

2001-07-18  
Incomplete version  
Not to be quoted

## Activity-based model development to support transport planning in the Stockholm region

*Staffan Algers<sup>1</sup>, Jonas Eliasson<sup>2</sup> and Lars-Göran Mattsson<sup>1</sup>*

<sup>1</sup>Division of Transport and Location Analysis

Department of Infrastructure and Planning

Royal Institute of Technology

SE-10044 Stockholm, Sweden

<sup>2</sup>Transek AB

Solna Torg 3

SE-171 45 Solna, Sweden

### Abstract

The environment in which transport analysis and infrastructure planning takes place has changed dramatically during the last years. The focus is now, to a considerable extent, on how to transform the transport system in a direction that could be sustainable in the long run, rather than on planning for infrastructure investment to meet new demand. At the same time information technology penetrates all sectors of the society. This will change how the transport system will be used by travellers and conveyers, both directly, through new products and services, and, indirectly, through a spatial reorganisation of many activities that govern the demand for transport. In such a situation it must be questioned whether the modelling tools that may have functioned reasonably well in the past, also are appropriate, or possible to adapt, to be useful for the issues that urban and regional planners will face in the future.

A literature survey of activity based models is made. The main interest is on such models that could be useful for a medium-sized city like Stockholm. It stresses demands that might be raised on modelling tools with a background in the planning issues that can expected to be central within the next ten-year period. The development of activity-based models is discussed with respect to their theoretical appeal, empirical validity, usefulness for planning, need for data, and obstacles of implementation.

To be presented at the 41<sup>st</sup> Congress of the European Regional Science Association  
Zagreb, August 29 – September 1

## 1. Introduction

For some decades now, transport researchers have put considerable efforts on developing what is called activity-based approaches to transport analysis (for some overviews, see e.g. Arentze and Timmermans, 2000; Ettema and Timmermans, 1997; McNally, 2000). The basic idea is that travel demand is a derived demand based on people's desire to take part in different activities. In particular, the interrelationships among different activities with respect to temporal and spatial constraints are in focus. It means that such models treat the activities and the travelling of the households with respect to where and when the activities can be carried out, and how they may be scheduled, given characteristics of the households and potential opportunities, the transport networks and various institutional constraints.

Conventional transport demand models, on the other hand, usually take trips as their starting point. The trips are then modelled with respect to their generation, distribution, modal split and assignment to the networks for road and public transport. These models have turned out to be very useful in many respects, and also reasonably easy to furnish with data and to implement. It has also been possible to extend their area of application as new policy issues have come to the forefront. Yet there is a growing awareness that there are many important policy issues for which these models are inadequate. Conventional transport demand models have often been criticised for being more suitable for construction of new infrastructure than for analysis of how an efficient management of existing demand can be accomplished (cf. Kitamura et al., 1997). For the evaluation of travel demand management strategies related to environmental and societal costs of private car use such as congestion, pollution, noise and accidents other tools may be necessary (see e.g. Gärling et al., 2001; Pendyala et al., 1997).

Activity-based models with their explicit focus on how people organise their activities in time and space, seem to apply a very appropriate perspective in evaluating such strategies. This may also be true for other very urgent issues such as the implications of information technology (IT) on the efficiency of the transport system and on mobility and land use. This involves how IT can be used for short-term traffic control, how IT may change short and long term demand for transport and also how IT may transform the urban structure and hence the supply of activities that give rise to mobility and goods movement. However, conventional models rely on a long history of experience, while activity-based models are still close to the research frontier and so far there is only limited experience of applying them in policy studies.

In this paper we will review some activity-based approaches with the aim of identifying promising models for possible application in the Stockholm region. The selection will be guided by the usefulness of the models for handling planning issues that can be expected to be central within the next ten-year period. The different approaches will be evaluated with respect to their theoretical appeal, empirical validity, usefulness for planning, need for data and easiness of implementation.

## 2. Background

The activity-based approach is not new. The intellectual roots go back to Hägerstrand (1970) and his time-geography and to Chapin (1974) and his emphasis on individuals' desires and personal characteristics behind their engagement in different activities. Whereas Hägerstrand stressed various forms of constraints, Chapin's interest was more on opportunities and choice. One can make some parallels to the methodological discussion on decision-making principles for travel and activity decisions. Whereas activity-based approaches often very explicitly advocate bounded-rationality principles, state-of-the-art travel demand models are usually based on (random) utility-maximisation (McFadden, 1974; Ben-Akiva and Lerman, 1985). This distinction should not be overemphasised. Some popular discrete choice models, not the least the logit model, can be perceived both as the result of stochastic utility maximisation and as a way of representing procedural bounded rationality (Mattsson and Weibull, 2001).

Hägerstrand (1970) introduces three kinds of constraints on the activities an individual can undertake: “capability constraints” are biological constraints related to, e.g., the need of eating and sleeping; “coupling constraints” reflect that certain activities like a meeting require people to be at the same place at the same time; and “authority constraints” are external institutional constraints set by various kinds of regulations such as opening hours of shops and working hours of employees. This in combination with the location of the opportunities where different activities can be performed, and considering by which velocity the individual can move given available means of transport, determine the prisms in the time-space, inside which the individual has to act. As we will see, activity-based models are very much about to make this brilliant but simple idea operational.

Among recent operational activity-based models three major classes are discernible: random-utility-based models (e.g. the Portland model by Bowman and Ben-Akiva, 2000), computational process approaches (e.g. PCATS by Kitamura and Fujii, 1998, and AMOS by Pendyala et al., 1998), and rule-based approaches (e.g. ALBATROSS by Arentze and Timmermans, 2000).

The daily activity schedule model for Portland, which was originally applied for Boston, is the model among those mentioned above that is most closely related to the conventional trip-based travel demand models. On important extension, though, is that it is tour-based rather than trip-based. This means that all travel is viewed as round-trips starting and ending at home and that the temporal and spatial connections among the activities within one tour are handled. A daily activity pattern, which consists of a set of tours, forms the basis for the individual’s activity and travel demand. The tours are divided into primary and secondary tours, for which the choices of destination, mode and time of day are modelled. The theoretical framework of the model is random utility maximisation in form of nested logit models. The approach has clear resemblance with some Scandinavian model development, notably the SIMS model by Algiers et al. (1995) and the PETRA model by Fosgerau (2001).

Also the PCATS model, which is an acronym for the Prism-Constrained Activity-Travel Simulator, is based on utility maximisation. More interesting is, however, that it is an attempt to represent Hägerstrand’s concept of time-space prism in a model system of activity engagement and travel simulation. Activities are considered fixed or flexible depending on whether they have to be carried out at predetermined locations or not. Flexible activities can be performed during the intervals that are open between ending and starting times of fixed activities. The speed of travel, locations and timing of the fixed activities together define the time-space prisms, within which the flexible activities are constrained. The individual is assumed to select those flexible activities and mode of travel that maximise the sum of the utilities of the activities and of the associated trips, where exogenous variables are assumed to influence these utilities.

The AMOS (Activity-Mobility Simulator) model, also developed by Kitamura and others, has a different purpose. It is meant as a system for analysis of behavioural adjustment, e.g., in response to proposed travel demand management policies like parking tax, improved bicycle/pedestrian facilities and congestion pricing as in its prototype application in Washington. The idea is to start with an observed daily activity pattern for an individual; generate a modified pattern as an adaptation to some change in the external conditions; evaluate whether this new pattern is good enough with respect to some satisficing rule; accept it or continue the search process until an acceptable pattern has been found. The decision criterion belongs to the class of bounded rationality rather than utility maximisation. It is thought that this should reflect the cognitive process of decision-making individuals as they adapt to changes in the environment.

Also the TRANSIMS model development at the Los Alamos National Laboratory should be briefly discussed. The purpose of this model system is to provide the transport planners with detailed information on how changes in transport policy or infrastructure might affect travel behaviour, congestion and pollution (see Bush, 2000; Arentze and Timmermans, 2000). As compared to previously discussed models, TRANSIMS has a much stronger emphasis on simulating the traffic and can also be useful for analysis of traffic safety

(Kosonen and Ree, undated). It combines activity-based travel demand generation with models for mode and route choice using a very advanced micro-simulation methodology, which makes use of cellular automata models (see Simon and Nagel, 1998). The system creates a synthetic population for a region from available data sets. Activity patterns are generated for the individuals in the population and later transformed into individual trip plans. These plans are inputs to the travel micro-simulation model. The trips are assigned to the network to satisfy a compromise of different goals concerning typically cost, travel time, departure time, congestion and safety.

ALBATROSS, finally, is a prototype of a rule-based system for predicting transport demand. It has been developed by a team at Eindhoven University of Technology. The purpose has been to develop a tool by which it should be possible to better assess the consequences of external changes like flexible work hours or longer opening hours of shops than with conventional models. It combines much of the experience that has been accumulated in the research on activity-based approaches over the last years. The system is available as a software. It consists of a scheduling engine that communicates with an inference engine. In addition there are a reporting agent and a scenario-building agent. The individuals are assumed to come to a choice about their daily activity pattern that satisfy their preferences while meeting given temporal and spatial constraints. ALBATROSS is well documented in a monograph, which also illustrates the possible use of the system for policy analysis (Arentze and Timmermans, 2000). Four examples of analysis are given: decrease of two-adult households, change of work start times, increase of part-time workers and Friday afternoon off.

