Land use planning of big traffic generators: the need for new tools

P. Giordano, MACS-Lab, University of Lugano
A. Vancheri, MACS-Lab, University of Lugano
D. Andrey, PhysArch sagl

Conference paper STRC 2008

STRC | **8th Swiss Transport Research Conference**
Monte Verità / Ascona, October 15-17, 2008

Land use planning of big traffic generators: the need for new tools

P. Giordano
MACS-Lab
Mendrisio

Phone: +41586665729
Fax: +41586665868
email: pgiordano@arch.unisi.ch

A. Vancheri
MACS-Lab
Mendrisio

Phone: +41586665729
Fax: +41586665868
email: avancheri@arch.unisi.ch

D. Andrey
PhysArch sagl
Viganello

Phone: +41919722468
Fax: +41919703163
email: dandrey@physarch.ch

October 2008

1. Abstract

We present a mathematical model based decision support system for the placement of big traffic generators (BTG) like shopping centres. The model aims to forecast the traffic flows induced by BTG on the road network of Canton Ticino in Switzerland, and is beginning to be used by the public administration as a support to decide master plan rules for zones selected for the reception of BTG. In general, it can also be used for the planning of placement or rejection of new projects of BTG. The system has been modelled as an Interaction Space, a new type of mathematical structure generalizing both multiagents systems and cellular automata and aiming to axiomatically define complex systems made by several interacting entities. The presented model use methods more near to the way of thinking of Physics, like the use of populations of individuals, but with several positive elements inherited from multi-agents systems. Indeed, we can say that the model is an activity-based modelling of the whole private transportation system of Canton Ticino, but with the use of populations of agents instead of individual agents. Fuzzy logic based methods, instead of the classical random utility theory, have been used to define several indicators, both to model agents' behaviour or attractiveness indicators for zones containing economical activities. The use of these indicators permits to obtain agents' location choices of secondary activities with a great and scalable computational efficiency. Suitably organized simulations obtained from the model include several environmental indicators, and can be used to support decisions in the placement of new BTG or in the definition of areas intended to receive BTG. A partially incomplete, but anyhow meaningful, version of the model has also been used to test a classical planning strategy consisting in the imposing of a maximum level of commercial surface suitable for BTG in a prefixed number of zones (called divisions). We performed a suitable simulation where every division has been filled uniformly in time until to arrive at its filling, with the aim to forecast the situation of worst environmental impact caused by BTG located in divisions. Some final results of this simulation has been unexpected, like the decreasing of NO_x in some points (but not all) with respect to the present situation. This underline the need for new planning strategies in the displacement of BTG.

Keywords

environmental risk management – activity based modelling – location of secondary activities – fuzzy logic based indicators – multiagents systems.

2. Introduction

A decision support system (DSS) to help public administrations or private companies in the placement of big traffic generators (BTG; e.g. shopping centres, tourists attraction shows, very attended public places like important monuments, etc.) must include several features:

- it has to be based on a mathematical model of traffic;
- it has to consider traffic related decisions of agents at the regional level, including abroad-based trips;
- it has to adapt the behaviour of households in hypothetical situations with different BTG placed in different locations with respect to the present one. Indeed, the DSS has to be able to accept, at least exogenously, new projects of BTG to evaluate them from an environmental and economical point of view. Therefore, the agents of the model must adapt their behaviour due to the presence of these new BTG.
- It has to include several types of movements, i.e. both based on different sets of goals performed in chains of zones, amount of spent time or spatial ranges and more systematic trips like classical home-work or home-shopping.
- It has to describe the attractiveness of BTG by means of faithful (i.e. injective) attractiveness indicators depending on features of BTG and on the types of movements we are considering.
- It has to include the estimation of important environmental indicators, like the level of NO_2, some measure of roads' level of service, an estimation of the covered distance in making use of a BTG and the possibility to separate the contribution of each BTG in the total amount of a given environmental observable.
- It has to include the estimation of economical indicators, like the average cost of the shopping, the amplitude of the catchment area of each shopping mall or some measure of the risk to lose clients with the arrival of new competing shopping centres.

The kind of observables we are interested in (see above), focuses our attention about the type of available model for this problem toward a multi-agents system (MAS). Indeed, models that use methods well-known in Synergetics, like in Weidlich and Haag 1999 (see also Helbing and Nagel 2004 and references therein), because they lose the identity of each started trip, seem not suitable to consider the effects of chains of movements or to evaluate indicators like the amount of NO_2 ascribable to a specific shopping centre among two acting in synergetic cooperation. It can also be useful to consider that, on the basis of recent statistical survey about the behaviour of users of shopping centres in Canton Ticino (see Alberton and Guerra 2008) a non negligible percentage between 27.7% and 32.7% do not realize only a home-shopping-home trip, but a more complex one, with several goals performed in a chain of different zones.

MAS for regional transportation problems including activity-based simulations of secondary activities (as opposed to home-work or home-school trips, e.g. to do one's shopping, to move for leisure or having a touristic goal, etc.) is beginning to be completed in these years, see e.g. Arentze and Timmermans 2007, Balmer 2007, Balmer et al. 2004, Marchal and Nagel 2005. In the present paper we used fuzzy logic based indicator instead of the classical random utility theory. On the basis of our experience (see Vancheri et al. 2008a, 2008b) this permits to obtain a more realistic modelling of agents' decision processes. As we will see, the use of fuzzy logic based attractiveness indicators for BTG permits to obtain good and scalable computational performances in the problem of the allocation of secondary activities by agents. Moreover, this allocation does not assume, as in Marchal and Nagel 2005, an individual's partial knowledge of locations, which seems questionable in a small region like Canton Ticino.

To complete the state of the art of works related to the present one, we can say that integrated models of traffic and urban land use (see e.g. Bürgle et al. 2005, Salvini and Miller 2005, Strauch et al. 2002, Arentze and Timmermans 2000, Waddell 2000, Landis and Zhang 1998 and references therein) do not include an endogenous or at least exogenous sub-model for the insertion of new BTG and the consequent adaption of households in their weekly agendas.

2.1 Populations of agents instead of individual agents

Our approach, even if it incorporates several positive features of a MAS, like the level of details in the modelling of trips, can be considered as methodologically different from a pure agent-based model. Indeed, from the point of view of several physicists and mathematicians, MAS try to reconstruct a piece of reality in a very detailed way (which sometimes is a different point of view with respect *to model* a piece of reality). To satisfy this aim, agent-based model obviously need a great amount of statistical data. Frequently these data come from the last census, but due to the time interval of 10 years between two subsequent censuses, for some types of problems census data are not sufficiently updated. E.g. if we want to estimate the environmental impact of a BTG built in 2008, and we use the data from the 2000 census to construct the initial condition of our model, we are forced to start the simulation in 2000 and to last it 10 years to have two year of environmental impact. Moreover, we are also forced to include a urban development model. Finally, it is questionable if from a census it is really possible to obtain *reliable* information about economic management of families, life styles, leisure habits of different members of each family and even less about their stochastic nature and reciprocal correlations.

Even worst is the situation if viewed from partial surveys, particularly for their representativeness in situations, like ours, strongly depending on individuals' home position (e.g. how many times an individual do a big shopping instead of a small local shopping depends on the distance from individual's home to shopping centers and the availability of small shops near home).

For all these reasons, we do not feel up to affirm that the state of each single agent, e.g. her/his habits expressed using a personal agenda of activities, do correspond to reality (a position related to the one expressed here, about the validation of MAS, can also be found in Casti 1997). In our point of view, it seems more realistic to think that the initial state assigned to each single agent in a pure MAS, could be *only one among several possible states compatible with available statistical data*. For example, suppose we have an empirical statistical distribution about (in a given zone z) "the number of times an individual does local shopping in a week", about "the number of times an individual does non-local shopping in a week", and about "the number of times an individual does an evening free-time movement in a week". Suppose that each one of these random variables is given by means of a histogram with respectively N_{loc}, $N_{\text{non-loc}}$ and N_{free} intervals/bars, so the possible combination (i.e. a priori all the possible initial states of agents living in the zone z) are $A \cdot N_{\text{loc}} \cdot N_{\text{non-loc}} \cdot N_{\text{free}}$, where A is the number of agents living in z. Of course, some of these combinations will be improbable, but in a survey usually we cannot have a reliable estimation of the correlations between these random variables.

In our opinion, choosing only one of these combinations (initial states) means to include in a model a knowledge going over the pure statistical data and the usual hypothesis always present in every mathematical model. This knowledge (correlations) not necessarily corresponds to reality, and hence it seems methodologically not correct to include it in the model. Unless to prove that our conclusions drawn from the use of the model, do not depend on the particular chosen initial state, e.g. doing several simulations with several initial states.

For these reasons, as we will see in this article, in our approach we use temporary agents whose state is generated on the basis of the pure statistical data only, without assuming unknown correlations. Temporary agents are destroyed exactly after they finish the trip they are performing. In other words, if at time t we have generated a trip of type "doing a non-local shopping", and after at $t' > t$ we generate a trip of type "free time", in our model it is not possible to affirm that these two movements are really performed by the same agent. As regards the aim to support decisions about BTG, this correlation seems not important, because the important thing is to register two movements and not that the driver is the same.

In this way, we lose some correlations and even if we have a complete individual identity of each trip[1], we can say that we are following a point of view more similar to the one used in Weidlich and Haag 1999 and not characteristic of a pure MAS.

3. Complex mobility patterns

The principal aim of our model is to develop a DSS for helping the administration of Canton Ticino to evaluate planning strategies related to BTG. Nevertheless, the model could also be used to help private firms to estimate the environmental impact of a given BTG (e.g. a new shopping centre) or in choosing the best placement of a new BTG among a set of given locations.

The objective of the cantonal administration is to define:

1. what is a BTG in a normative sense;
2. what zones of Ticino are suited for hosting new BTG;
3. what thresholds should be applied for limiting the amount of BTG in each zone.

