CHAPTER 97

Research Avenues

Kai Nagel, Kay W. Axhausen, Benjamin Kickhöfer and Andreas Horni

The on-going work documented and interest expressed in the various scenarios proves that this system has not at all exhausted its possibilities as a platform for research both on and with it. This chapter highlights chances for further discussion and action.

97.1 MATSim and Agents

97.1.1 Complex Adaptive Systems

The core MATSim architecture, where agents learn utilities for plans, was originally derived from the field of Complex Adaptive Systems (CAS; e.g., Axelrod, 1984; Holland, 1992; Hraber et al., 1994; Palmer et al., 1994) (also see Section 46.1 of this book). Arthur (1994) addresses a coordination problem where agents receive a payoff only when less than 60 out of 100 go to an event. He addresses this by first generating a large number of heuristic predictors for the next round's attendance, such as "same as in last round" or "trend from last four rounds". He next gives each agent a randomly selected handful of these strategies, so that agents have different sets of predictors. Then, many rounds of the game are played, where the score of each predictor is updated based on its prediction quality, and agents act based on their currently best predictor. Simulations demonstrate that the approach leads to successful coordination, i.e., around 60 agents show up in every round. That approach, in turn, builds on work by Palmer et al. (1994), who simulate a stock market, Holland (1992), whose classifier systems have more structure than Arthur's model, but a similar model of performance learning, or Axelrod (1984), who investigates adaptive agents in the face of repeated non-cooperative games.

Arthur (1994) keeps each agent's predictors fixed after initialization. In contrast, Hraber et al. (1994) simulate an artificial ecosystem, where individual agent strategies are based on so-called genes, adapted over the rounds/iterations by genetic algorithms (Goldberg, 1989).

How to cite this book chapter:

Nagel, K, Axhausen, K W, Kickhöfer, B and Horni, A. 2016. Research Avenues. In: Horni, A, Nagel, K and Axhausen, K W. (eds.) *The Multi-Agent Transport Simulation MATSim*, Pp. 533–542. London: Ubiquity Press. DOI: http://dx.doi.org/10.5334/baw.97.

97.1.2 Artificial Intelligence

CAS focuses on many agents, agent interaction and emergence. Artificial Intelligence (AI) in contrast, concentrates on single agents. In AI terms, the original MATSim agents (those doing day-to-day learning) are very simple reinforcement learning agents (Russel and Norvig, 2010, Chapter 21.3). Since these MATSim agents have only one state (the initial/nightly state) and each action is simply a plan, the distinction between Q-learning and utility learning (as defined by Russel and Norvig, 2010) actually collapses; what remains is the temporal difference learning (again as defined by Russel and Norvig, 2010) scheme for the utility, which translates to the MATSim situation by updating the score/performance/utility value of each plan every time it is selected.

97.1.3 Synthesis

The original MATSim system thus took the focus on large systems, interaction, emergence, and strategy innovation from CAS, while the score updating comes from the AI field. In consequence, a clear path to move on is the inclusion of more modern AI aspects into the MATSim agents. Examples include:

- Extend MATSim to agents that can react immediately, rather than having to wait for the next iteration or round. In transport, this is sometimes called en-route or within-day replanning (e.g., Emmerink et al., 1995; Balijepalli et al., 2007; Axhausen, 1990). See Chapters 30 and 23 as well as Section 97.2.
- Improve the MATSim agents with respect to choice set generation. This may include both better creative capability for the agents to come up with innovative new strategies to handle their virtual lives, as well as consistency considerations between choice set generation and estimated choice models. See Sections 49.2 and 97.3.