### 3. Extending Random-Utility Approaches

An assumption in the mainstream approach to travel demand modelling has for long been to assume independence between different travel purposes. This is of course one of the main criticisms from activity-oriented modellers. Research has demonstrated, however, that it is possible to extend the modelling of trip generation from purpose specific frequencies to frequencies of combinations of trip purposes (or, as would be more appropriate to put it, combinations of activities). Such an approach adds of course to the computational task, but is fairly straightforward. It makes it necessary for the frequency models to be based on disaggregate data. It is difficult to see any specific obstacles to such a development. The gain is that a more realistic travel behaviour can be represented.

The lack of trip chaining in conventional travel demand models has also been criticised by activity-oriented modellers (see e.g. Ettema and Timmermans, 1997). This is another example of extension to mainstream modelling that should be possible to include in a better way in the future. This could for example be done by the inclusion of trip type choice and the concept of a secondary destination, conditioned on the movement between two other points, such as home and work, as is suggested by Daly and Lindveld (1995). Such an extension, will increase the computation burden substantially. This may have been an obstacle earlier, but as computers grow faster this will change. So far, the obstacles have been circumvented by simplified implementation of such extensions. The gain is a more realistic modelling of travel behaviour and its economic and environmental impacts.

#### The Portland Model

For Portland Metro (Oregon, US) a model system has been developed and implemented, containing an advanced example of how the choice of a combination of activities during a day can be modelled together with other choice dimensions (Bowman et al., 1998; see also Bowman and Ben-Akiva, 2000):

In this model system, trip frequency for different trip purposes is modelled as a choice of a combination of activities. The Day Activity Pattern Model, as is it called, contains 114 alternatives, differing with respect to activities involved and the order in which the activities are performed. The choice set is described in the following way:

##### Primary activities:

- Subsistence (work or school) on tour
- Subsistence (work or school) at home
- Maintenance (shopping, personal business etc.) on tour
- Maintenance at home
- Discretionary (social, recreation, entertainment etc.) on tour
- Discretionary (social, recreation, entertainment etc.) at home

If the primary activity is on-tour, the day activity pattern model also determines the trip chain type for that tour. There are eight possible types for work/school tours and four possible types for maintenance and discretionary tours. The trip chain type is defined by the number and sequence of the stops in the tour. The alternatives that apply for all trip purposes are:

- Simple tour
- One or more intermediate activities on the way from home to the primary destination
- One or more intermediate activities on the way from the primary destination to home

- Intermediate activities in both directions

For work/school tours, types five through eight are define as above with the addition of a work based sub-tour.

Simultaneously with primary activity and primary tour type, the day activity pattern model predicts the number and purposes of secondary tours. There are six alternatives:

- No secondary tours
- One secondary tour for work or maintenance
- Two or more secondary tours for work or maintenance
- One secondary tour for discretionary purpose
- Two or more secondary tours for discretionary purpose
- Two or more secondary tours: at least one for work or maintenance and at least one for discretionary purposes

Since not all of the tour types apply to all of the primary activity types, there are 19 possible combinations of primary activity/tour types. Each of the six secondary tour alternatives are possible for all primary activity/tour types, so the model has a total of  $19 \times 6 = 114$  alternatives.

### **Other choice dimensions**

Time of day is a choice dimension that will be increasingly important, as trends towards time differentiated pricing (phrased in many ways) makes this a more common policy option. This is related not only to urban car traffic, but also to public transport, and to long distance travel. Time of day is a central choice dimension in activity-based approaches. This kind of extension is not unusual in mainstream modelling either. It can be expected to become a standard part of conventional trip-based models in the next decade. It is usually modelled as an additional choice of time periods at the bottom of a nested logit model (Daly et al., 1990). In principle, there is only the (declining) obstacle of computational effort.

As information technology advances, it will be possible for individuals to substitute trips with other types of contacts, predominantly through the use of Internet. This concerns not only work or business trips, but all kinds of trips (activities) that do not require a physical presence. Thus, the ranges of substitution possibilities will in reality be extended to include also non travel alternatives. So far, some attempts have been made to survey communication in a broader sense (including travel as well as non travel contacts) (Zumkeller, 1996; SIKa, 1998), but how to model it, and how (if possible) to include it in the mainstream approach is still a challenge. The main obstacle here seems to be that research has not come far enough, although research is ongoing (Mokhtarian and Salomon, 1997). It seems however to be of vital importance to make achievements in this respect during the next decade. This holds for activity-based as well as conventional models.

## 4. Computational Process and Stated-Behaviour Approaches

### PCATS

“The development of PCATS was motivated by the recognition that various constraints imposed on individuals’ activity and travel are not well represented in conventional models of travel behavior. Emphasized in PCATS, therefore, are the constraints imposed on the individual’s movement in geographical space along time. Because the speed of travel is finite while the time available for travel and activity is limited, the individual’s trajectory in time and space is necessarily confined within Hägerstrand’s prism. PCATS first identifies the set of prisms that govern an individual’s behaviour, then generates activities and trips within each prism while observing coupling constraints involving private modes and operating hours of public transit” (Kitamura et al., 2001).

PCATS simulates the activities and trips of an individual during a day. The simulation is undertaken such that space-time constraints are satisfied. The starting point of the simulation is the separation of the day into *open periods* and *blocked periods*. Blocked periods are periods where the individual has committed to engage in certain activities (called *fixed activities*) at certain locations. This applies to for example work and sleep. Open periods are periods in which the individual has the option of travelling and engaging in activities (called *flexible activities*). Given the speed of travel, the ending time and location of a blocked period and the beginning time and location of the subsequent blocked

Time 0	Sleep	At home
6	Open	
9	Work	At work
12	Open	
15	Work	At work
18	Open	
21		
24	Sleep	At home

**Figure 4.1.** A typical activity schedule in PCATS

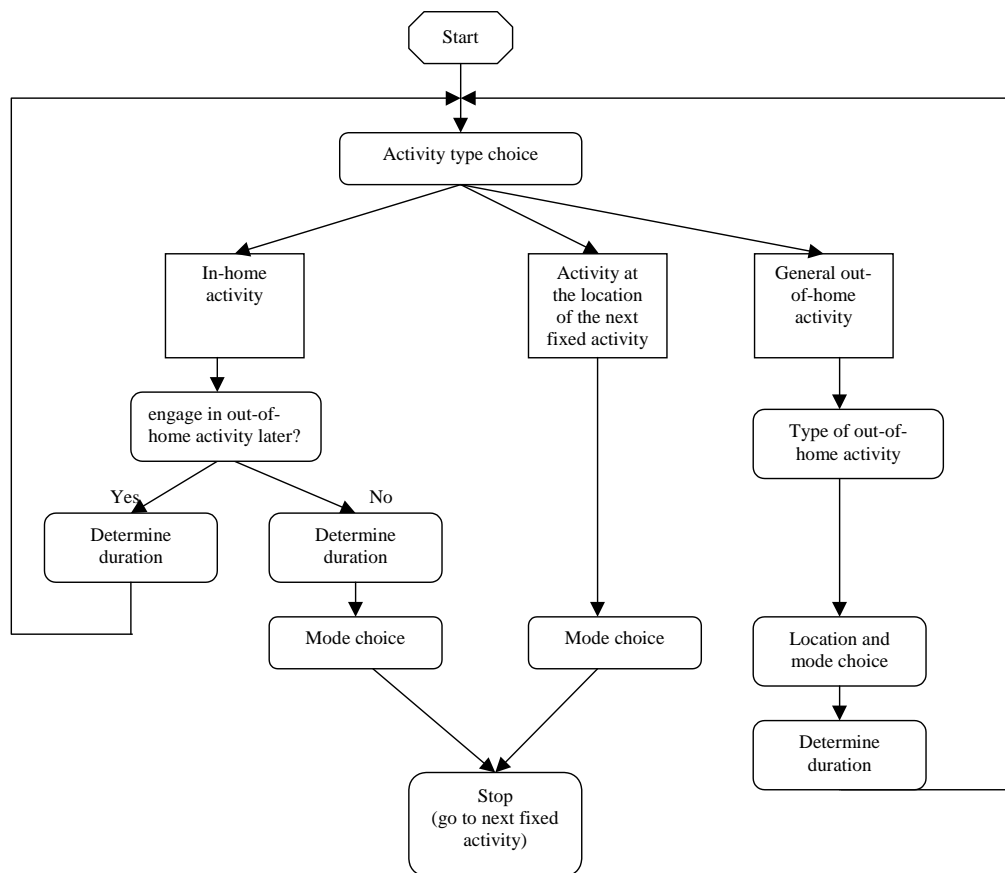
period, define a *time-space prism* in which the individual’s activity and travel are constrained. The simulation is carried out by starting in the morning, generating possible activities and trips sequentially until the day is over.

Figure 4.1 shows a typical activity schedule. Blocked periods are grey, and activities and locations are given before the simulation starts. Open periods are white, and activities and locations will be determined by the simulation.