These elements should be defined in such a way to maximize both the attainment of environmental sustainability of new BTG and the economical benefits derived from BTG. The environmental sustainability has been measured through the concentration of NO_2 and an estimation of roads' level of services (LOS) using the ratio $\frac{q}{q_{\max}}$, where q is the simulated flux inputted in a given link of the network, and $q_{\max}$ is the maximum value of this input flux, determined on the basis of the physical characteristics of the link (the so-called *capacity* of the link). Economical benefits can be estimated using the money spent by each agent in each type of BTG during the simulation. The BTG model is expected to support decisions in all the above-mentioned tasks, especially in evaluating the nonlinear effects of synergy or antagonism among different BTG, roads' traffic and structure of the demand. The system has been modelled as an Interaction Space (IS), a new mathematical structure aiming to axiomatically define complex systems made by several interacting entities, see Giordano et al. 2007. We shall give an intuitive introduction of IS in section 4, sufficient for the comprehension of this article.

[1] Indeed, we can say that our model is a pure MAS, but where "agents" are single trips.

3.1 General data about the model

The case study area has been subdivided into 517 zones, coming from a VISUM© model (see PTV 2008) and corresponding to towns, districts of larger cities and border zones. Each internal zone is described by the number of resident individuals and by the presence of economical activities, the last ones described by means of the number of employees or the commercial surface. Activities relevant with respect to BTG problems have been subdivided into 45 categories accordingly to the *General Classification of Economic Activities* (see NOGA in www.bfs.admin.ch). Custom points have been also included and called external zones; they are described by their position only. The main road system is given in terms of roads' segments (called *links*) described with data like position of starting and ending points, speed limits, slope, number of lanes, etc. Agents' decision processes are based on the main time step of a week from Monday to Saturday, so that throughout this article we will always have $\Delta t = 6$ days, called, for the sake of brevity, a *week*. The week is subdivided in *time windows* $w_1, \ldots, w_{n_{\mathrm{tw}}}$, with $w_i \subset \mathbb{R}_{>0}$, $w_i \cap w_j = \emptyset$ and $\max \cup_j w_j - \min \cup_j w_j = \Delta t$. In the present implementation of the model, which is only partially complete, we have considered, conceptually, two time windows, corresponding to ferial days $w_1 = [0, 4 \cdot \mathrm{day})$ and the weekend $w_2 = [5\,\mathrm{day}, 6\,\mathrm{day}]$. Therefore, the k-th week of the simulation will be $(k-1) \cdot \Delta t + \cup_j w_j$. Of course, the decisional process of agents, e.g. with respect to BTG related trips, strongly depends on the time window w_k in which they occur, and indeed, the time windows have to be chosen as statistically relevant subdivisions of the week with respect to these decisional processes. For this reason, we can also consider situations like $w_1 = \cup_{d=1}^{6} [15^{00} \text{ of day } d, 17^{00} \text{ of day } d)$ for taking into account trips of type "back to home from school".

3.2 Definition of complex mobility patterns

One of the problems in the modelling of BTG related dynamics is that individuals follow "complex mobility patterns", i.e. movements that are not only origin-destination, but can include more than one destination for more than one goal. For these reasons, we introduce the notion of *complex mobility pattern* (CMP), but before introducing it formally, we need some elements that define a sort of grammar useful to describe a huge family of different trips. Using this grammar we will be able to automatically associate to the description of each trips an attractiveness indicator to identify the best zones for the accomplishment of that trip. A first step to describe a BTG related trip is to list its goals. Roughly speaking, goals coincide with categories of goods that can be bought (e.g. "first necessaries" or "complements"), with ways to spend free time (e.g. "shopping" or "tourist trip") or with more systematic movement (e.g. "go to work" or "go to school"). We hence introduce a set of labels $G := \{g_1, \ldots, g_{n_g}\}$

representing these goals. In the present version of the model we have considered $n_{\text{g}} = 8$ different goals:

- to do shopping for first necessaries/foodstuffs,
- to go shopping in shopping centers,
- to go shopping in central urban contexts,
- to do shopping for interior decoration/small furnishing,
- to do shopping for furnishing,
- to go in "do it yourself centres",
- to go in a clothing shop
- to realize a systematic movements like "home-work" and "home-school".

The second step is to describe some attributes of these goals, like:

- a fuzzy predicate to express the truth value of the sentence "I have spent little time for this trip". The membership function of the fuzzy set of time intervals expressing this sentence is well represented by a decreasing function μ, whose characteristic parameters can be the threshold value t_{thr} over which $\mu(t) < 0.9$, and the saturation value of time $t_{\text{sat}} > t_{\text{thr}}$ over which $\mu(t) < 0.1$. For example a "regional" trip, in Canton Ticino, can be expressed taking t_{thr} between 30 and 40 minutes depending on the statistical data one can consider. In the present version of the model we have considered three different membership functions μ_{loc}, μ_{sem} and μ_{reg} for trips of type "local", "semi-local" and "regional".
- Another type of attribute for trips' goals is the amount of estimated money for the accomplishment of the goal $g \in G$. If the range of all the money spent by all the individuals in a given socio-economical class c, is $[m, M]$, then we can model this attribute with a number $\sigma_g \in [0, 1]$ expressing the percentage of the money range that the agent want to spent in the considered trip, so that for $\sigma_g = 0$ our agents of class c will spend m, more money for greater σ_g, till to arrive at M for $\sigma_g = 1$. In the present version of the model we have only one average socio-economical class c, and we have considered only three values of $\sigma_g \in \{0.1, 0.5, 0.9\}$, corresponding to "small expense", "average expense" and "great expense".

We can now define a CMP as a 4-tuple $m = (k, \omega, \{I_j\}_{j=1}^{n}, c)$, where:

- $k = z_0 p_{01} z_1 p_{12} z_2 \dots z_m p_{m0} z_0$, called *kernel*, is a spatial polygon made of an ordered cyclic sequence of zones[2] $(z_0, z_1, \dots, z_m, z_0)$ connected by paths p_{ij} (i.e. sequences of connected links)[3]. The kernel defines *where* the trip will be accomplished.

[2] Usually one can restrict the kernel to a maximum of three visited zones, i.e. $m \leq 3$.

- ω is the preferred time window to start the trip, i.e. an index $\omega \in \{1, \ldots, n_{\text{tw}}\}$. This time window define *when* the trip will start.
- $\{I_j\}_{j=1}^{n}$ is a family of subsets of goals, i.e. $I_j \subseteq G$. Each I_j has to be thought as a set of goals that have to be realized in a single zone and the aim of the trip will be to accomplish all the sets of goals $I_1, \ldots, I_n$, in general in different zones. The subsets of goals define *what* will be done in the trip.
- Finally, the *characteristics of movement* (CM)

[Beginning of the document][Automatic section break] $c = (\alpha, i_1, \ldots, i_n, \mu, \sigma)$

is a family of parameters and functions of $g \in G$ defining boundary conditions under which the movement takes place (i.e. *how* to accomplish the trip). The parameters we considered are:

1. the *accomplishment* of each goal $g \in I_j$ in a given subset of goals, i.e. real numbers $\alpha_g^j \in [0, 1]$ summing to 1:

   $$\sum_{g \in I_j} \alpha_g^j = 1$$

   The interpretation of α_g^j is that each goal $g \in I_j$ will be accomplished at the level $\alpha_g^j \in [0, 1]$. For example if in $I_1 = \{g_{11}, g_{12}\}$ we want to realize "g_{11} a little, but principally g_{12}", we can set $\alpha_{g_{11}}^1 = 0.1$ and $\alpha_{g_{12}}^1 = 0.9$. More synthetically we can write that the subset of goals I_j is accomplished as the formal weighted sum

   $$\sum_{g \in I_j} \alpha_g^j \cdot g$$

2. The relative *importance* of each subset I_j of goals, i.e. real numbers $i_j \in [0, +\infty)$, each i_j interpreted as a fuzzy modifier for the attractiveness of the subset of goals I_j.
3. A function $\mu : [0, +\infty) \longrightarrow [0, 1]$, called the *evaluation of time*, is a decreasing fuzzy membership to define how much time has to be spent for the accomplishment of the trip. In the present implementation we have $\mu \in \{\mu_{\text{loc}}, \mu_{\text{sem}}, \mu_{\text{reg}}\}$.
4. $\sigma_g^j \in [0, 1]$ is a number expressing the expected value of the ratio

   $$\frac{S - m}{M - m},$$

   where S is the expected money that will be spent in the trip. The numbers σ_g^j are called the *coefficients of expense.*

[3] In a kernel we have $m \geq 0$ and $m = 0$ identify very local trips.

In general a CMP $m = (k, \omega, \{I_j\}_{j=1}^n, c)$ represents formally the trip:

- to do in $k = z_0 p_{01} z_1 p_{12} z_2 \ldots z_m p_{m0} z_0$ the goals $\sum_{g \in I_1} \left(\alpha_g^1 \cdot g \text{ spending } \sigma_g^1\right)$ and ... and $\sum_{g \in I_n} \left(\alpha_g^n \cdot g \text{ spending } \sigma_g^n\right)$,
- in the time window ω.
- The importance of the goals $\sum_{g \in I_1} \alpha_g^1 \cdot g$ is i_1 and ... and the importance of the goals $\sum_{g \in I_1} \alpha_g^1 \cdot g$ is i_n.
- With respect to the duration, the trip is of type μ.

