answer important questions ("Is that new bus line an important alternative for everyone? Are residents or immobile people put at disadvantage? etc.). Nowadays consultants look mostly at overall results, but not on individual steps.

Maybe the biggest advantage of multi agent modelling over traditional models is its dynamic environment. Since it is able to change circumstances (i.e. pricing, variable message signs or even the weather) it gives new finding to compare different alternatives.

But what can be done to spread the usage of multi agent systems and to work with them on daily routine?

Most important is to get to know them. The more people know about this modelling option the more often it with be discussed as a real alternative to traditional modelling. Probably the best way to demonstrate suitability for daily routine is to launch lighthouse projects like analyzing "Westumfahrung Zürich" using MATSim. Raising a new user-generation is another good step.

By running more and more multi agent projects a specific kind of data standardisation will take place also effecting a homogenization of data (what kind?; from what source?; in what format? etc.). This is extremely important because multi agent systems need a huge database to satisfy diverse agent classes. This will also help to increase the number of applications while enabling a (partial) transferability to other regions or related questions.

Referring to figure 1 we have to take "one giant leap" by increasing the number of applications. This will change the meaning of multi agent models to a niche product or even a "demand creator", maybe generating snowballing.

And by the way: There are already multi agent applications in daily routine. You just have to take a closer look…

6. References

Cirillo, C. and Axhausen, K.W. (2006) Evidence On The Distribution Of Values Of Travel Time Savings From A Six-Week Diary. *Transportation Research A*, Vol. 40, No. 5, 444-457

Franklin, S. and Graessner, A. (1997) Is it an agent, or just a program? A taxonomy for autonomous agent, in Jennings and Wooldridge (Ed.) Intelligent Agents III, Berlin, 21-35

Gorr, H. (1997) Die Logik der individuellen Verkehrsmittelwahl. Theorie und Realität des Entscheidungsverhaltens im Personenverkehr. Gießen

Harder, F. (2011) Straßenbenutzungsgebühren und nachhaltiger Stadtverkehr: Eine agentenbasierte Mikrosimulation, Mannheim

Neff, C. (2011) Werkstattbericht. Auswirkungen des demographischen Wandels auf die ambulant hausärztliche Versorgung im ländlichen Raum. in Helbich, et.al. (Ed.): *Theorie und quantitative Methoden in der Geographie.* Kolloquiumsbeiträge (=Heidelberger Geographische Bausteine Heft 19), Heidelberg, 103-112.

PTV AG (2002) Benutzerhandbuch VISEM 8.0. Karlsruhe

Rauh, J., Schenk, T., Schrödl, D. (2008) Agentenbasierte Simulation von Konsumentenhandeln, in Mandl und Koch (Ed.) *Modellierung und Simulation komplexer geographischer Systeme*, Salzburg, 77-104

Lehrstuhl für Künstliche Intelligenz und Angewandte Informatik (2012) SeSAm - Multi-Agent Simulation Environment, *http://www.simsesam.de/*, University of Würzburg