**The PCATS flow chart**

When a fixed activity is over, the individual is at a certain location at a certain point in time. The open period that now follows (lasting until the next blocked period) is filled with activities by the simulation system in the following way (see Figure 4.2).

First, one of four possible activity types is chosen. The first possible activity type is to stay at home for a while and engage in an out-of-home activity later. In this case, the home-stay



**Figure 4.2.** Flow chart of PCATS duration is determined, and a new activity type is chosen, this time with the two “at home”-alternatives unavailable.

The second activity type is to stay at home until it is time to leave for the next fixed activity. The duration is then determined by the starting time of the next fixed activity minus the time it takes to get there. Lastly, the mode to get to the next fixed activity is chosen. This ends the simulation of the open period at hand.

The third activity type is to engage in an activity at or near the location of the next fixed activity. This will determine the location and the duration of the activity, so only mode choice will have to be simulated.



The fourth activity type is to engage in a general out-of-home activity at a general location. In this case, the sub-type of activity is also simulated (meal, social, grocery shopping, comparison shopping, hobbies/entertainment and sports/exercise). Then, the location of the activity and the mode to get to this activity is simulated, and the duration of the activity is determined. The process is then repeated from the top.

Durations are always determined such that the activity fits into the time available. Likewise, activity types, modes and locations are determined such that it will be possible to arrive in time to the next fixed activity. In this way space-time prism constraints are used to exclude certain activity types, modes and locations from the respective choice sets. This is what guarantees that the prism constraints are satisfied. Furthermore, so-called “coupling constraints” are used, primarily for car availability: if car was not used to work, for example, it cannot be used from work.

### *The PCATS sub-models*

The PCATS model is composed of many sub-models, corresponding to the boxes with rounded corners in the flow chart (see Figure 4.2). The activity type choice model is a nested logit model with the two types of in-home activities in one nest and the different types of general out-of-home activities in one nest. The duration models are hazard models with truncated Weibull distributions (the truncation ensures that the activity will fit in the open period with probability one). The mode choice is a multinomial logit model. The location and mode choice model is a nested multinomial logit model, where the mode choice level appears to be the same as the mode choice model just mentioned. At present, the mode and location choice models are identical for all trip types, which is an admitted weakness that will be corrected in later versions.

The variables in the models are specified in Table 4.1.

### *Data and estimation*

The models are estimated from activity diary data. This type of data is an extension of the usual trip diary in that not only trips but also activities are registered in the diary.

**Table 4.1.** Variables in PCATS

Model	Variable type	Variables
<b>Activity type</b>		
Out- of home activity type	Socio-economic	Age, sex, home-maker
	Time of day Available time	Probability that the activity fits within the open period
Duration	Socio-economic	Age, sex, employment, income, license, no. of vehicles
	Past activities	Total time spent on the activity previously the same day
	Time of day Available time	Time to next fixed activity
	Location type (home or non-home)	Indicators of the combination of the current location type and the location of the next fixed activity
Mode	Socio-economic	Age, sex, employment, income, license, no. of vehicles
	Travel characteristics	Travel times, travel costs, no. of transfers, intra-zone dummies
	Time of day	

Location	Location type (home or non-home)	Indicators of the combination of the current location type and the location of the next fixed activity
	Zone characteristics	Population, number of commercial establishments, intra-zone dummies
	Travel time	Minimal travel time to the location and then to the next fixed activity
	Available time	Probability that the activity fits within the open period

The activity categories used in the travel diaries are sleep, personal care (other than taking bath), personal care (bath), child care, meal, domestic chore, work and work-related, school and study, social, grocery shopping, comparison shopping, hobbies and entertainment, sports and exercises, TV viewing, reading, resting, medical and dental, and others. Not all of these were deemed relevant to have been explicitly represented in the model. The activities which are actually used in the model can be found under “activity types” above. The rest of the activity types appear to have been grouped together under for example “in-home activities”.

A crucial point is what activities should be considered fixed. A set of assumptions was adopted to determine whether an activity is fixed or flexible. Sleep is always classified as a fixed activity. Personal care (other than taking bath), personal care (bath), TV viewing, reading, and resting, on the other hand, are always classified as flexible. Activities of the remaining types are classified as fixed if the respondent indicated in the survey that the activity was subject to both temporal and spatial constraints; otherwise they are regarded to be flexible.

It appears as though each sub-model is estimated separately, which is probably the only way to estimate a model of this complexity (at present, at least). All sub-models are estimated with maximum-likelihood estimation. For complete estimation results, the reader is referred to an unpublished M. S. thesis in Japanese (Otsuka, 1996).

The estimated model that has been reported does not seem to replicate estimation data too well, although most key indicators have the right magnitude. Table 4.2 lists some key figures.

**Table 4.2.** Results from the estimation of PCATS

	Predicted		Observed		t	R <sup>2</sup>
	Mean	S.D.	Mean	S.D.		
Total travel time	116.3	70.7	127.9	87.1	-2.00	0.622
In-home flexible activity duration	314.5	152.9	288.7	191.0	2.04	0.673
Out-of-home flexible activity duration	28.4	72.3	39.6	75.6	-2.07	0.329
Number of non-work destinations	0.071	0.61	0.31	0.58	-5.42	0.169
Number of non-work trip chains	0.059	0.28	0.013	0.11	2.86	-0.027
Number of trips	2.89	1.56	3.38	1.79	-4.00	0.576

S.D.: standard deviation across sample individuals

t: t-statistics associated with the difference between the predicted and observed values (not based on the standard deviations associated with “predicted” values)

R<sup>2</sup>: Pearson correlation coefficient between predicted and observed values

Source: Kitamura and Fujii (1998)

The most striking discrepancies between observed and predicted values are the differences in number of non-work destinations and number of non-work trip chains. Since this touches the heart of what activity-based models are supposed to handle better than traditional models, namely how activity and trip patterns are generated, this seems problematic. However, this table is from an early version of PCATS, which is more of an “initial prototype” than a real, operational model meant for forecast use, according to the authors (Kitamura and Fujii, 1998).

### *Extensions*

Recently, PCATS have been integrated with DEBNets, a dynamic network model, and HAGS, a model forecasting household ageing, formation and fixed activities (see Fujii, 2001).

## **AMOS**

“Activity-Mobility Simulator (AMOS) is a micro-simulation model system of individuals’ adaptation behavior which predicts changes in travel behavior that will follow a change in the travel environment. The individual’s adaptation behavior is characterized as a trial-and-error experimentation process” (Kitamura, 1996).

What separates AMOS from nearly all travel models is first that rather than simulating an activity-travel pattern “from scratch”, it starts with a real activity-travel pattern taken from someone’s activity diary. The model then predicts, using a neural network, how this activity-travel pattern will change if a certain change takes place (say, car tolls or better bicycle lanes). This also highlights the second particular characteristic of AMOS, namely that it is not meant to be able to analyse just about any change in the transport system, but is explicitly constructed to analyse consequences of a relatively small number of pre-specified policies.

At the heart of the AMOS model lies the neural network that predicts how people’s activity-travel pattern will be changed by a particular policy. The data needed to “train” the neural network to do this is collected by interviewing people how they would change their activity-travel pattern the previous day if a certain change had taken place. Data collection through this sort of open-ended interviews is the third characteristic that separates AMOS from most other models.

### *What issues can be handled?*

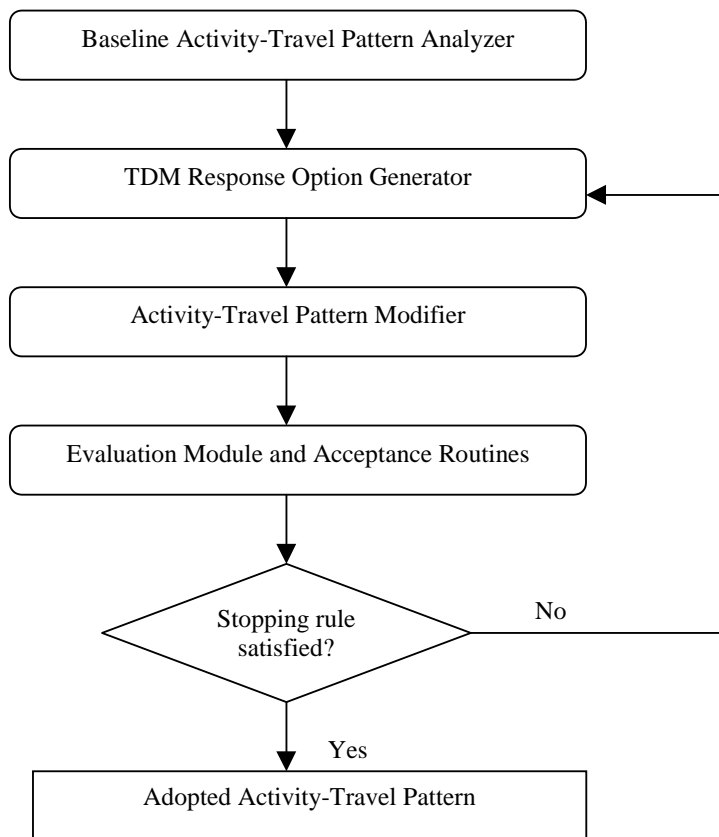
As pointed out above, this question is actually more crucial for AMOS than for many other models. The heart of AMOS is the “Response Generator”, which simulates how an individual reacts to a given travel demand management policy (TDM). Consequently, the TDMs of interest must be identified in advance, before data collection can start<sup>1</sup>.