Or, more synthetically:

$$\text{IN } k \ \text{ AND } \ \omega \ \text{ AND } \ \mu \ \text{ AND } \ \bigwedge_{j=1}^{n} \left(\sum_{g \in I_j} \alpha_g^j \cdot g \ \text{ SPENDING } \sigma_g^j \right)^{[i_j]}$$

The pair $(\{I_j\}_{j=1}^n, c)$ is called a *schema of CMP* (SCMP) or, more quickly, a *type of movement*. We can say that the SCMP $(\{I_j\}_{j=1}^n, c)$ is realized in the CMP $m = (k, \omega, \{I_j\}_{j=1}^n, c)$, that is in a specific kernel k of zones and paths and in a specific time window ω. In general, a SCMP can be realized in several ways and this motivates the word *shema*. A CMP is called *finalized* if each subset of goals is made of one goal only, $I_j = \{g_j\}$, and $\alpha_{g_j}^j = 1$. They represent trips where agents do goals as very important single aims. For example:

$$\text{IN } k \ \text{ AND } \ \omega \ \text{ AND } \ \mu \ \text{ AND } \ \bigwedge_{j=1}^{n} \left(g_j \ \text{ SPENDING } \sigma_{g_j}^j \right)^{[i_j]}$$

We will call a CMP *diffused* if we have only one subset of goals, i.e. $n = 1$ and $i_1 = 1$:

$$\text{IN } k \ \text{ AND } \ \omega \ \text{ AND } \ \mu \ \text{ AND } \ \sum_{g \in I_j} \left(\alpha_g^j \cdot g \ \text{ SPENDING } \sigma_g^j \right)$$

We will call a CMP *semi-finalized* if at least one subset of goals is made of one goal only:

$$\text{IN } k \ \text{ AND } \ \omega \ \text{ AND } \ \mu \ \text{ AND } \ \bigwedge_{j=1}^{n_1} \left(g_j \ \text{ SPENDING } \sigma_{g_j}^j \right)^{[i_j]} \ \text{ AND } \ \bigwedge_{j=n_1+1}^{n} \left(\sum_{g \in I_j} \alpha_g^j \cdot g \ \text{ SPENDING } \sigma_g^j \right)^{[i_j]}$$

A typical example of finalized CMP is the following:
- to buy necessaries, complements and clothes (g_1, g_2, g_3) in the zones z_1 and z_2;
- starting from home z_0, using the paths p_{01}, p_{12} and p_{20} (hence $k = (z_0 p_{01} z_1 p_{12} z_2 p_{20} z_0)$);
- on Saturday ($\omega = 2$ in the present version of the model);
- with necessaries "much more important" than complements and clothes (e.g. $i_1 = 2$ and $i_2 = i_3 = 1$);
- in a regional trip ($\mu = \mu_{\text{reg}}$) and with an average expense ($\sigma^j_{g_j} = 0.5$).

It is clear that using CMP we have a language to express a huge family of non-trivial movements. It is also intuitively clear that a SCMP $(\{I_j\}_{j=1}^n, c)$ contains all the information one need to evaluate if a given BTG is attractive or not for a movement of this type. Indeed, at each CMP we can always associate a fuzzy logic based attractiveness indicator (see section 6). Therefore, we can say that SCMP are the basic components of an algorithm able to probabilistically and dynamically generate new agendas for our agents (see section 5 for the mathematical details about this generation). On the other hand, the use of CMP needs a certain amount of statistical data about the habits of households in BTG related movements; to find these data, the *micro-census about traffic behaviour* of the Swiss Federal Office of Statistics can be of help.

Finally, we have 16 types of movements, obtained combining goals in different ways and characterized by different value of the parameters $c = (\alpha, i_1, \dots, i_n, \mu, \sigma)$:
- *Semi-regional*, that is with a typical connectivity time with the destination zones of about 20 minutes,
- *Regional*, that is with a typical connectivity time with the destination zones of about 30-40 minutes.
- *Local*, that is taking place in a neighborhood of the starting zone.
- *Finalized*, characterized by the presence of only one main goal.
- *Semi-finalized*, characterized by the presence of one or more additional goals with a medium degree of importance beside a more important main goal.
- *With diffuse finalisation*, characterized by several goals with comparable level of importance.

In the following we shall suppose to have fixed a set

$$\tau_1 = (\{I_j^1\}_{j=1}^{m_1}, c_1), \dots, \tau_\chi = (\{I_j^\chi\}_{j=1}^{m_\chi}, c_\chi)$$

of $\chi > 0$ different SCMP, that is of types of movements.

4. Interaction spaces theory

Interaction Spaces (see Giordano et al. 2007) can be seen as a good interpolation between AI based methods, more typical of multiagents systems (MAS), and Physics' methods more typical of Synergetics and Econophysics (see e.g. Haag, 1989). Indeed, as a generalization of MAS, IS are a good environment for the use of AI based methods in modelling decision processes of agents. On the other hand the mathematical setting given by IS' axioms makes possible the use of modelling tools nearer to Mathematics, e.g. continuous state variables instead of qualitative ones, populations of agents instead of single agents and differential equations for extensive variables and their probability distributions (see Vancheri et al 2008 a, b). IS aim at modelling complex systems enclosed in the following general informal frame:

1. The system is made of *interacting entities* described by *state vectors* evolving in time.
2. This state vectors' dynamics is either an effect of *interactions* occurring at stochastic times or a deterministic function of time, depending only on the state of the considered entity (in the latter case we will use the name *internal evolution of the state*). We can also say that interacting entities are modelled like dynamical systems driven by input currents produced by entities interacting with it.

   Examples of internal evolution can be: a bouncing billiard ball; a pedestrian between two subsequent interactions with other pedestrians or obstacles; the process of building of a house after its starting time and before its end, and the internal evolution of a box in a flow chart representing a computer program.
3. Each interaction can be described as a *causally directed process* in which a set of *agent* entities, modify the state of a set of *patient* entities through *propagator* entities. Propagators can be thought as entities activated by agents, and carrying the signals/currents sent by agents to patients to modify their state. The advantage in introducing propagators currying signals (instead of only signals) consists in the possibility to involve signals in interactions like if they were ordinary entities. This is relevant e.g. in cases where signals can be interpreted as information flows that can be modulated and influenced by other entities.

   Examples can be: a physical interaction between one particle that sends a signal to another particle; or a firm (agent) sending an advertisement (propagator) and hence changing the state of several people (patients); a suitable set of goods (agents) in a market sending a signal (propagator) that carries information useful for buyers (patients); a biological entity (agents) sending a chemical signal (propagator) to another entity (patients) having receptors able to recognize that signal, an object in an object oriented program sending a message to another object and a developer in an urban system (agent) that build a house in a lot of terrain (patient). In the latter example the propagator is given by the planned action, that can be identified with the building request.
4. An interactions can occur only if all the agents involved in a given interaction are *active*. In the state of each interacting entity there is always a Boolean variable indicating if, with

respect to a given interaction, the entity is active. Intuitively an agent is active with respect to an interaction if it can start that interaction and all the other agents of that interaction are also active. Active agents can be also interpreted in biological terms as entities sending some kind of chemical signals and to patients entities having receptors act to recognize it; in this description propagators are entities carrying the signals. Agents pass to an active state as a consequence either of an interaction or of the internal dynamics.

Examples: in the above mentioned example about firm's advertisement, only people activated in some way for the advertised products will have a state modification; only the biological entities having suitable receptors are active for the corresponding interactions; only hungry predators are active for hunting preys; only software objects with a suitable public state variable can receive a message to change that variable.

5. The occurrence of an interaction and its effects depends only on the “information” collected from a (possibly small) set of entities called the *neighbourhood* of the interaction. The neighbourhood of an interaction includes always agent, patient and propagator entities. Because in general there are no explicit interaction connecting entities in the neighbourhood of the interaction i with agents, patients and propagators of i, we can say that neighbourhood is a way to simplify a model without being forced to make explicit these interactions regarding the collecting of information. Of course, it is always possible to define all the entities in the neighbourhood as agents of the given interaction, so that we can always think empty the neighbourhood, but this may not always be the best model.

   Examples: if an agent is searching for a new house, only the information collected in some order in its memory will affect its future decisions; only the state of the cells belonging to the neighbourhood can influence the future state of a given cell in a CA; only the objects in the visual field of a pedestrian may influence its goal-oriented path; the information collected in a graphical user interface may influence the starting or not of a given program.

6. Interactions occur at random times, whose distribution depends on the history of interaction's neighbourhood only.

   Examples: an interaction where an agent chooses a shop on the basis of its information about quality, prices, and goods availability occurs at random times with a suitable distribution depending both on objective and subjective characteristics; an interaction describing the leasing out of a house occurs at random times depending on several factors, e.g. the rate of birth, of marriage, of immigration, etc; the starting of a program randomly depends on the interaction of the user with the program's interface.

7. The change of the state of a patient entity depends only on the internal evolution of the interaction and on a set of stochastic variables, the already mentioned signals, temporary produced by the agents at the occurring time of the interaction. More precisely we can say

that agents probabilistically extract a state σ from the state space of each propagator p and assign the state σ to p. The change of the state of each patient of the interaction depends only on this σ and on the internal evolution of the interaction. The intuitive name *signal* is only another name for the state σ. Signals' probability distribution depends only on the history of interaction's neighbourhood.

Examples: an excited electron (agent) produces a photon (propagator) that changes the state of another electron (patient) in a scattering interaction. The input currents of a neuron are the signals that will be integrated to produce a suitable changing of the output synapses. A developer decides to build a new house and produces as signal the house's project, hold in the state of a suitable abstract propagator entity. Starting from this project the state of the building plot will change in a suitable amount of time, unless the municipal administration blocks the project.

In urban models, agents can be individuals acting in the urban space, patients can be lots of terrain, signals can be volumes and surfaces produced for different uses so that the state space of propagators is tied with the available surface and volume at disposal depending on the master plan (see Vancheri et al. 2008a and 2008b).