97.2 Within-Day Replanning and the User Equilibrium

Within-day replanning, i.e., the ability of the agents to respond to the immediate context, is the standard mode of operation for simulation models. In the transport domain, note the traffic flow models as an example, where aspects such as acceleration/braking, or lane changing, are (obviously) computed reactively, while the simulation is running and not before the simulation starts (e.g. Wiedemann, 1974). Many traveler-oriented or agent-based models of travel demand adopt the same approach, cf. ORIENT (Sparmann and Leutzbach, 1980), ORIENT/RV (Axhausen and Herz, 1989), MobiTOPP (Schnittger and Zumkeller, 2004). For many aspects of Intelligent Transport Systems (ITS) systems, within-day replanning is indispensable (e.g., Hall, 1993; Emmerink et al., 1995; Dobler, 2013). None of these systems aim for equilibrium in the same way as MATSim, carried forward from TRANSIMS and originally inherited from static assignment.

One may argue that, if supplied with a learning approach, these within-day models should approach equilibrium after many iterations, as agents with a suitable memory structure would avoid plans that could put them at a disadvantage. This memory, which would need to be agent-specific and covering the very large set of choice options, makes the approach costly to implement. Importantly, the solutions may be different: when faced with a stochastic environment, an agent able to react within-day could be better off than an agent following a pre-computed plan. This is important: finding a plan with the highest expected score is not the same as finding a conditional strategy with the highest expected score.

Still, there are contexts where this immediate response ability can be used within MATSim to explore the choice set more effectively, especially if the choice alternatives are limited and within geographic reach. Waraich et al. (2013a) proposes within-trip replanning to find the best parking space near a destination. This localized search reduces the need for full iterations considerably and allows addition of behavioral detail at this point (here the type of parking), walking distance to final destination, and parking fee trade-off.

While within-day replanning can be used as described above within the framework equilibrium search, it can also be added to open up the MATSim framework to contexts where such an equilibrium is inappropriate. Dobler (2013) uses the MATSim calculated equilibrium as the starting point for his model of evacuations and the behavior of evacuees. He also explains that his approach finds a set of executed plans close to the MATSim equilibrium, but for much lower computing cost. While the benefits of an equilibrium solution in comparison with an approximation have been extensively discussed for aggregate assignment models, for MATSim the issue is whether these fast approximations could be used to speed up the overall equilibrium search; similar to starting a Frank-Wolfe search based on four or five incremental loadings of the network (Jourquin and Limbourg, 2006).

While aggregate assignment can identify the routes chosen as belonging to the equilibrium, research in the agent-based context is needed to see: a) if the approach is indeed faster and b) if the resulting set of plans is unbiased by the fast initialization.

97.3 Choice Set Generation

As described at several places in this book (e.g. Sections/Chapters 3.1, 4.5.1, 49.2, 27, 47, 49), the MATSim iterative process in its standard version modifies each agent's choice set (= each agent's set of plans) over the iterations. Clearly, an agent can only select a plan generated by this process. Thus, search space definition is important.

97.3.1 The Statistical Weight of Each Plan

Econometric research (e.g., Ben-Akiva and Lerman, 1985, Chapter 8 and 9) points out that it is not sufficient if certain alternatives are eventually discovered by the search process; rather, it is important that they are generated with probabilities consistent with the choice model. This, however, is at odds with the CAS approach, where solutions are generated rather arbitrarily. For example, Arthur (1994) "create[s] 'an alphabet soup' of predictors" that are "randomly ladle[d] out".[1] Research is needed to clarify when statistical properties of the choice set need to be tightly consistent with the choice model and when not.

As a result, in the MATSim context, it is important to look not only at plans generation/innovation (e.g., Section 4.5.1.1), but also at plans removal. The default MATSim approach is to remove the plan with the worst score. This is, however, problematic both from a CAS and an econometric perspective. From a CAS perspective, such an approach simply does not generate enough diversity, since similar scores rather often mean similar plans; thus, the approach has a tendency to remove the most different plan, typically leading to a set of plans that are all quite similar. From an econometric/discrete choice perspective (cf. Section 49.2), the combination of plans generation and plans removal needs to ensure that each plan's probability of being in the choice set corresponds to its weight used in the choice model estimation.