### *AMOS flowchart*

The Baseline Activity-Travel Pattern Analyzer reads individual trip records, compares them with network data for logical consistency and missing information, and then generates a coherent baseline activity-travel pattern (called activity pattern from now on for brevity) of all documented individuals. These individual activity patterns forms the base of the subsequent modelling, since it is those that are modified by the later modules. This module is not really a part of the model as such, since it is not used when the model is run.

---

<sup>1</sup> AMOS also uses ordinary trip diary data and network supply data, apart from the dedicated interviews particular for AMOS. The former data can of course be collected in advance, and is often already present.



**Figure 4.3.** The flowchart of AMOS

The TDM Response Option Generator creates the “basic” response of an individual to a travel demand management policy (TDM). The module consists of a neural network trained by using stated and revealed preference data (see below under Estimation). The module takes as input the activity patterns from the database and the corresponding socio-economic data and generates a stochastic “response” to the given TDM. The response can be “do nothing” (which is often the case), switch work trip mode, change departure time of work trip or work at home.

The response is then fed into the Activity-Travel Pattern Modifier. This module executes the necessary adjustments of the whole activity pattern (if any). For example, a shopping stop on the way home may no longer be viable if the response was to switch from car to bicycle. Instead, a new (home-based) shopping trip may emerge. This adjustment process follows a complex set of rules and criteria.

The modified activity pattern is then fed into the Evaluation Module and Acceptance Routines. This module evaluates the new activity pattern, and determines whether it is “sufficiently good” based on a set of rules. If this is not the case, the process is repeated with a new response generation.

**Data and estimation**

AMOS needs three data sources. First, travel supply data are needed, such as network data, travel costs and travel times etc. Second, standard trip diary data are needed. These two data sources are likely to be present at most planning agency, so there is no need for additional dedicated data collection if one wants to set up an

AMOS installation in a particular area. The third survey, though, is the survey that constitutes the essence of AMOS. The survey is conducted as a telephone interview. Respondents are first asked to describe their activity pattern the previous day. The interviewer then describes a possible TDM policy, such as increased fuel prices, congestion pricing, improved bicycle lanes etc. For each of those policies, the respondent indicates how his or her activity pattern would have changed if this TDM policy would have been in effect. Responses include for example “nothing would change”, “travel at a later time”, “go by bus instead” etc. The interviewer then, together with the respondent, investigates how this would affect the entire activity pattern. For example, stops on the way home from work might be home-based trips instead, or perhaps postponed to another day.

These responses are then used to train the neural network. (The neural network is the module of the model that is used to generate responses to given TDM policies – see above.) Travel supply data, characteristics of the activity pattern and socio-economic characteristics of the respondent are used as explanatory variables to forecast the response obtained at the interview.

## 5. ALBATROSS – A Rule-Based Activity Approach

### Background

ALBATROSS is a rule-based activity model for predicting travel demand. It has been developed by a team of researchers of the Eindhoven University of Technology on the commission of the Ministry of Transportation, Public Works and Water Management in the Netherlands. The project has been led by professor Harry Timmermans. The development of the prototype model is documented in an extensive book (Arentze and Timmermans, 2000). The review that follows will be based on this book without any further reference except when exact citations are made.

The researchers behind the model system have for a long time been active in the field of travel demand analysis, including the development of activity-based models. ALBATROSS is in many respects a synthesis of their accumulated experiences.

One reason for starting the project was the view that traditional trip-based travel demand models have become less suitable for handling many urgent emerging transport policy issues. Examples of this are changes in the institutional environment of the households such as changing work hours (the amount of hours or the extent to which they are flexible), changing opening hours of shops and increased participation of women in the workforce. It is easy to imagine that new trends towards a 24-hour society may alter the behaviour of the travellers in a way that has consequences for the performance of the transport system, especially for the level of congestion and for the time profile of the demand. Another example is the increasing interest in evaluating various travel demand management strategies related to environmental and societal costs of private car use such as congestion, pollution, noise and accidents. This has accentuated the need for models of great policy sensitiveness. Conventional models have often been criticised for being more suitable for the construction of new infrastructure than being useful for the management of existing demand (cf. Kitamura et al., 1997). In particular, temporal, spatial and institutional constraints are difficult to handle in an adequate way with these models. It is reasonable to think of travel demand as the outcome of a complicated process where such constraints interfere with people's desires to participate in different activities at different times and locations, and given the properties of the transport system. In this perspective it is easy to understand that activity-based models often have been put forward as a alternative promising approach for the next generation of modelling tools.

In contrast to traditional travel demand models rooted in discrete choice theory, activity-based models are specifically designed to handle the complex interplay between the needs of the individual or the household and physical, temporal, economic and social constraints such as where and when different activities can be performed, how easy activities can be combined into tours, which modes of transport there are available and how fast an individual can move from one activity in one place to another activity in another place given the properties of the transport system. Traditional travel demand models usually treat trip generation, distribution, mode choice and route assignment separately for each trip type. The ambition with ALBATROSS is to go beyond this by predicting which activities will be conducted where, when, for how long, with whom and which modes and routes that will be used for the travelling that will be the result of the activity pattern. In this meaning the travel demand will be derived from the activity pattern.

### The structure of ALBATROSS

An activity pattern is the result of a complex interaction between the needs of the individual or the household and various constraints. One obvious such constraint is the *physical environment* that describes where different activities can be conducted and the attributes of these locations. Because of limited or imperfect information a particular individual  $r$  in a household  $h$  will only be familiar with a limited number of locations,

his *cognitive environment*  $L_{rh}(\tau)$ . Its dependence on  $\tau$ , the sequencing variable, indicates that the cognitive environment is not time-invariant but may vary with when in a schedule the activity will take place. Activities are also constrained by institutional context. It could be rules or regulations concerning the number of work hours and flex-time, opening hours of shops or requirements on individuals such as age limits to be allowed to have a driver's license. Such constraints for activity  $a$  at location  $l$  are summarised in the *land-use pattern*  $G_l(a)$ .

In the modelling of the activity decision process a distinction is made in ALBATROSS between different time frames. An *activity agenda*  $A_h$  is a set of activities that the individuals in household  $h$  need or wish to carry out within some time period. It is expressed at the household level to indicate that some activities like shopping or leaving or picking up children at the day-care centre may be allocated to different members of the household. An important classification of household activities is between *in-home* and *out-of-home* activities and between *mandatory* and *discretionary* activities. From the activity agenda activities are drawn that are planned to be conducted on a specific day. This subset defines the *activity program*. The *activity schedule* is then the precise sequence of these activities in time and space during the day. Included in the sequencing is also a choice of transport mode from the set of available modes for individual  $r$  of household  $h$  during day  $d$ ,  $M_{rhd}(\tau)$ . In the present version of the system the modes public transport, slow transport and car passenger are always available, while the availability of car as driver depends on the characteristics of the individual and the household for that particular day. During a particular day, an individual may also be involved in unplanned activities. The activities that are actually performed may differ from those planned. The realised activities are referred to as the *activity pattern*.

The execution of an activity program involves the decision about which activities to carry out, where, when, with whom and which modes to use to go to the different locations. More precisely, let  $A_{almw}^{rhd}(\tau)$  denote that individual  $r$  of household  $h$  on day  $d$  will carry out activity  $a$  at location  $l$  together with travel party  $w$  and use mode of transport  $m$ . To predict the choice of a specific activity pattern, ALBATROSS models the choice probability of this bundle

$$\Pr\{A_{almw}^{rhd}(\tau) \mid A_h; L_{rh}(\tau); M_{rhd}(\tau); G_l(a)\} \quad \forall \tau, r, h, d$$

as a function of the activity agenda, the cognitive environment, the available modes and the land use pattern including institutional context.

ALBATROSS involves different time frames. The generation of activity agendas is to a large extent dependent on a number of long-term decisions such as family status, choice of occupation, workplace, residence and car ownership. These decisions are fundamental for which activity patterns will be possible to perform for different household types. The generation of activity programs will depend on the nature of activity (mandatory or discretionary) and on how necessary it is to complete a particular activity on a specific day. When the activities in a program are scheduled in an activity pattern, it is necessary to observe a number of logical, institutional, household (such as bringing children to school), spatial, time and spatial-temporal (related to the transport system) constraints. Given these constraints the question is how the individuals will choose between different feasible activity patterns.

A distinct feature of ALBATROSS is that it assumes that the decision-making is based on some *learning mechanism*. This is a deliberate choice of the developers of the system, since such a mechanism is believed to better represent actual choice behaviour than the often applied principle of utility-maximisation. Consider an individual who is in a new choice situation. We assume that he comes to a decision in some way. If he is satisfied with the outcome he may continue to make the same decision in similar choice situations in the future. If he is not satisfied, he may try another alternative the next time. The idea of a learning mechanism is that behaviour leading to positive experiences will be reinforced. Gradually such experiences are transformed

into ready-made heuristics that the individual will apply in specific choice circumstances. If these circumstances change, the individual will not start to examine all possible alternatives once again, but rather be guided by some generic rule that has turned out to work satisfactory in the past. In that case the individual evaluates different alternatives according to some mental simulation model of the choice situation including its associated constraints. To summarise, the individuals make long-term adaptations of their behavioural rules in response to interactions with the environments. They also make short-term adaptations based on the outcome of the mental simulation of generic rules.