Frequently individuals are statistically represented in our models as populations. Intuitively speaking we can think a population $\mathcal{D}$ as a set of interacting entities, collectively described in statistical terms; its state is given by a probability distribution P and by the number N of agents belonging to the population. From a formal point of view it is possible to see populations as further interacting entities. This description enables us to calibrate conveniently the level of detail used to describe entities of the system. Usually, populations interact with other entities not directly, but through temporary interacting entities, which are probabilistically generated instances of individuals belonging to that population. The evolution of populations is a consequence of suitable *feedback interactions* by temporary entities, which statistically change the parameters of the probability distribution generating them. Populations contain only statistically relevant information needed to generate temporary entities that survive only for a limited time span before sending a feedback to the parent population. Using populations instead of permanent individuals offers some advantages with respect to ordinary MAS. First, it enables to deal with situations where the number of agents is very large and the detailed description of each agent is prohibitive from a computational point of view. Second, the interactions, involving only temporary individuals with a fully defined state, can be described with a degree of faithfulness comparable with MAS. Furthermore, temporary states have the advantage of being simpler than complete identities of agents in MAS, because they encompass only information relevant for the current interactions. Third, a suitable choice of the number of populations enables us to disregard irrelevant information in a systematic way, contributing to solve a typical problem in the construction of MAS. The number and the state of the populations may even change in time because of interactions: the centre of a population can shift because of feedback interactions. New populations can be generated when the change of the state of a temporary entity is such that the parent population does not recognize it any more. A population can split into two populations when the variance related to a feature of the population becomes too large. Two populations can coalesce when the moments of their probability distributions are at a distance small with respect to the dispersion of the features.

For a formal mathematical theory of IS, see Giordano et al. 2007.

5. Presentation of the BTG model using IS formalism

We would like to take the opportunity of this article to give also an intuitive but practically clear exposition of the concepts useful for constructing an IS, so that we shall present the BTG model using IS language.

5.1 Interacting entities

Interacting entities play in IS the same roles of agents in MAS. As mentioned above, the term “agent” is used in IS in a more technical way to denote a special kind of interacting entity. A suitable configuration space is associated to each interacting entity.

Examples of interacting entities from the BTG model are:

Commercial surfaces and other BTG displaced on the territory. Configuration space: amount of commercial surfaces and number of employees subdivided into 45 different categories considered as relevant for BTG related movements, spatial position of the corresponding building, and number of available parking places.

Links of the transportation network. Configuration space: geo-referenced position, speed limit, slope, number of lanes, a classification into 32 functional categories, capacity $q_{\max}$ of roads’ links (non-dynamical state variables). The description of the transportation network obviously must include one-way links, forbidden turnings, time penalties due to traffic lights, etc. In our model, all these data are exactly the same used by the cantonal administration for the implementation of the VISUM© model. Really, the only interesting dynamical state variable of a link is the number of vehicles $N^k_{\text{veh}}(l,t)$ passed on the link during each time window w_k (see the subsequent subsection *Routing*) at time t. From this state variable we obtain the input flux $q_k(l,t) = \frac{N^k_{\text{veh}}(l,t)}{|w_k|}$

where $|w_k|$ is the amplitude of the time window w_k.

Paths: generally speaking, to connect each pair of zones z, z' each agent i can choose in a set $\mathrm{Paths}(z, z')$ of paths starting in the zone z and ending in the zone z'. The set $\mathrm{Paths}(z, z')$ can be defined in different ways with respect to the degree of faithfulness we want the models respects, or the computational resources one has. E.g. it can be static and contain only the quickest path with respect to a hypothetical situation with empty roads, or it can contain the quickest path with respect to the traffic experienced in recent past $[t - \Delta t, t)$ (this requires non trivial computational resources), or finally it can contains other detours between the two given zones (see Shibuya et al. 1995 for a possible precise definition of detour). A dynamical switching between different detours from z to z' can also be dynamical and based on the past traffic experience (i.e. duration time) that, over a certain threshold, can increase the probability to choose for a different detour. This dynamical exploration of the possible paths from z to z' requires less computational resources with respect to the searching of the quickest path given a duration time-based penalty matrix. In the present version of the model in $\mathrm{Paths}(z, z')$ we have only the quickest path in situation of unloaded roads (maximum speed given by the highway code).

Each path $p \in \mathrm{Paths}(z, z')$ is an interacting entity whose state is given, for each time window, by the run time necessary to cover that path. These entities enter in the neighbourhood of several decision processes of agents.

Populations of individuals residing in a given zone of the system and sharing statistically similar socio-economical features. Configuration space: spatial coordinates of the polygon enclosing the population, data describing at a statistical level the socio-economical status of the individuals of the population and parameters describing goals and constraints driving the mean behaviour of the individuals belonging to the population. In the present version of our model we considered the following population related goals that will be achieved using a stochastic controller (see section 5.3):

1. average number of movements $\bar{n}_k \in \mathbb{N}$ for each given type τ_k of movement in a week, $k = 1, \ldots, \chi$;
2. preferred time window $\omega_k \in \{1, \ldots, n_{\mathrm{tw}}\}$ for each given type τ_k of movement in a week,
3. average time $\bar{t}_k > 0$ spent in a given type of movement τ_k in a week;
4. average money $\bar{m}_k > 0$ spent in a given type of movement τ_k in a week.

The socio-economical status is expressed giving:

1. the coefficients of expense $\sigma_g^j \in [0, 1]$ (see section 3.2, item 4 in the definition of CMP) for each goal g in the subset of goals $I_j^k \subseteq G$ present in the type of movement $\tau_k = (\{I_j^k\}_{j=1}^{m_k}, c_k)$;
2. the intervals $[m_g^j, M_g^j]$ representing the range of money spent by the considered socio-economical class in the accomplishment of the goal g in the subset of goals $I_j^k \subseteq G$;

3. a number $s_w \in [0,1]$, called *suitability*, representing the fuzzy truth value of the sentence "for the socio-economical class c it is good to accomplish some BTG related goal in the time window w".

Complex mobility patterns $m = (k, \omega, \{I_j\}_{j=1}^n, c)$ are themselves interacting entities, and the configuration space is the set of all possible values for kernels $k = z_0 p_{01} z_1 p_{12} z_2 \dots z_m p_{m0} z_0$, index identifying a time window $\omega \in \{1, \dots, n_{\mathrm{tw}}\}$, subsets of goals $\{I_j\}_{j=1}^n$ and characteristics of movement $c = (\alpha, i_1, \dots, i_n, \mu, \sigma)$.

Temporary moving entities (TME), i.e. members of a population and currently involved in a trip chain. The configuration space is given by: specific residential location, the socio-economical status of the entity and a pointer to a CMP.

5.2 Interactions

The next step to define an IS is to list a finite set of interactions. We can think an interaction as a causally directed process between agents and patients mediated by propagator entities. The general form of an interaction is hence:

$$a_1, \dots, a_n \text{ HAVE AN INTERACTION } \alpha \text{ WITH } p_1, \dots, p_m \text{ THROUGH } r_1, \dots, r_p$$

Where $a_1 \dots, a_n$ are called *agents*, $p_1, \dots, p_m$ are called *patients*, $r_1, \dots, r_p$ are called *propagators* and "HAVE AN INTERACTION α" is a linguistic description of the interaction. Formally the label α is associated with every mathematical objects describing the properties of the interaction. If we represent an interaction by means of a graph, like in the following figures 1 and 2, and connect two graphs when they share an entity, we obtain a network representing causal flows in the system.

Figure 1: Grafical representation of a single interaction in which two agents (red) modify the state of a patient (blue) trough a propagator (grey). One of the agents is modified by the interaction and hence plays simultaneously the role of patient.

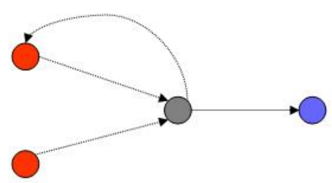
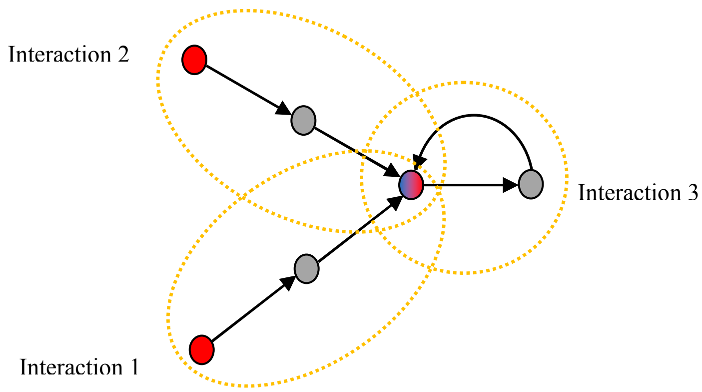


Figure 2: A network of three interactions corresponding to a decision process. Two objects (red) send (interaction 1 and 2) information to an individual (red/blue) and subsequently the individual chooses an object using collected information (interaction 3). The decision process is a self-interaction of the individual with the information collected in its working memory.