Section 49.2 discusses a version of the plans' generation/removal process, but makes rather strong assumptions about the capability to compute best-response plans. Here, let us instead consider a heuristic argument. Assume that plans i for a person n are created with a certain probability $p_{n,i}^{create}$; the person index n will be dropped in the following. Also assume that plans are removed with probability p_i^{remove}. The master equation for the probability q_i of plan i to be contained in the choice

[1] To be precise, Arthur uses strategies that eventually *generate* choices.

set becomes in leading order[2]

$$\frac{dq_i}{dt} = -q_i \cdot p_i^{remove} + p_i^{create} \ .$$

The steady state solution, obtained from $dq_i/dt = 0$, is given by

$$q_i = \frac{p_i^{create}}{p_i^{remove}} \ . \tag{97.1}$$

That is, quite obviously, if one wants to control the statistical distribution q_i within the MATSim process, one needs to look not only at plans generation, but also at plans removal.

MATSim's plans generation model can be approximately described by a creation probability of $p_i^{create} \sim \exp(\beta^{create} S_i)$, with relative large β^{create}, corresponding to an approximate best-response model.[3] At the same time, removal of the worst plan corresponds to $p_i^{remove} \sim \exp(-\beta^{remove} S_i)$ with a very large β^{remove}. Overall, thus

$$q_i \sim \exp((\beta^{create} + \beta^{remove})\, S_i) \ .$$

Combining this with a choice model that selects with $\sim \exp(\beta^{choice} S_i)$ from the set of plans, i.e. `ChangeExpBeta` or `SelectExpBeta`, leads to

$$p_i \sim \exp(\beta^{choice} S_i) \cdot q_i \sim \exp((\beta^{choice} + \beta^{create} + \beta^{remove})\, S_i) \ . \tag{97.2}$$

Let us again stress that this is not an exact statistical analysis of the MATSim dynamics, but instead an illustrative approximation to gain insight. From this approximation, it becomes clear that MATSim in its current form, because of the strong additional effects of plans generation, expressed through β^{create}, and plans removal, expressed through β^{remove}, strongly over-weighs plans with high scores. *It is thus important to include plans removal in all considerations,* since otherwise the very large β^{remove} in Equation (97.2), coming from always removing the worst plan, will dominate the statistical distribution.

97.3.2 Heterogeneity in Plans Removal

Clearly, removing not the worst plan, but instead according to some logit model with a smaller β^{remove} would improve the situation. In addition, to increase diversity and simultaneously correct for correlations between alternatives, one could use an (inverse) path-size logit (e.g., Frejinger and Bierlaire, 2007; Prato, 2009; Schüssler, 2010) model, i.e.,

$$p_i^{remove} \sim \exp(-\beta^{remove} S_i + \alpha\, PS_i)$$

where PS_i would be an index of similarity of plan i to all other plans in the plans set. As a result, plans very similar to other plans in the set would have a greater chance of being removed. The last of such similar plans would no longer be similar to any other plan, thus PS_i would be small; that plan would be less likely to be removed.

Such an approach is experimentally available as `PathSizeLogitSelectorForRemoval`. It possesses an ad-hoc similarity computation of one plan to all other plans in the set (Grether, 2014). Further investigations using this approach should be performed.

---

[2] In higher order, one would have to correct for the possibility that a plan may appear more than once in the choice set.

[3] The operator $\sim$ means "proportional to in leading order". It neglects, for example, the effect of the denominator in a logit model.

97.3.3 Heterogeneity in Plans Generation

No extended research has yet been undertaken to see whether MATSim could adopt the strategy of regularly introducing new random starting solutions to avoid local minima. One challenge: the generation of such random plans would result—in most cases—in nonsensical plans, which would need to be removed through computationally expensive iterations. See Feil (2010) for difficulties in constructing optimal alternative plans for the number and sequence of activities. One possibility is to initially allocate a small set of randomly chosen alternative day plans and measure their similarity throughout the simulation. There is no research on a replanning technique involving switching of the day plan activity order, which again would produce more dissimilar plans than currently possible.