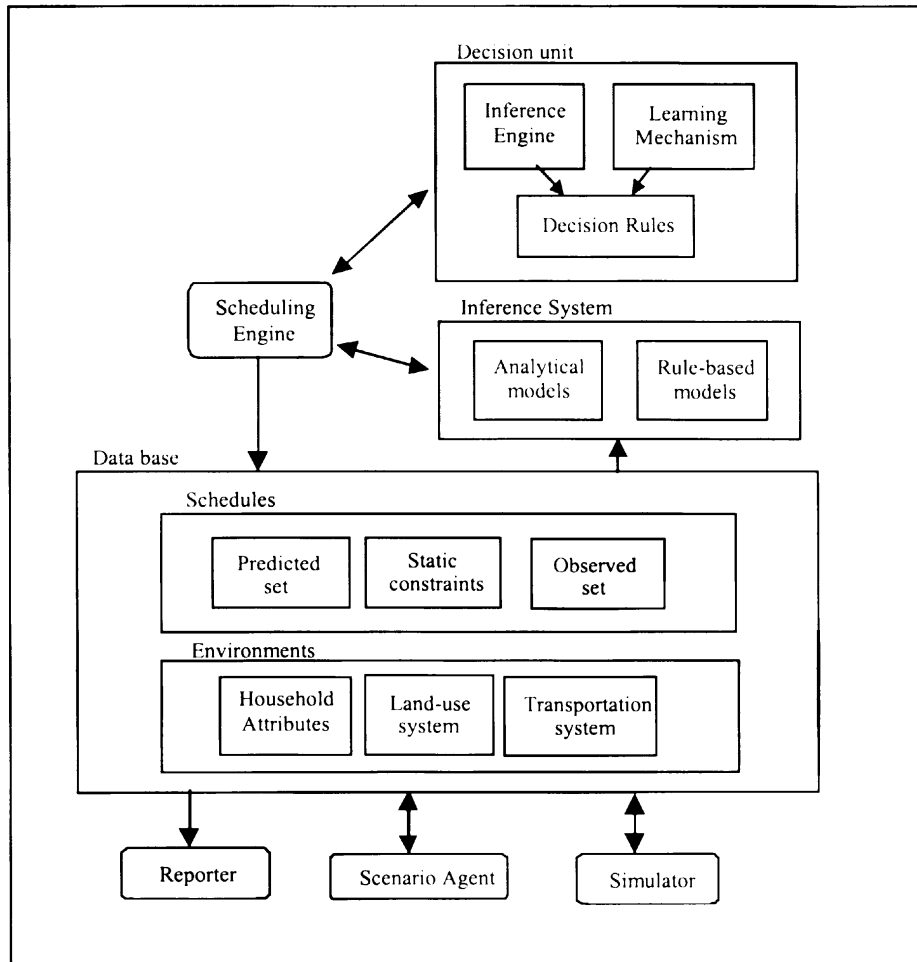
The decision process implemented in ALBATROSS is sequential. This may introduce some problems, since earlier decisions in a schedule could unnecessarily reduce the options for later decisions. To minimise this effect, an elaboration of the present system is planned. In this new version the scheduling process will be carried out in two stages, a planning and an execution stage. In the planning stage a preliminary schedule will first be formed, which is later rescheduled to resolve possible conflicts between earlier and later decisions. In the execution stage the activity pattern will be rescheduled once again, because unplanned events may introduce new conflicts with earlier decisions or open for possible improvements of the original schedule. For the moment only the preliminary schedule module is in place.

Next we will discuss in some detail how the activity scheduling process has actually been operationalised and implemented. The architecture of ALBATROSS is shown in Figure 5.1. The scheduling process is performed in a number of steps that are controlled by the Scheduling Engine. The intention is that this should imitate how individuals or households are solving their scheduling problems. The Scheduling Engine identifies relevant information for the Decision Unit, arranges that this information is produced by the Inference System and employs the decisions from the Decision Unit in forming an activity schedule. In the Decision Unit an Inference Engine selects and applies relevant decision rules from a rule base that in the present version of the system could be built up in a pre-processing stage by the Learning Mechanism. The Inference System uses a number of built-in analytical and rule-based models that should capture available knowledge about such scheduling constraints that are stable and not varying with individuals or environments.

The Database keeps information about the study area (household and population attributes, land-use data including location and opening hours of facilities for different activities, transport system characteristics in form of travel time matrices by mode) and schedule information by household and day of week. The latter information is divided into an Observed Set used for comparing the activity patterns generated by the model and observed data, a set of Static Constraints that contains information about static household- or individual-related constraints that are needed by the Scheduling Engine, and a Predicted Set that stores the generated activity patterns. Finally there are some modules that perform certain tasks. The Reporter generates statistics for impact analysis and also measures of goodness-of-fit between generated and observed activity patterns, while the Scenario Agent helps the user to define and manipulate input data for the scenarios to be studied. This can be done by changing the composition population segments, possibly introducing direct behavioural change through exogenous variables (e.g. change the start time for the journey to work) or alter the impact assessment by changing other exogenous variables. The Simulator should among other things simulate traffic flows on the road network and dynamically adjust travel times to network capacities. This module still remains to be implemented, however.

The structure of ALBATROSS is further discussed in connection with the description of the software.





**Figure 5.1.** Architecture of the ALBATROSS system  
 Source: Arentze and Timmermans (2000, Fig. 3.2, p. 83)

**The scheduling process**

A specific scheduling problem could and would probably be solved in many different ways by different individuals. Such individual variations can to some extent be covered by the decision rules of the Decision Unit. To make the modelling task more reasonable, however, the scheduling process has been simplified in several respects in the present version of the system. First it is assumed that the household is the fundamental decision unit. Further the scheduling is restricted to activities for the adults only. If there are children in a household this will be treated as part of the conditions for the scheduling problem of the adults rather than making any attempt to generate the activity patterns separately for the children.

The problem now is to schedule the activities for a particular (adult) individual  $r$ , who is a member of a particular household  $h$ , during a particular day  $d$  of the week. The outcome of the process, the activity pattern,

will be a schedule in form of a list  $S_r^{h,d}$  of activities, ordered sequentially ( $\tau = 1, 2, \dots, \Gamma$ ), and an activity profile for each activity  $\tau \in S_r^{h,d}$ , i.e. information about activity type  $a(\tau)$ , travel party  $w(\tau)$ , duration  $\Delta(\tau)$ , start time  $t^s(\tau)$ , location  $l(\tau)$ , transport mode  $m(\tau)$  and travel time  $t^t(\tau)$ .

The different activities that can be scheduled are exogenously separated into fixed and flexible activities. This reflects the previously discussed distinction between mandatory and discretionary activities. The fixed activities are those that have to be included in the schedule and for which the location, start time and duration are treated as given. They represent activities generated by the individual's long-term plan. Work would be a typical example. Flexible activities are those that may or may not be included in the schedule. If they are included, their activity profiles will be determined during the scheduling process. ALBATROSS also distinguishes between in-home/out-of-home activities, with only one type of in-home activity. It should be noted that travel and waiting time are not considered an activity of its own right but is rather treated as part of the profile of the associated activity.

The set of fixed activities, denoted  $S_r^{x,h,d} \subseteq S_r^{h,d}$ , constitutes the schedule *skeleton*, which expresses the fact that these activities form the basic structure of the schedule. The scheduling process will complement this skeleton by adding flexible activities to available time slots. It is assumed that the activity types are ranked according to some pre-defined priority list. The priorities are used when choices among different activities have to be made.

The scheduling process is carried out in a number of steps (our description here follows closely that in Arentze and Timmermans (2000, p. 88-90)). The term *choose* will be used to indicate when the Decision Unit gives input to the Scheduling Engine (see Figure 5.1).

- Step 0*      *Initialise the schedule with the given set of fixed activities*  
For each  $r$  set  $S_r^{h,d} = S_r^{x,h,d}$ .
- Step 1*      *Select the transport mode for each primary work activity*  
For each  $r$  and each primary work activity  $\tau_w \in S_r^{h,d}$ , *choose* the transport mode  $m(\tau_w)$ .
- Step 2*      *Select and add instances of flexible activities to the schedule and specify the travel party*  
For each  $r$  and for each flexible activity  $a$ :
1. initialise an instance  $\tau$  of activity type  $a$  by setting earliest and latest start and end times and minimum and maximum duration;
  2. *choose* whether  $\tau$  should be added to  $S_r^{h,d}$  or not;
  3. if  $\tau$  is added then set  $a(\tau) = a$ , *choose* travel party  $w(\tau)$  and duration  $\Delta(\tau)$ , re-define minimum and maximum duration given  $\Delta(\tau)$  and add  $\tau$  to  $S_r^{h,d}$  in a possible position given time constraints;  
[  $w(\tau) \in \{\text{alone, only inside household, outside household}\}$ ;  $\Delta(\tau) \in \{\text{short, average, long}\}$ ]
  4. if an instance has been added go back to 1 and decide whether a new instance of the same activity type should be added or not.