The interactions of the BTG model are:

Generation of temporary moving entities: a population $\mathcal{D}$ (agent), representing the households residing in a given zone, generates a TME i (patient) assigning to it a state. If $P_{\mathcal{D}}$ is the probability associated to $\mathcal{D}$ (see section 4) and describing the socio-economical features of the population (in our model these are: the coefficients of expense σ_g^j, the range of money for BTG-related expenses $[m_g^j, M_g^j]$ and the coefficients of suitability s_w, see section 5.1), then the state of i is extracted using $P_{\mathcal{D}}$. E.g. in this state we can consider the average amount of money spent by i for shopping. This is an example of interaction where the propagator is an abstract entity whose configuration space is the set of all possible states of temporary entities generated by $\mathcal{D}$, i.e. the support space of the probability $P_{\mathcal{D}}$. The signal (see item 12, section 4) from the propagator to the patient i is the specific initial state of the temporary entity i. This interaction is also responsible of the generation of types of movement $\tau_k = (\{I_j^k\}_{j=1}^{m_k}, c_k)$ (see section 5) for each TME i. To understand what can be a sufficiently correct probability distribution to generate the number of individuals that in a week accomplish a trip of type τ_k, let $N_{\mathcal{D}}$ be the number of individuals in the population $\mathcal{D}$ and let M_i be the random variable counting the number of movements for some type of movement τ_j of the individual $i \in \mathcal{D}$ in a week. Obviously, it is not possible to faithfully model the distribution of each different M_i, but we always have $N_{\mathcal{D}}$ at least of the order of hundreds of individuals, hence we can approximate the total number of movements $\sum_{i\in\mathcal{D}} M_i$ with a normal distribution $N(\mu \cdot N_{\mathcal{D}}, \sigma)$. The average number μ of movements (related to BTG, i.e. of some type τ_j) per individuals and per week, and the standard deviation σ can be obtain from statistical studies about traffic behaviour of the population $\mathcal{D}$. In the present version of the model, on the basis of the statistical data owned by the cantonal administration, we have chosen[4] $\mu = 5\,\frac{\text{mov}}{\text{week}}$ and $\frac{\sigma}{\mu} = 5\%$. Of course, due to the mentioned approximation of $\sum_{i\in\mathcal{D}} M_i$ with a normal distribution, we are forced to cut off the negative tail of $N(\mu\cdot N_{\mathcal{D}},\sigma)$ and, because we have to respect the meaning of total number of BTG related movements, we also have to consider the integer part of this $N(\mu\cdot N_{\mathcal{D}},\sigma)$. Using this distribution we can extract the total number of BTG related movements in a week, let $m_{\mathcal{D}}\in\mathbb{N}$ be this number. Considering the probabilities

$$p_k^{\mathcal{D}} := \frac{\bar{n}_k}{\sum_{j=1}^{\chi}\bar{n}_j}$$

we can use a multinomial distribution with parameters $(m_{\mathcal{D}}, p_1^{\mathcal{D}},\dots,p_\chi^{\mathcal{D}})$ to extract the number $n_k^{\mathcal{D}}\in\{0,\dots,m_{\mathcal{D}}\}$ of TME in the population $\mathcal{D}$ that realize, in some time window of the considered week, the type of movement $\tau_k = (\{I_j^k\}_{j=1}^{m_k}, c_k)$.

At the end of this interaction in the state of the TME i there are a type of movement $\tau_k = (\{I_j^k\}_{j=1}^{m_k}, c_k)$ and her/his socio-economical data. Now the TME starts to associate to the type of movement τ_k a kernel $k = z_0 p_{01} z_1 \dots z_m p_{m0} z_0$.

Selection of a sequence of zones and a time window by a TME: after its creation, and association with a type of movement τ_k, a TME (patient) collects information (signals) coming from BTG (agents). This "information" is concretely given by fuzzy logic based measures of the attractiveness $a(z_1,\dots,z_m,\omega\,|\,t,\tau_k)\in[0,1]$ of the BTG for different kind of goals of the TME and to environmental features like the congestion of the commercial surfaces (depending on the time window ω of the past week $[t-\Delta t, t)$; see section 6 for more details about the definition of the attractiveness indicators). As a patient entity, the TME uses this fuzzy attractiveness indicators sent by BTG and filters among them only some $(m+1)$-tuple[5] $(z_1,\dots,z_m,\omega)$ of zones with a probability proportional to the attractiveness indicators $a(z_1,\dots,z_m,\omega\,|\,t,\tau_k)$. Note that of course more than one BTG can be attractive for a given type of movement $\tau_k = (\{I_j^k\}_{j=1}^{m_k}, c_k)$. Case studies with a high number of BTG may see a huge number of possible kernels $k = z_0 p_{01} z_1 p_{12} z_2 \dots z_m p_{m0} z_0$, so that it can be impossible to manage them from a computational point of view. Analogous problems are possible in case of a high number of exogenous events (i.e. insertion during a simulation of new BTG), because

---

[4] Recall that in our types of movements we include also "local" trips of short duration.

[5] We can suppose that only those with $m \le 3$ are statistically relevant.

each new insertion will influence the attractiveness indicators $a(z_1, \ldots, z_m, \omega \,|\, t, \tau_k)$ and hence the extraction of visited zones z_j carrying these new BTG $b_j \in z_j$. In this case, we can think to use an evolutionary inspired algorithm to manage all the possible CMP. These algorithm will provide a source of new CMP. CMP that will be selected with a low frequency by populations will be discarded by a specific annihilating interaction.

Routing: All the TME going from z to z' select a path $p \in \mathrm{Paths}(z, z')$ with a probability

[Automatic section break]

$$P(p) := \frac{\mu(\mathrm{time}_p(t - \Delta t, w))}{\sum\limits_{q \in \mathrm{Paths}(z,z')} \mu(\mathrm{time}_q(t - \Delta t, w))}$$

where $\mathrm{time}_p(t - \Delta t, w)$ is the time necessary to go from the zone z to the zone z', along the path p, with respect to the traffic experienced in the time window w of the past week $[t - \Delta t, t)$. The function $\mu : [0, +\infty) \longrightarrow [0, 1]$ is a decreasing fuzzy membership function expressing the idea that agents i will select paths with similar high probability if the duration $\mathrm{time}_p(t - \Delta t, w)$ is under a certain threshold, with a decreasing probability for higher duration and with lower probability if this time is over a certain saturation value. Once we have extracted the path p for each pair of zones, we have completed the CMP associated with the TME i. These interactions have really the same formal structure than the previous two interactions for the selection of BTG, where the TME (patient) firstly selects the sets of paths $\mathrm{Paths}(z, z')$ interacting with roads, and secondly selects single paths with the above-mentioned probability.

Updating of the variable counting the number of vehicles: each link l of each path p_{jk} connecting the zone z_j to the zone z_k in the kernel $k = z_0 p_{01} z_1 \ldots z_m p_{m0} z_0$ interact with all the temporary entities that cover that link and increase the number of vehicles $N^k_{\mathrm{veh}}(l, t)$ present in the given link l during the time window w_k at time t. Of course, this counting variable is the core for the evaluation of several interesting observables.

Updating of the paths: The present implementation of the model is not based on micro-traffic simulations, and the duration of the run of a link has been estimated using the capacity restraint function used in the VISUM© model, i.e.

$$\mathrm{time}_p(t, w_k) = \sum_{l \in p} t_0^l \cdot \left\{ 1 + a_l \cdot \left[\frac{q^k(l,t)}{q^l_{\max}} \right]^{b_l} \right\} \tag{5.2.1}$$

where a_l and b_l are parameters depending on the type of link $l \in p$, t_0^l is the run time of the given link in unloaded conditions, $q^k(l,t) = \frac{N^k_{\mathrm{veh}}(l,t)}{|w_k|}$ is the flux inputted in the link in the time window w_k and $q^l_{\max}$ is the capacity of the link. Using (5.2.1) we can update the state variable $\mathrm{time}_p(t, w_k)$ of the paths in $\mathrm{Paths}(z, z')$ for each pair of zones.

Systematic movements: each link of the roads network is also loaded with a certain number of movements not related to BTG, like home-work or home-school. The counting of these movements has been taken equal to the one determined by the VISEM© model. The routing of these origin-destination matrices follows the idea already explained above. Also coming from the VISEM© model, is a background of movements of heavy vehicles; this background is held constant during all the simulations, even if this is not correct in case of insertion of new BTG. We plan to solve this defect in a future version of the model.

5.3 A stochastic controller to achieve populations' goals

[Automatic section break]*BTG related activities*: a TME achieves the schema of movement through a sequence of interactions (e.g. shopping or free-time related activities) in which makes use of the selected BTG and roads. These interactions are described by a set of quantities (the so-called signals) like the amount of spent money and time and the amount of benefit and punishment derived by these activities.

Feedback of the TME to the originating populations: after having accomplished the trip chain, the TME (agent) i sends to the original population $\mathcal{D}$ (patient) the extracted signals and is destroyed (recall that a TME has a lifetime limited to a trip chain). The population $\mathcal{D}$ adds all the signals extracted by all the TMEs in a given span of time obtaining a total value of benefit and punishment.

Optimization of a strategy by a population: using information collected in the previous feedback interactions a population $\mathcal{D}$ modifies the parameters of the probability distribution used for extracting schemata of movement, i.e. it modifies the frequencies with which the TME realize different types of movements. This is a self-interaction where $\mathcal{D}$ is both agent and patient. This interaction has been realized implementing a stochastic controller that tries to minimize, resp. maximize, the functions of total punishment and benefit of the population. An analytical formula for the gradient of these functions has been obtained so that the gradient method has been chosen in this optimization process. Intuitively, the idea of the stochastic controller is, e.g., that if an agent lives in a zone far from shopping centres, then on average she/he tries to decrease the number of big shopping and to increase the amount of spent money. We can hence say that the stochastic controller start with an average value of the frequencies of different types of movements and adapts these frequencies to the particular zone where the considered population is resident. The adaptation is performed optimizing the benefit and punishment functions.

The general idea behind the previously described interactions is that each population is an interacting entity that attempts to maximize goals under constraints. In order to achieve its global goals, a population produces TME that are driven along a trip chains by CMP and collect signals that are eventually sent back to the population. The selected CMP determine the performances of TME. A population can improve the attainments of goals optimizing the choice of CMP.

In general, decision processes where an individual selects an object in a set of possibilities are decomposed in our approach into two steps (interactions):

1. flows of information (signals) sent by the objects (agents) reach the individual (patient) modifying its knowledge. Since many objects are sending signals at the same time, the individual will be modified with a higher probability by objects sending a signal that match the internal requirements of the agent itself (in other words the agent is able in general to filter information and direct its attention);
2. using the collected knowledge the individual (agent) selects an object.

This decision schema can reproduce a multinomial logit model at the most trivial level of implementation, but is suited for a wide range of generalizations. The first phase can include indeed a chain of interactions with objects that progressively modify the filters applied by the individual (and hence the way it directs attention) and its final decision criteria. In this way, the decision process is turned in a stochastic process of partial exploration of the decision space where accessibility to information, recent experiences and all kind of bias play a role as relevant as pure utility. Typical phenomena like sensitivity and habituation induced by recent detected signals could bias the decision process in a critical way. This could lead to decisional dynamics more realistic than in usual utility-based approaches. Properly speaking we have a decision process only in the phase 2. This decision process involves only a limited amount of information concerning a small number of objects and stored in the working memory of the individual.