Moyo Oliveros and Nagel (in press) and Nagel et al. (2014) report computational experiments where a randomized Pareto router is used to generate a different route every time it is called. The Pareto router randomly draws a trade-off between different utility function contributions, such as fare/toll, travel time, access/egress time, then computes an optimal route based on the resulting generalized cost. The randomized approach considerably reduces the requirement that the router be consistent with the scoring function. The randomized Pareto router generates a collection of possible routing solutions; each agent then can select one that best suits its own trade-off between monetary budget and time pressure. Heterogeneity is generated by each synthetic traveler having a different trade-off.

The approach of Horni (2013, also see Chapter 27 of this book) can also be seen in this sense: attaching a random error term to each location-person-pair means that two persons—at exactly the same home location with exactly the same activity pattern—will select different locations for their activities. So far, this describes heterogeneity between persons. However, the approach also generates more heterogeneity per person, since the destinations attractive to each synthetic person will be spatially more spread out than they might otherwise be.

97.3.4 Deliberate Search Strategies

The need for a strategies meta-search, as sketched by Arthur (1994), remains an open question. In the MATSim context, all decisions, based on explicit search for alternatives, can be studied to see how far apart choice set generation strategies of discrete choice modeling (which draw from the universal choice sets), are from explicit construction strategies. One idea would be to observe the second step to see what impact these would have on the results and the policy conclusions.

A good example is parking search (Waraich et al., 2012), for which multiple strategies have been documented and which explains the empirical observations (Shoup, 2005). In a discrete choice model context, distribution of parking preferences can mimic choice strategies, but the approach could not capture the context-specific strategy choice. In MATSim, this set of strategies could become the object of a meta-search to see which agents would retain which strategies and how these would be used by the agents. Empirical work could be conducted to see whether these sets and their distributions match travelers' practices.

In the same vein, one could look at the leisure destination choice, where different strategies can be observed, although they have not yet been subject of empirical study. If longer-term choices were added to the MATSim framework, residential and workplace choice could also be considered.

MATSim plans' convergence towards a single optimal structure can be seen as the absence of search strategies on the plan level. This overlaps strongly with the question of number choice and activities sequence, where these alternative plans are needed.

97.3.5 Transients Versus the Notion of "Learning"

As in other similar simulations, interpretation of the relaxation procedure (iterations) of MATSim is unclear. Sometimes, the relaxation process is ascribed a behavioral interpretation: for example, day-to-day learning, where the transition process, as well as the final equilibrium, has a meaning (Liu et al., 2006, p.128), (Nagel and Barrett, 1997, p.523). An opposite viewpoint exists, where the relaxation procedure is just a numerical method to compute the equilibrium state, or states, without a behavioral basis of the transitions. Although this interpretation ambiguity has not hampered development process so far—also because, in discrete choice modeling, the same ambiguity exists—it is obvious that future questions about adoption of behavioral versus statistical methods require MATSim interpretation.

97.4 Scoring/Utility Function and Choice

97.4.1 Discussion of the Present Scoring Function Mathematical Form

The current logarithmic MATSim activity scoring function,

$$S_{act,q} = \beta_{dur} \cdot t_{typ,q} \cdot \ln(t_{dur,q}/t_{0,q})$$

(cf. Equation (3.2), with $t_{0,q}$ as defined by Equation (3.7), is not suitable for modeling activity addition and dropping (Feil, 2010, p.127f). As already stated in Section 3.3.1, the problem is that, at the typical duration, i.e., at $t_{dur,q} = t_{typ,q}$, all activities generate the same score, independent of their actual duration; thus, it makes sense to first drop the longest activity, since that generates the least amount of utility *per time unit*. This is typically the home or work activity; dropping this first clearly is nonsensical.

The property that all activities have the same utility at their typical duration is obtained by computing the value of the parameter $t_{0,q}$ from the condition[4]

$$const \cdot \beta_{dur} \stackrel{!}{=} S_{act,q}\Big|_{t_{dur,q}=t_{typ,q}} = \beta_{dur} \cdot t_{typ,q} \cdot \ln(t_{typ,q}/t_{0,q}) \tag{97.3}$$

and therefore

$$t_{0,q} = t_{typ,q} \cdot \exp\left(-\frac{const}{t_{typ,q}}\right) \tag{97.4}$$

(cf. Equation (3.7) with $10h \to const$ and $prio \to 1$).