- Step 3 For each flexible activity, determine the time of day in which the start time should fall and place the activity in an appropriate position in the schedule*  
For each  $r$ :
1. define  $S^* = S_r^{x,h,d}$  ;
  2. for each flexible activity type  $a$  and instance  $\tau \in S_r^{h,d}$  , such that  $a(\tau) = a$  , choose the time-of-day for the start time of  $\tau$  , given this re-define earliest and latest start and end times and add  $\tau$  to  $S^*$  in an appropriate place and determine an initial start time  $t^s(\tau)$  ;  
[  $t^s(\tau) \in \{\text{before 10 AM, 10-12 AM, 12-2 PM, 2-4 PM, 4-6 PM, after 6 PM}\}$  ]
  3. set  $S_r^{h,d} = S^*$  .
- Step 4 For each activity, determine the position in the schedule and trip link with previous and/or next activities or insert in-home activities if needed*  
For each  $r$ , determine the organisation of trips into trip chains:
1. define  $S^* = S_r^{x,h,d}$  ;
  2. for each flexible and out-of-home activity type  $a$  and instance  $\tau \in S_r^{h,d}$  , such that  $a(\tau) = a$  , choose whether the activity is conducted directly after, before or in-between specific activities in  $S^*$  , or as a separate trip; add  $\tau$  to  $S^*$  in a position consistent with the choice of trip type; if needed add an in-home activity before or after the activity;
  3. set  $S_r^{h,d} = S^*$  ;
  4. for each consecutive pair of fixed activities in  $S_r^{h,d}$  , add an in-home activity in-between the two activities only if the time gap is larger than estimated travel time plus a pre-defined maximum waiting time.
- Step 5 For each trip chain in the schedule determine the transport mode*  
For each  $r$  and  $\tau \in S_r^{h,d}$  determine the transport mode  $m(\tau)$  :
1. identify trip-chains  $TC_i \subseteq S_r^{h,d}$  as any sequence of activities beginning and ending at home and including at least one out-of-home activity;
  2. for each  $TC_i$  set  $m(TC_i) = m(\tau_w)$  if  $\tau_w \in TC_i$  or else choose  $m(TC_i)$  ;
  3. set  $m(\tau) = m(TC_i)$  for each out-of-home activity  $\tau \in TC_i$  and also for the last in-home activity of  $TC_i$  (i.e., the return trip).
- Step 6 For each flexible out-of-home activity determine the location and travel time*  
For each  $r$  and each  $\tau$  such that  $a(\tau)$  is a flexible out-of-home activity choose the location  $l(\tau)$  and determine the travel time  $t^t(\tau)$  .

In case of two adults in the household, the scheduling process is carried out in parallel to take interdependencies into account. This means, however, that the allocation process is represented somewhat implicitly. In addition, there is no mechanism to ensure that coupling constraints between members in a travelling party will be met. As indicated in step 1, the mode choice for the primary work activity is considered a high level decision that determines how the car, if there is only one, will be used for different activities by different members of the household.

When flexible activities are considered for inclusion in step 2, which is done according to the pre-defined priority list, the locations are unknown and time constraints are set according to the widest possible opening hours across available facilities. If an activity is added, the position is chosen according to a minimax rule to keep maximum flexibility for remaining activities. This means that the new activity is placed in the smallest possible time slot between any two activities given the minimum duration of the activity. The position of the flexible activities is further determined in step 3.

In step 4 out-of-home activities are combined into trip chains, unless they are separated by an in-home activity. New in-home activities may be added to divide a trip chain into smaller chains or separate trips. The choices are based on imprecise information about temporal constraints since location, travel time and duration are still not finally determined. As indicated in step 5, mode choice is made at the trip chain rather than the trip level, i.e. it is not possible to change mode within a trip chain. All trip links in a chain that includes the primary work activity will use the mode of that activity. Finally, locations and travel times are determined in step 6. Different activity types are then assigned to locations according to the priority list. The system identifies location alternatives by means of generic rules that represent heuristics that individuals are assumed to use. These rules express different ways of making trade-offs between travel time and attractiveness of the locations.

### Decision rules

The choices that are made at various stages in the scheduling process are governed by decision rules. These rules play a very significant role in ALBATROSS, which is the reason to say that it is a rule-based activity model system. Rules may be derived in many ways. A very interesting feature of the ALBATROSS system is that these rules are derived in a formalised way from empirical data.

A decision rule is represented in form of a decision table, *DT*, which consists of a list of condition variables and connected to these a list of action variables. The condition variables are related to characteristics of the individual, the household, the activity program, the physical environment, the transport system, the institutional context and other schedule information. The action variables represent available choice alternatives for each choice situation. For each combination of condition variables, the *DT* expresses what action is taken with respect to the different action variables.

The idea is simplest illustrated by an example, see Table 5.1. Here we have two condition variables: travel distance to a location and parking facilities at the location. The action variables are travel modes: bike, car and public transport. The *DT* expresses whether a particular travel mode will be chosen or not for each particular combination of condition variables. It should be noted that a *DT* must meet some rather obvious formal requirements in order to be useful: exhaustiveness, exclusiveness and consistency.

**Table 5.1.** Example of a decision table

C1	Distance (m)	D < 500	500 ≤ D < 1000		D ≥ 1000	
C2	Parking facilities	Bad or good	Bad	Good	Bad	Good
A1	Bike	×	×	-	-	-
A2	Car	-	-	×	-	×
A3	Public transport	-	-	-	×	-

The symbol × means that the action will be executed whereas - means that it will not be executed

Source: Arentze and Timmermans, 2000, p. 109, Table 4.1

The rules should represent behavioural heuristics in a realistic way. The developers discuss two alternative ways of deriving decision rules from empirical data. In the first case the idea of a learning algorithm is explored, while in the second case a Chi-square Automatic Interaction Detection (CHAID) algorithm is studied.

The idea of a learning algorithm is illustrated for a simplified scheduling problem concerning a particular individual a particular day. An activity program,  $AP$ , is assumed given. From the beginning the schedule is empty. If we would like to evaluate all possible sequences of activities, the number of evaluations would soon be overwhelming as the number of activities increases. Instead a simplifying algorithm is used in which the activities are added to the schedule one by one. The scheduling decision at a particular step consists of two parts, the selection of an activity from those not already scheduled, and the selection of a position in between the activities already scheduled. This reduces the number of evaluations drastically for a typical number of activities.<sup>2</sup>

For each part of a decision step, constraint rules and preference rules are applied. Constraint rules represent constraints on feasible schedule positions for an activity and have the highest priority. They could include temporal as well as sequence constraints, and they are specified a-priori by the system designer. The idea behind this is that they should represent basic knowledge about the decision situation common for all individuals. The preference rules, on the other hand, are the counterpart to the utility functions of the discrete choice models in conventional travel demand models. Similar to these, they are typically derived or “estimated” from behavioural data. An interesting property is that they offer a possibility to capture how individual behaviour may be subject to adaptation over time.

One way of deriving the preference rules is through a “supervised inductive learning algorithm”. It is then assumed that we have data on activity programs  $AP_p$  and associated “observed” schedules  $S_p$  for a sample of individuals and days, or person-days,  $p$ . It is also assumed that the set of decision tables  $DT_t$  are given, where we use  $t$  to index this set. It should be noted that this algorithm can, in principle, optimise over all  $DT_t$ :s simultaneously.

We consider a situation where only one action alternative can be chosen for each column in each  $DT_t$  as in Table 5.1. The idea of the algorithm is to “train” the  $DT_t$ :s by applying them to the conditions of the activity programs  $AP_p$  of the “observations” and compare the resulting schedules  $S'_p$  to the observed schedules  $S_p$ . The aim is to find the actions in the  $DT_t$ :s that minimise some aggregate measure of the distance between observed and generated schedules. More precisely, considering a particular  $AP_p$ , then for each relevant column of the  $DT_t$ :s an action is selected by a probability that is related to how well this action has explained observed schedules for the same conditions for previous observations (initially all actions are assumed equally good). A schedule  $S'_p$  is then generated according to the selected actions of the  $DT_t$ :s, and depending of how close this  $S'_p$  is to the observed schedule  $S_p$ , the choice probabilities for the selected actions will be increased or decreased. These new choice probabilities are then applied when the next observation is processed. Possibly, this process will converge and we have derived, or estimated, the actions in the  $DT_t$ :s.

The algorithm can be characterised as an incremental learning process. There is a positive feedback from how successful an action is in generating schedules close to observed ones and the probability that this action will be included in the final  $DT_t$ :s. Obviously the outcome will depend on how the algorithm is specified in detail:

---

<sup>2</sup> Let it be  $n$  activities in the  $AP$ . These activities could be arranged in a sequence in  $n!$  different ways. To find the best sequence would require  $n!$  evaluations. The suggested decision procedure will only require  $2(n + (n - 1) + \dots + 2) = n^2 + n - 2$  evaluations.