6. Attractiveness indicators

As mentioned above, probability distributions selecting kernels of movement, are based upon indicators associated to zones.

6.1 Attractiveness indicators for a single zone and a single goal:

A first set of indicators measures the attractiveness of each zone z for each goal $g \in G$. This group of indicators is based on *measures of presence* $\pi(a \mid g) \in \mathbb{R}_{>0}$ of each kind of activity

$a \in z$ in the zone z at time t. Recall that during the simulation we can exogenously insert new activities in some zones, so that we will underline this dependence on time using the notation $a \in z(t)$, with the meaning of "the economical activity a is present in the zone z at time t". The presence has been measured either by the commercial surface into the zone or by the number of employees or finally by the number of firms. We can hence define the *offer* corresponding to the presence $\pi(a \mid g)$ as an increasing function $\Delta_a^{\text{off}} : \mathbb{R}_{>0} \longrightarrow [0, M_a]$ which is directly proportional to the presence in case of low values of presence, $\Delta_a^{\text{off}}(\pi) \simeq \alpha \cdot \pi$ if $\pi < \pi_{\text{thr}}$, and saturates to a maximum value M_a of offer for high value of presence, $\Delta_a^{\text{off}}(\pi) > \beta \cdot M_a$ if $\pi > \pi_{\text{sat}}$. We can also include subtle effect in the offer function Δ_a including an inflexion point representing a value of presence around which we have a strong increasing of the perceived offer. Of course, this way to proceed introduces several parameters that has to be calibrated. See section 7 to see the calibration methods we used. The offers of all the activities are added up obtaining $\sum_{a \in z(t)} \Delta_a^{\text{off}} [\pi(a \mid g)]$, called the total offer of the given zone, which is the input variable of a fuzzy membership increasing function $\mu_{\text{off}} : \mathbb{R}_{>0} \longrightarrow [0,1]$ measuring the attractiveness of a zone for a given goal g. Finally, the term $\mu_{\text{off}} \left(\sum_{a \in z(t)} \Delta_a^{\text{off}}(\pi(a \mid g)) \right)$ has to be modified with a fuzzy modifier of type VERY in case the agent that has to evaluate the zone z intends to do a great shopping expense. Indeed, the idea is that she/he cannot do this type of shopping if the attractiveness is not sufficiently high (and hence the offer is sufficiently high). We can turn this idea in mathematical terms using a power $b_g(\sigma_g) \in [1, +\infty)$ depending on the coefficient of expense σ_g. For example, we can define $b_g(\sigma_g) := 2$ if $\sigma_g > 0.75$ (great expense) and $b_g(\sigma_g) := 1$ for the lesser value of σ_g. Therefore, the definition of this part of the attractiveness indicator is:

[Automatic section break]
$$a_{\text{off}}(z \mid t, g, \sigma_g) := \mu_{\text{off}} \left(\sum_{a \in z(t)} \Delta_a^{\text{off}}(\pi(a \mid g)) \right)^{b_g(\sigma_g)} \tag{6.1}$$

The function μ_{off} permits to include threshold and saturation effects, from a decisional point of view, for the value of the overall measure of presence.

A second group of indicators, called *contextual indicators* detect further information that is not directly related to the supply for each goal but that can turn to be relevant during a movement. These indicators measure:

- the availability $a_{\text{ser}}(z \mid t, g) \in [0,1]$ (where the subscript "ser" recall the word "services") of some facilities and services $s \in z$ (like petrol stations, bank branches, cash dispensers, etc. We can use an idea similar to the one present above, but using the number of services as a measure of presence:

$$a_{\mathrm{ser}}(z \,|\, t, g) := \mu_{\mathrm{ser}} \left(\sum_{s \in z(t)} \Delta_s \left(\mathrm{card}\{s \,|\, s \text{ is a service useful for } g\} \right) \right)$$

where μ_{ser} is another increasing fuzzy membership function. If one cannot consider the (exogenous or endogenous) creation of new services, then the availability $a_{\mathrm{ser}}(z \,|\, t, g)$ does not depend on time.

– the expected density of visitors in each time window is estimated considering two components: the first one is the number of visitors $N_{\mathrm{vis}}(z, w \,|\, t)$ in the zone z per unit of surface and unit of time that have been present during the time window w of the present week:

$$n(z, w \,|\, t) := \frac{N_{\mathrm{vis}}(z, w \,|\, t)}{S(z,t) \cdot |w|},$$

where $S(z,t) := \sum_{a \in z(t)} S_a$ is the total commercial surface available in the zone. The second is, intuitively, a memory that each agent has about the congestion experienced in the past in that zone. More formally, this is turned into mathematical terms considering a memory variable having the following recursive dynamics:

$$\theta(z, w \,|\, 0) := n(z, w \,|\, 0)$$
$$\theta(z, w \,|\, t + 1\mathrm{week}) := \gamma \cdot \theta(z, w \,|\, t) + (1 - \gamma) \cdot n(z, w \,|\, t + 1\mathrm{week})$$

The parameter $\gamma \in [0,1]$ represents the inertia of this memory variable: for high values of γ only the old value $\theta(z, w \,|\, t)$ is important; on the contrary for low values of γ only the new value $n(z, w \,|\, t + 1\mathrm{week})$ gives the more important contribution. High values of $\theta(z, w \,|\, t)$ must decrease the attractiveness of the zone, so we can consider an increasing fuzzy member function μ_{cong} to evaluate using a fuzzy negation $a_{\mathrm{cong}}(z, w \,|, t) := 1 - \mu_{\mathrm{cong}}(\theta(z, w \,|\, t))$ to measure this negative component in the attractiveness of the zone.

– The *urban context* $a_{\mathrm{urb}}(z \,|\, t, g, \sigma_g) \in [0,1]$ is another contextual indicator, measuring the fact that in a central urban zone (in Ticino, or in a shopping centre that tries to imitate such a type of central urban context) it is easier to find small highly specialized shops. Therefore, a small number of them is able to arrive to a sufficient level of offer (corresponding to a sufficient number of agents attracted in that zone for the goal g) , but we need a higher value of total offer to arrive at a high value of attractiveness (corresponding to a high number of agents attracted in that zone for the same goal). For these intuitive reasons, we defined the urban context $a_{\mathrm{urb}}(z \,|\, t, g, \sigma_g)$ using a formula similar to (6.1)

$$a_{\text{urb}}(z \mid t, g, \sigma_g) := \mu_{\text{urb}} \left(\sum_{a \in z(t)} \Delta_a^{\text{urb}}(\pi(a \mid g)) \right)^{b_g(\sigma_g)}$$

but with a lower value of threshold π_{thr} for the corresponding offer functions Δ_a^{urb}, and a higher value of saturation for the membership function μ_{urb}. Note that this indicator does not aim at measure the real urban central nature of the zone (as it is possible to do using other indicators, see Vancheri et al. 2008b), but to measure the attractiveness of the zone z as if it was in a urban central context. Indeed, this indicator assumes very high value in central zone e.g. of Lugano, but a sufficiently high value also in zone corresponding to very big shopping centres, like the Parco Commerciale Grancia.

We can hence define the contextual indicator as:

$$a_{\text{context}}(z, w \mid t, g) := a_{\text{ser}}(z \mid t, g) \cdot a_{\text{cong}}(z, w \mid t) \cdot a_{\text{urb}}(z \mid t, g),$$

expressing that to have a high value of this indicator we must have a high value of presence of service $a_{\text{ser}}(z \mid t, g)$ and a low evaluation of congestion $a_{\text{con}}(z, w \mid t)$ and a high value of urban context $a_{\text{urb}}(z \mid t, g)$.

Finally, the attractiveness indicator for a single zone and a single goal is defined as

$$a(z, w \mid t, g, \sigma_g) := a_{\text{off}}(z \mid t, g, \sigma_g) \cdot a_{\text{context}}(z, w \mid t, g)$$

corresponding to a high value of offer and of the contextual part.

6.2 Attractiveness indicator for a single zone and a subset of goals

Recall that our purpose is to define meaningful attractiveness indicators to evaluate when a certain kernel $k = z_0 p_{01} z_1 p_{12} z_2 \ldots z_m p_{m0} z_0$ is attractive for a SCMP $(\{I_j\}_{j=1}^n, c)$ in a certain window ω. This will permits to select the best zones and paths to accomplish this SCMP, that is this type of movement. Hence, let us now consider a CMP $m = (k, \omega, \{I_j\}_{j=1}^n, c)$, with characteristics of movement $c = (\alpha, i_1, \ldots, i_n, \mu, \sigma)$. As we already have mentioned in section 3, the diffuse movement

$$\sum_{g \in I_j} \left(\alpha_g^j \cdot g \text{ SPENDING } \sigma_g^j \right)$$

has to be thought as accomplished in a single zone z. Hence, it is natural to associate to this movement the indicator

$$a\,(z, \omega \,|\, t, I_j, c) := \sum_{g \in I_j} \alpha_g^j \cdot a(z, \omega \,|\, t, g, \sigma_g^j).$$

6.3 Attractiveness indicator for a sequence of zones and a type of movement

In this last step we have to measure how much a whole sequence of zones $z = (z_1, \ldots, z_m)$, instead of a single zone, is attractive to accomplish all the subsets of goals in a given SCMP $\tau = (\{I_j\}_{j=1}^n, c)$. In the present version of the model we have separated the choice of a sequence of zones to accomplish τ_k, from the choice of the paths connecting each pair of subsequent zones[6]. For this reason, we start evaluating the zones $z = (z_1, \ldots, z_m)$ with respect to the shortest paths connecting subsequent zones. Let $\mathrm{best}(z) = \{p_{01}, \ldots, p_{m0}\}$ be the set of these shortest paths. The evaluation of $z = (z_1, \ldots, z_m)$ from the point of view of the best time that can be spent passing from one zone to the next is given by

$$\mu\,(\mathrm{time}_z(t, \omega)) := s_\omega \cdot \mu \left(\sum_{p \in \mathrm{best}(z)} \mathrm{time}_p(t, \omega) \right)$$

where μ is the evaluation of time contained in the characteristic of movement $c = (\alpha, i_1, \ldots, i_n, \mu, \sigma)$ (see item 3, section 3.2) and s_ω is the suitability of the time window ω associated to the socio-economical class of the agent that, intuitively, is evaluating the kernel. Let us note that a different permutation of the sequence of zones $z = (z_1, \ldots, z_m)$ can have a different evaluation of time, so that the sequence with the highest value of this evaluation will also be the best among all the permutations.