How the distance between generated and observed schedules are measured? How the  $DT_i$ s are specified and how the action probabilities are updated during the process?

This algorithm is tested in some numerical experiments. Some of the parameters that govern the detailed behaviour of the algorithm are also varied. The algorithm seems to work satisfactory and is in a series of experiments capable of “deriving” the  $DT_i$ s that were used for producing the “observed” schedules. The subject of modelling learning and adaptation processes has been further elaborated on in a later work by Arentze and Timmermans (2001).

A drawback with the discussed inductive learning algorithm is that it requires a pre-defined condition structure, i.e. the subdivision of the decision tables into columns must be defined in advance. We would rather prefer that this subdivision and the selection of actions could be optimised simultaneously to attain the best fit with observed data. An alternative algorithm that can accomplish this is described below. It is based on the previously mentioned Chi-square Automatic Interaction Detection (CHAID) algorithm, originally developed by Kass (1980). Also this algorithm has a drawback. It can only handle one  $DT_i$  at a time.

To apply the algorithm we need a sample of person-days for which we have observations of the condition states, with respect to a number of pre-defined condition variables, and the chosen actions, as exemplified in Table 5.1. In contrast to this table, however, we should know the exact values for the condition variables. Assume first that there is only one column in our potential decision table, i.e. that all observations are in one condition category. The distribution of actions across the alternatives represents the heterogeneity of the observations. Considering some condition variable, the question is if we can split the observations into two or more columns by means of some condition expressed in terms of the condition variable, so that the observations in each column will be more homogenous than the original column. If so, the process is continued for each condition variable as long as the homogeneity can be significantly increased and the number of observations in each column will not be too small.

Finally, the action for each column should be determined. This could be done in a deterministic way by assigning each column the action that is most frequently chosen. It could also be done in a probabilistic way by setting the probability of choosing a particular action equal to its relative frequency in the particular column.

In the applications reported on, the developers eventually decided to use the CHAID-algorithm to derive the rule-based decision tables that should represent the choice heuristics of the individuals in the scheduling processes.

### **How ALBATROSS can be used**

We will here briefly discuss how ALBATROSS can be used for impact analysis by reviewing some prototypical examples of application.

In the present version of the system no simulation model for synthesising input data is implemented. It is therefore necessary to assume that the sample of households and schedule skeletons on which the development of the system is based, are representative for the entire population of interest. The system has some special modules for impact analysis (see Figure 5.1). The Scenario Agent is used for building relevant policy scenarios by changing the composition of households, the schedule skeletons or both. By means of the Reporter output from the simulations will be generated according to a pre-defined format.

Four policy scenarios defined by the Dutch Ministry of Transport are studied:  
*Scenario 1. Decrease of two-adult households*

A decrease of two-adult households by 10 % is simulated. The number of single-adult households is increased so that the total size of the population and also the number of person-days are kept constant.

*Scenario 2. Change of work start times*

For a 15 % random selection of work activities with work start time in the interval 7.00 AM – 9.30 AM, the start time is changed to 7.00 AM or 9.30 AM whichever is closest. End times are changed similarly so that the work activity duration times remain the same. Some other activities must also be moved on the time scale so that the duration of the activities can be kept constant.

*Scenario 3. Increase of part-time workers*

The number of full-time workers is reduced by 10 %. This is applied with the same probability to each of the categories dual-earner two-adult households, single-earner two-adult households and single-earner one-adult households. The number of part-time workers is increased so that the number of person-days remains the same for each household category. Full-time is interpreted as more than 32 hours per week and part-time as 17-32 hours per week.

*Scenario 4. Friday afternoon off*

For Friday schedule skeletons with more than 6 hours of work time, the duration of the work activity is reduced by 50 % by changing the end time while letting the start time remain the same.

The policy scenarios and the null scenario were based on around 2,200 households, slightly varying dependent on scenario conditions. Because of the stochastic character of the simulations, they must be repeated to even out random fluctuations. The simulations were repeated 20 times.

**Results**

The output data generated by the system can be divided into three parts: travel demand indicators (total distance, number of tours, trips per tour, etc.), activity patterns and household and individual characteristics of the sample population. Table 5.2 and 5.3 summarise the results in form of percentage change as compared to the null scenario for a few travel demand indicators and for activity type frequencies.

In addition to the changes displayed in the Tables 5.2 and 5.3 a few other results should also be mentioned. In Scenario 1 less activities start before 10 AM and more in the interval 4 – 6 PM. As a consequence of the change in household composition both the oldest households (> 65 years) and the households without children increase by 10 %. In Scenario 3 there is an increase of the share of activities starting in the interval 10 – 12 AM. There is also a shift from younger (25-45 years) to older households (>45 years) by about 2 % and related to that an increase of households without children by 1 %.

**Table 5.2.** *Percentage change for selected travel demand indicators as compared to the null scenario*

Indicator	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Total travel distance	-1.5*	0.3	-3.0*	0.4
Distance car (driver + passenger)	-1.3*	0.7	-2.8*	1.3*
Number of tours	-0.7	0.3	-0.2	1.6*
Number of trips per tour	0.3*	-0.1	0.1	0.1
Fraction of total travel distance				
by car driver	0.5	0.6	0.0	0.4
by car passenger	-4.4	-1.6	2.9	7.2*
by public transport	0.8	-0.3	-5.8	-8.9*
by slow mode	-0.2	-2.0	3.7*	0.0

\* Significant change at 5 % level according to a two-sided *t*-test

An important consequence of Scenario 1 is a significant decrease of total travel distance. One possible explanation could be that single-adult households tend to live closer to work than two-adult households. The result could also be related to the other changes in household composition mentioned above.

For Scenario 2 no significant effects are found. The simulated change is obviously too small to lead to any measurable impacts.

**Table 5.3.** *Percentage change in activity type frequency as compared to the null scenario*

Activity type	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Work out-of-home	-0.4	0.0	-3.2*	0.1*
Bring/get	-4.0*	-0.0	2.1*	0.1*
Non-leisure	-0.4	0.0	7.6*	0.0
Daily shopping	-0.6	-1.5	-0.2	3.1*
Service	3.8*	2.0	4.6*	8.4*
Non-daily shopping	0.4	-0.6	0.8	5.5*
Social visits out-of-home	0.4	1.4	0.7	3.8*
Leisure out-of-home	0.3	-0.0	-2.1*	-0.2
Other (medical, personal business etc.)	2.8*	0.1	0.5	0.5*
Total	-0.2	0.0	0.0	1.7*

\* Significant change at 5 % level according to a two-sided *t*-test

Scenario 3 implies an even greater decrease in total travel distance than Scenario 1. The reduction of work activities is partly neutralised by an increase in other out-of-home activities. Also in this case the change in household composition might have influenced the impacts.

In Scenario 4 the total distance by car and the number of tours both increase significantly. A pronounced increase in shopping, service and out-of-home social activities are also noticeable. The increased car use might be explained by the fact that due to less work hours, there is an increased availability of the car for other activities and for non-working members of the household. It should be noted, however, that the present version of ALBATROSS does not allow for a re-scheduling of activities among the days of the week.

Before ending it should be noted that the system is capable of producing much more detailed results than indicated above. For instance, the simulation results can be subdivided by weekday. Effects on car ownership, travel party composition, activity duration and start time are available as well as on origin/destination behaviour. Detailed information on household composition including size, socio-economic class, age, number of children, work hours and gender shares are also available.

### **Data requirements, calibration and validation**

The data requirements for activity-based models are in general more demanding than for conventional travel demand models. This is rather obvious since such a model should be able to predict the travel behaviour in detail including how the activities are selected and scheduled. This holds also for ALBATROSS, in which case predictions are made about which activities will be conducted where, when, for how long, with whom and by which mode of transport.

To calibrate such a model, behavioural data are first of all needed about activities. In the present case a full (i.e. including also in-home) activity two-day diary was used for collecting such data. For each specific activity the respondent had been involved in, he was asked to report kind of activity according to a pre-defined scheme, day, start and end time, location of the activity, transport mode, travel time, accompanying



persons and if the activity was planned or not. In addition, conventional socio-economic data were also collected.

For the calibration of the model and for the policy analyses additional data are needed about the location of facilities where activities can be carried out, the institutional context and the transport system. Location data include floor space and number of employees for daily and non-daily shopping activities, other service activities and out-of-home leisure activities. Data about the institutional context relate to opening hours for the same kind of activities. The transport system was represented by model generated travel time matrices for the modes car, public transport, bike and walking assuming a non-congested situation. No travel cost data were used in the present calibration of the model system. Some of these data that were used had to be collected by fieldwork.

Respondents tend to consider the completion of activity diaries a demanding task. Typically there are problems with errors and inconsistencies. Related to the data collection for ALBATROSS, the developers have also developed an interactive computer system for the logical test of consistency and completeness of activity diaries, SYLVIA<sup>3</sup>. The system helps the user to identify such errors and to some extent automatically correct them.

The calibration, or estimation, of the system was carried out by applying the CHAID-based algorithm to derive necessary decision tables from the described data. This was done in several steps, starting with the transport mode choice in connection with the primary work activity. The predictive ability of the derived decision table for the mode choice was 65 % as compared to 53 % for a null-model.

In the next step decision tables for activity selection, travel party and duration decisions were derived. This was followed by the derivation of decision tables for activity start time and trip-chaining decisions. The hit rate of start time was reported to be about 40 % and that of trip-chaining to be around 80 %. Finally, the decision tables for the transport mode and location decisions were derived. The predictive ability was reported to be satisfactory also in this case.

Validation of behavioural models is a difficult task. Even if the absolute predictive ability according to the kind of measures discussed above is satisfactory, this does not exclude that another specification of the model would have performed better. Activity-based models are also very demanding in many respects as compared to conventional travel demand models. Is this extra effort worthwhile? Do activity models perform better than conventional models? As an additional way of validating ALBATROSS such a comparison was performed. As the competitor model a combined multinomial logit and Poisson model was developed. To avoid any prejudices this second model was developed by a different team, however, using the same data sets as for the development of ALBATROSS. The resulting Linked Logit – Poisson Model, hereafter referred to as LLPM, describes choice of activity calendar, number of tours, allocation of activities to tours, activity sequence, destination and transport mode.

The model systems are compared in two ways. First, origin-destination trip matrices are derived from the generated activity patterns and compared with observed matrices and, second, the similarity between observed and predicted activity patterns are evaluated by different measures including some based on sequence alignment techniques.

ALBATROSS performs consistently better than LLPM for five tested matrix specifications with respect to both a contingency and a correlation coefficient measure. Typically LLPM achieves 65 to 90 % of the scores for ALBATROSS. Also with respect to the second comparison ALBATROSS performs consistently better than LLPM for all measures applied. In sum, ALBATROSS outperforms LLPM indicating that it represents a considerable improvement as compared to more state-of-the-art approaches.

---

<sup>3</sup> SYSystem for the Logical Verification and Inference of Activity diaries

## **Computer software**

ALBATROSS is available as a software system which can be run interactively under Windows. The intention with the software is that it should be generally applicable and it is accompanied with a fairly detailed manual and data test sets. It is unknown whether the system actually has been transferred to any place outside the original study area in the vicinity of Rotterdam, or has been applied by anybody outside the original development team.

The system has tools for reading, exporting and viewing schedule data and for selecting schedules. New scenarios can be defined by a scenario builder that allows the user to multiply cases, to change household, activity and environment attributes, and to use pre-defined scenarios. A reporter agent allows the user to generate frequency tables and OD-matrices and to calculate performance indicators based on observed and/or predicted schedules. Results from a scenario can be stored in Excel 5.0 format for comparison with other scenarios and further analysis. A tool for goodness-of-fit analysis is available for comparing observed and predicted schedules based on the previously discussed sequence alignment technique. As also discussed earlier, there is a tool, SYLVIA, for checking input diaries for errors and inconsistencies and for automatic data repairing. This tool can be used in connection with the collection of new sets of activity data.

The ALBATROSS software has not been applied by the authors. It is therefore difficult to judge how user-friendly it is. One question mark is about computing time. Nothing is said in the documentation on this point other than that the simulations may take much time dependent on the number of cases.

## References

- Algers, S., Daly, A., Kjellman, P. and Widlert, S. (1995) "Stockholm model system (SIMS): Application", Paper presented at the 7th World Conference of Transportation Research, Sydney.
- Arentze, T. and Timmermans, H. (2000) *ALBATROSS: A Learning Based Transportation Oriented Simulation System*, European Institute of Retailing and Service Studies.
- Arentze, T. and Timmermans, H. (2001) "Modeling learning and adaptation processes in activity-travel choice: A framework and numerical experiments", Urban Planning Group, Eindhoven University of Technology.
- Bowman, J.L. and Ben-Akiva, M.E. (2000) "Activity based disaggregate travel demand model system with activity schedules", *Transportation Research A*, 35: 1-28.
- Ben-Akiva, M. and Lerman S.R. (1985) *Discrete Choice analysis: Theory and Application to Travel Demand*, MIT Press, Cambridge, MA.
- Bowman J.L., Bradley M.A., Shiftan Y., Lawton T.K. and Ben-Akiva M.E. (1998) "Demonstration of an activity based model system for Portland", In Selected Proceedings from the 8th World Conference on Transport Research, Antwerp, Belgium, 12-17 July 1998, Volume 3: Transport Modelling/Assessment.
- Bush, B.W. (2000) "An algorithmic overview of TRANSIMS", <http://transims.tsasa.lanl.gov>.
- Capin, F.S. (1974) *Human Activity Patterns in the City*, Wiley, New York.
- Daly, A. and Lindveld, K. (1995) "Forecasting non-home-based trips as components of a tour model", PTRC proceedings, P392 ISBN 0-86050-282-1.
- Daly, A. et al. (1990) "Peak-period proportions in large-scale modelling", PTRC proceedings, P335 ISBN 0-86050-221-X.
- Ettema, D.F. and Timmermans, H.J.P. (1997) (eds) *Activity-Based Approaches to Travel Analysis*, Elsevier Science Ltd.
- Fosgerau, M. (2001) "PETRA: An activity-based approach to travel demand analysis", in Lundqvist L. and Mattsson, L.-G. (eds) *National Transport Models: Recent Developments and Prospects*, Forthcoming Springer-Verlag.
- Fujii, S.R. (2001) ...
- Fujii, S.R. Kitamura, R. and Monma, T. (1997) A micro-simulation model system of individuals' activity-travel patterns based on random utility theory. Mimeo.
- Gärling, T., Eek, D., Loukopoulos, P. et al. (2001) "A conceptual analysis of the effects of travel demand management on private car use", Mimeo, Department of Psychology, Göteborg University.
- Hägerstrand, T. (1970) "What about people in regional science?", *Regional Science Association Papers*, 24:7-21.
- Kass, G.V. (1980) "An exploratory technique for investigating large quantities of categorical data", *Applied Statistics* 29: 119-127.

Kitamura, R. (1996) "Applications of Models of Activity Behavior for Activity Based Demand Forecasting", Presented at the TMIP meeting, Washington.

Kitamura, R. (2001) ...

Kitamura, R. and Fujii, S. (1998) "Two computational process models of activity-travel choice", in Gärling T, Laitila, T and Westin, K (eds) *Theoretical Foundations of Travel Choice Modeling*, Elsevier Science Ltd, 251-279.

Kitamura, R., Fujii, S. and Pas, E.I. (1997) "Time-use data, analysis and modeling: toward the next generation of transportation planning methodologies", *Transport Policy*, 4:225-235.

Kitamura, R., Fujii, S. Yamamoto, T. and Kikuchi, A. (2001) "Application of PCATS/DEBNets to regional planning and policy analysis: Micro simulation studies for the cities of Osaka and Kyoto, Japan", Presented at the Transportation Research Board.

Kitamura, R., Lula and Pas E. (1993) "AMOS: An activity-based, flexible and truly behavioral tool for evaluation of TDM measures", PTRC 1993.

Kitamura, R., Pendyala, R.M., Pas, E. and Reddy, D.V.G.P. (1995) "Application of AMOS, an activity-based TCM evaluation tool to the Washington DC area. Proceedings of the 23<sup>rd</sup> summer annual meeting", PTRC Education and Research Services, London.

Kosonen, I. and Ree, S. (undated) "The potential of microscopic simulation in traffic safety and conflict studies", Unpublished manuscript, Royal Institute of Technology, Stockholm, and Los Alamos National Laboratory.

Mattsson, L.-G. and Weibull, J.W. (2001) "Effort-based probabilistic choice", submitted to *Games and Economic Behavior*.

McFadden, D. (1974) "Conditional logit analysis of qualitative choice behavior" in Zarembka, P. (ed.) *Frontiers in Econometrics*, Academic Press, New York, 105-142.

McNally, M. (2000) "The activity-based approach", in Hensher, D.A. and Button, K.J. (eds) *Handbook of Transport Modelling*, Elsevier Science Ltd, 53-69.

Mokhtarian, P. and Salomon, I. (1997) "Emerging travel pattern: do telecommunications make a difference?" Resource paper presented at the 7th IATBR conference, Austin, Texas.

Otsuka, Y. (1996) "Development of a simulation model system of activity and travel which represents the individual's decision process", M.S. Thesis, Department of Applied Systems Engineering, Kyoto University.

Pendyala, R.M., Kitamura, R., Chen, C. and Pas E.I. (1997) "An activity-based microsimulation analysis of transportation control measures", *Transport Policy*, 4:183-192.

Pendyala, R.M., Kitamura, R. and Reddy, D.V.G.P. (1998) "Application of an activity-based travel demand model incorporating a rule-based algorithm", *Environment and Planning B*, 25:753-772.

SIKA (1998) "IT-development and transport", SIKAR Report 1998:1 (in Swedish).

Simon, P. and Nagel K. (1998) "Simplified cellular automaton model for city traffic", *Physical Review E*, 58:1286-1295.

US Department of Transportation/TMIP (Pendyala, Reddy, Hawks, Chen, Lula, Kitamura) (1995) "Activity-based modeling system for travel demand forecasting", US DoT-T-96-02.

Zumkeller, D. (1996) "Communication as an element of the overall transport context - an empirical study", Fourth International Conference on Survey Methods in Transport.