To evaluate the attractiveness of a sequence a zones for the movement

$$\mathrm{IN}\ k \quad \mathrm{AND} \quad \omega \quad \mathrm{AND} \quad \mu \quad \mathrm{AND} \quad \bigwedge_{j=1}^{n} \left(\sum_{g \in I_j} \alpha_g^j \cdot g \quad \mathrm{SPENDING}\ \sigma_g^j \right)^{[i_j]}$$

[6] Another possibility is to extract simultaneously, with a suitable probability, the sequence of zones and the paths connecting them.

the first idea is to ask a fuzzy formula of type: "$z = (z_1, \dots, z_m)$ is attractive for the type $\tau = (\{I_j\}_{j=1}^{n}, c)$ if and only if for every subset of goals I_j there exists a zone z_p which is attractive for that subset, i.e. for which the value of $a\,(z_p, \omega \,|\, t, I_j, c)$ is high". In symbols[7]:

$$a_1(z, \omega \,|\, t, \tau) \;:\Longleftrightarrow\; \forall j = 1, \dots, n \;\exists p = 1, \dots, m : \; a\,(z_p, \omega \,|\, t, I_j, c). \qquad (6.2)$$

This formula has the problem that is not sufficiently "intelligent" because is not able to avoid sequence where some zone z_k is useless, because is not attractive for any subset of goals. For this reason we will add to the formula (6.2) the condition of being without useless zones, i.e.

$$\mathrm{NoUs}(z, \omega \,|\, t, \tau) \;:\Longleftrightarrow\; \neg\exists p = 1, \dots, m \;:\; \forall j = 1, \dots, n \;:\; \neg a(z_p, \omega \,|\, t, I_j, c).$$

Finally, the formula (6.2) is not able to see if in the sequence there is a high number of abounding zones, that is pair of zones that are both attractive for the same subset of goals I_j. We explicitly admit the possibility that a certain I_j can be accomplished in a *low* number of different zones, but it does not seem to correspond to a realistic behaviour if this number is high. For this reason we will also add the last condition that the number of abounding zones is not high. A zone is abounding if there exists another zone which is attractive for the same I_j:

$$\mathrm{LowAb}(z, \omega \,|\, t, \tau) \;:\Longleftrightarrow\; \mu_{\text{low}}\left[\mathrm{Card}\,\{p \,|\, \exists j \,\exists q : \; p \neq q \,,\, a(z_p, \omega \,|\, t, I_j, c) \,,\, a(z_q, \omega \,|\, t, I_j, c)\}\right],$$

where $\mu_{\text{low}} : \mathbb{N} \longrightarrow [0, 1]$ is a decreasing fuzzy membership function to express the fuzzy notion of "low". At the end we can define the final indicator:

$$a(z, \omega \,|\, t, \tau) := a_1(z, \omega \,|\, t, \tau) \cdot \mathrm{NoUs}(z, \omega \,|\, t, \tau) \cdot \mathrm{LowAb}(z, \omega \,|\, t, \tau) \qquad (6.3)$$

Of course, we can affirm that the construction of this indicator was not easy and direct as in random utility theory, but at the same time we can also assert that it is able to include subtle non-linear effects. Moreover, every parameter that has to be calibrated has always a clear decisional meaning, so that frequently a pre-calibration value can be easily assigned. Finally, in the subsequent section we will see that the use of this indicator results in a great computational performance in the selection of the sequence of goals for BTG related activities.

[7] Recall that in our finite context the quantifiers correspond to finite conjunctions and disjunctions, that is to finite products and sums because $(x \text{ AND } y) := x \cdot y$ and $(x \text{ OR } y) := x + y - x \cdot y$.

6.4 Computationally efficient location of secondary activities

Using the indicator (6.3) we can implement an efficient algorithm for the construction of all the attractive sequence of zones:

1. Start with single zones. Filter the set of all the zones z (and time window ω) considering only those that are sufficiently attractive for some subset of goals, that with respect to the indicator $a(z, \omega \mid t, I_j, c)$. This first step usually select few zones, e.g. discarding all the zones that contain no BTG.

2. Among all the possible pairs of zones coming from the previous selection, filter the pair of zones that do not contain useless zones using a sufficiently high value of the indicator $\mathrm{NoUs}(z_1, z_2, \omega \mid t, \tau)$.

3. Among all the pairs filtered in the previous step, construct all the terns with a zone coming from step 1. Select from these terns only those with a low number of abounding zones (i.e. at maximum 2 abounding zones): $\mathrm{LowAb}(z_1, z_2, z_3, \omega \mid t, \tau)$.

Recall that we can neglect kernel with more than three zones because they are statistically non relevant.

6.5 Limitations of the present implementation of the model

We have frequently mentioned that the model presented in this paper is in a preliminary version. Nevertheless, we think that the corresponding simulations are anyhow interesting and meaningful. Moreover, as we have seen above, the mathematical theory of the model has been already completely developed. More precisely, the present version of the model has the following limitations:

- We have kernel with only one zone. The corresponding software has already been written and tested, but this section of the model has not been completely calibrated and validated, due to a lack of time.

- The stochastic controller has not been implemented.

- We consider only the shortest path, find with unloaded roads, between two zones.

- We only consider an average socio-economical class. This can be a problem if there are meaningful unhomogeneity in the territorial distribution of people among the different zones.

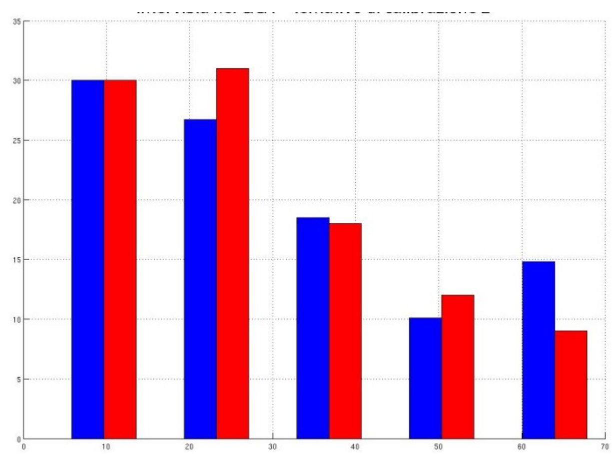
7. Calibration and validation

As we have already mentioned the use of fuzzy logic permits to obtain parameters that always have a decisional meaning, and this permits to fix a reasonable value for the majority of the parameters directly from statistical data without any optimization process.

Another method of calibration we have used for the parameters of the attractiveness indicator has been to calibrate the parameters so as to obtain exactly the classification of importance and number of clients of each BTG, coming from the data of the cantonal administration. In other words, after the calibration process, the zone with the best k-th value of the attractiveness is the best k-th zone with respect to the classification of the cantonal administration.

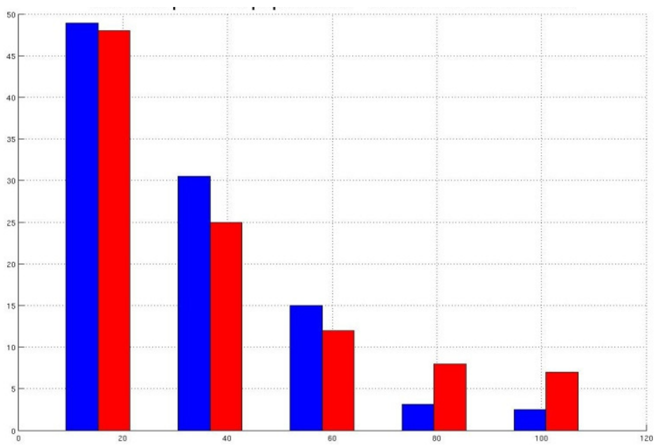
Finally, we have used the survey Alberton and Guerra 2008 to calibrate other parameters. This survey has been conducted with two different populations: the first one sampling the answers of a questionnaire from the clients of six among the biggest shopping centres of Canton Ticino, and the second one sampling the answers among a random choice of inhabitants of Ticino, stratified by residence town. The following is the comparison of the survey data and the simulated data about the number of clients in the first population that cover a certain class (among 5) of distance after calibration:

Figure 3: Comparison of the survey data (blue) and the simulated data (red) about the number of clients in the first population that cover a certain class (among 5) of distance after calibration.



After this calibration process we test the same quantity but using the second population, obtaining the following validation:

Figure 4: Comparison of the survey data (blue) and the simulated data (red) about the number of clients in the second population that cover a certain class (among 5) of distance.

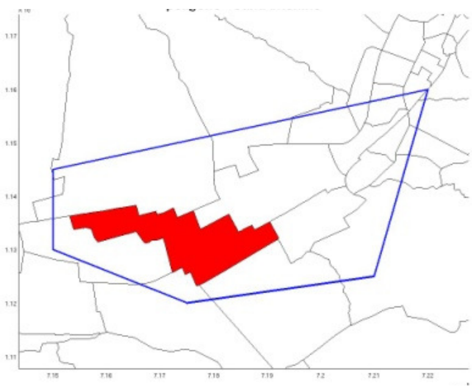


The average error is about 20%. We are now working to tempt a validation using the real counting data coming from the counters of the cantonal administration.

8. Simulations and conclusions

Preliminarily to the estimation of a given environmental indicator, one has to fix a zone of observation, typically including the zone interested by an exogenous intervention, e.g. the construction of a new shopping centre

Figure 5: A polygon inscribing a zone of observation for the estimation of environmental observables.



The model permits several types of indicators, among which we can consider those represented in the following figures:

Figure 6: Global inflow of NO_2 in a given observation zone. Before the graphical representation, the links contained in the polygon has been sorted with respect to this inflow. In red we represent the link that had an increasing of NO_2 and in blue those that had a decreasing of NO_2.

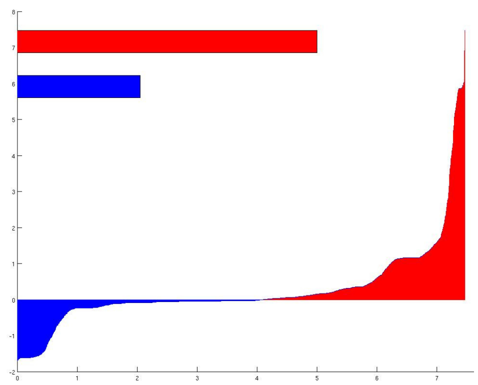
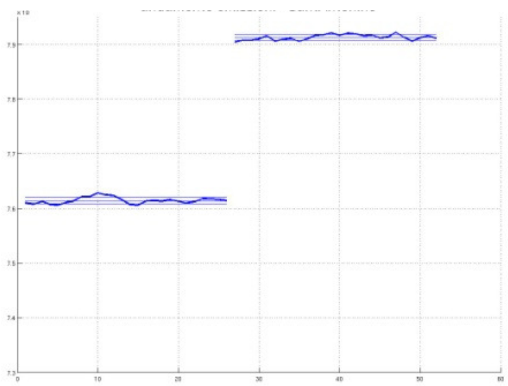


Figure 7: Inflow of NO_2 in a single link before and after an exogenous intervention.



The following are three plots about the representation of the pointwise representation of the inflow of NO_2 in the already mentioned simulation of uniform filling of all the divisions destined to accept BTG.

Figure 8: Pointwise representation of the inflow of NO2. In hot colours the points with and increasing of pollutant. In this particular area we have a preponderance of points with an increasing of pollutant.

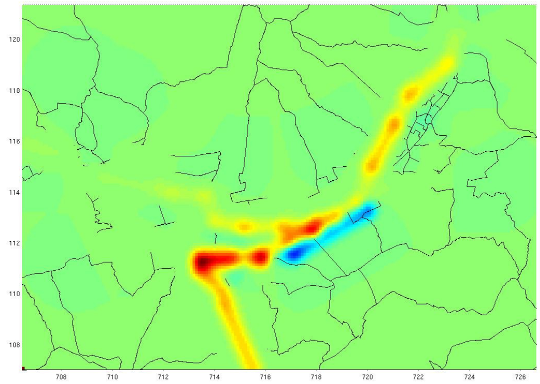


Figure 9: Pointwise representation of the inflow of NO2. In hot colours the points with and increasing of pollutant. In this particular area we have a preponderance of points with a decreasing of pollutant.

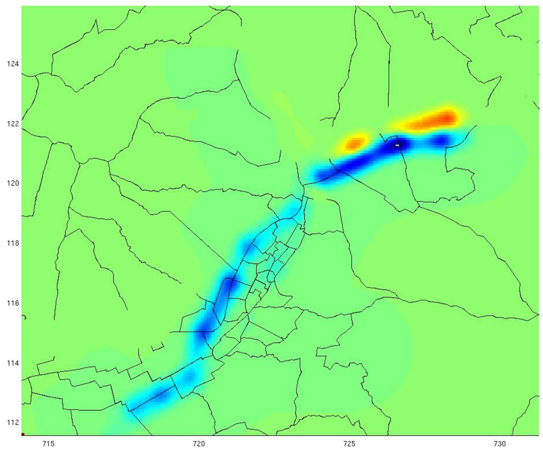
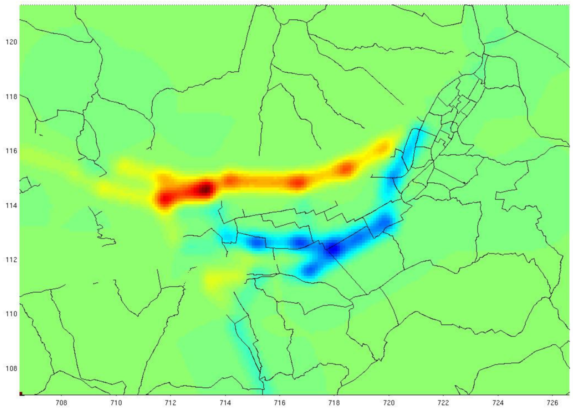


Figure 10: Pointwise representation of the inflow of NO2. In hot colours the points with and increasing of pollutant. In this particular area we have comparable areas of increasing and decreasing of pollutant.



These simulations shows that the classical planning method of fixing a certain number of divisions enabled to hold BTG can produce unexpected results. It is not possible to think that the hypothetical situation with all these divisions full of activities is the worst one, nor that it is the better one because the supply of retail activities is so diffuse to decrease the number of trips. In our opinion this shows that new planning ideas must be introduced to face the problem of environmental impact and economical best allocation of BTG, and a deeper use of these types of mathematical models can represent the core idea to introduce these innovative planning methods.

9. Acknowledgements

This research has been partially supported SNSF (contract number 200021-112420/1). We acknowledge the "Gruppo di lavoro grandi generatori di traffico" of the Cantonal administration for their help during the development of the BTG model. The data used in the simulations belong to the Cantonal administration, the Federal Institute of Statistics and the Cantonal Institute of Statistics.

10. References

Alberton, S. and Guerra, G. (2008), Il comportamento dei consumatori in materia di mobilità nei principali centri commerciali del Canton ticino. Report of the survey conducted by Instituto di Ricerche Economiche (IRE, Lugano) on behalf of Distributori Ticinesi (DISTI).

Arentze, T. and Timmermans, H. (2000) ALBATROSS: a Learning-based Transportation-oriented Simulation System, *ElRASS*, Eindhoven University of Technology, The Netherlands.

Arentze, T. and Timmermans, H. (2007), A Multi-Agent Activity-Based Model of Facility Location Choice and Use, disP 170, 3, pages 33–44.

Balmer, M., Cetin, N., Nagel, K., and Raney, B. (2004), Towards truly agent-based traffic and mobility simulations, AAMAS'04, July 19-23, 2004, New York, USA.

Balmer, M. (2007) TRAVEL Demand Modeling For Multi-Agent Transport Simulations: Algorithms And Systems, PhD thesis, ETH Zurich, ETH17238.

Bürgle, M., Löchl, M., Waldner U. and Axhausen, K.W. (2005) Land use and transport simulation: Applying UrbanSim in the Greater Zürich area, paper presented at *CUPUM*, London, June 2005.

Casti, J. (1997), Would-be worlds: toward a theory of complex systems, *Artificial Life and Robotics*, Volume 1, Number 1 / March.

Giordano, P. and Vancheri, A. (2007) Interaction spaces: an axiomatic theory for complex systems of interacting objects, working paper in www.mate.arch.unisi.ch/ACME/IS.pdf.

Helbing, D. and Nagel K. (2004), The physics of traffic and regional development, *Contemporary Physics*, September–October 2004, volume **45**, number **5**, pages 405–426.

Landis, J. and Zhang, M. (1998) The second generation of the California urban futures model Parts 1, 2 and 3, *Environment and Planning B: Planning and Design*, Vol. **25**, pages 657–666 and 795–824.

Marchal, F., and Nagel, K. (2005), Computation of Location Choice of Secondary Activities in Transportation Todels with Cooperative Agents, in Whitestein Series in Software Agent Technologies, Applications of Agent Technology in Traffic and Transportation, Birkhäuser Basel.

PTV (2008) – PTV traffic mobility logistics. http://www.ptv.de. Accessed October 2008

Shibuya, T., Ikeda, T., Imai, H., Nishimura, S., Shimoura and H., Tenmoku, K. (1995) Finding a realistic detour by AI search techniques, *Proc. 2nd Intelligent Transportation Systems*, November 1995.

Salvini, P.A. and Miller, E.J. (2005) ILUTE: An Operational Prototype of a Comprehensive Microsimulation Model of Urban Systems, *Networks and Spatial Economics*, Vol. **5**, pages. 217–234.

Strauch, D., Hertkorn, G., Wagner P., and Kühne, R. (2002) Neue Ansätze zu einer mikroskopisch-dynamischen Verkehrs- und Flächennutzungsplanung im Verbundprojekt ILUMASS.- Tavangarian, D. and R. Grützner [Hrsg.]: *Simulationstechnik - Frontiers in Simulation* **12**.- SCS-Publ., Gent: 511–516.

Vancheri, A., Giordano, P., Andrey, D. and Albeverio, S. (2008a) A model for urban growth processes with continuous state cellular automata, multi-agents and related differential equation. Part 1: Theory, *Environment and Planning B: Planning and Design*, vol. **35**, issue **4**, pages 723–739.

Vancheri, A., Andrey, D., Giordano, P. and Albeverio, S. (2008b) A model for urban growth processes with continuous state cellular automata, multi-agents and related differential equation. Part 2: Computer Simulations, *Environment and Planning B: Planning and Design*, vol. **35**, issue **5**, pages 863–880.

Waddell, P. (2000), A behavioural simulation model for metropolitan policy analysis and planning: residential location and housing market components of UrbanSim, *Environment and Planning B: Planning and Design*, vol. **27**, issue **2**, pages 247–263.

Weidlich, W. and Haag, G. (Eds.) (1999), An Integrated Model of Transport and Urban Evolution: with an Application to a Metropole of an Emerging Nation, Springer, Heidelberg.