97.4.2 Utility at Typical Duration Proportional to Typical Duration

As an alternative, Equation (97.3) could be replaced by the requirement that all activities at their typical durations yield a score proportional to their typical duration, i.e.,

$$const \cdot \beta_{dur} \cdot t_{typ,q} \stackrel{!}{=} S_{act,q}\Big|_{t_{dur,q}=t_{typ,q}} = \beta_{dur} \cdot t_{typ,q} \cdot \ln(t_{typ,q}/\tilde{t}_{0,q}) \ , \tag{97.5}$$

leading to

$$\tilde{t}_{0,q} = t_{typ,q} \cdot \exp(-const) \ . \tag{97.6}$$

That is, replacing Equation (97.4) by Equation (97.6) in the MATSim scoring function would make, in first order, all activities equally likely to drop. Starting with MATSim release 0.8.x, there will be a config switch

[4] The notation $S\Big|_{x=a}$ means that the expression S shall be evaluated at $x = a$.

```
<param name="typicalDurationScoreComputation" value="..." />
```

where `uniform` will mean the old behavior and `relative` the behavior suggested in this section. Consequences of this still need to be investigated.

97.4.3 S-Shaped Function

A new S-shaped function, proposed by Joh (2004), was tested by Feil (2010, p.129ff). It starts horizontally at zero duration, bends upwards with a positive second derivative and then changes curvature to the normal negative second derivative at longer durations. The function was motivated by the observation that utility functions with infinite (i.e., diverging) first derivative at duration zero lead to "doing a little bit of everything". This is also known from regular consumer theory, with activities replaced by goods. The S-shaped function avoids that problem, instead implying that activities below a certain duration should instead be dropped completely.

Estimates of the new function, based on the Swiss microcensus, were provided; this estimation, however, was difficult, which was attributed to the non-linearities of the function, and to the difficulty in generating sufficiently large choice sets. In addition, many daily activities and their durations were not chosen freely by the individual. Consequently, it is currently *not* advisable to replace the MATSim default scoring function with the Joh/Feil approach.

97.4.4 Heterogeneity of Alternatives and Challenges of Estimation

It is normal to differentiate between types of alternatives in the average; for example, trips by different modes, or with different purposes, are commonly assigned different time values. However, there are also large deviations from those averages between travelers. A possible approach to address this are so-called taste variations, i.e., to make some parameters of the utility function random, but fixed per agent; parameters of this randomness are made part of the choice model estimation. However, some of this apparent randomness may, in fact, be causal. For example, higher values of time for commuting than for leisure may be caused by the more crowded daily schedule on working days. Similarly, the strength of a preference for public transit may be caused by the walking distance to that transit stop serving the desired destination.

Simulation systems such as MATSim should be able to explicitly integrate alternatives' heterogeneity. Besides the aspects discussed in Section 97.4, it is desirable to know how the following aspects influence the scoring function:

- access/egress times to/from public transit,
- transfers between public transit lines,
- crowding in public transit vehicles,
- parking search,
- types of parking (on-street, guarded, sheltered, etc.), and
- personal or household income.

Clearly, this list is not complete.

For most of these aspects, initial studies within the MATSim context are available, see, e.g., Moyo Oliveros and Nagel (2012, in press) for access/egress times and transfers to PT (Public Transport), Bouman et al. (2013), Sun et al. (2014a) or Erath et al. (in preparation) for crowdedness, Waraich et al. (2013b) for parking search, or Kickhöfer et al. (2011) for income. In some cases, it is even possible to configure these elements through the standard config file, completely without Java programming. It is also quite clear that these issues were addressed outside the MATSim context.

A challenge, however, is that it is normally not possible to just collect and combine results from different studies, for the following reasons:

- It is not correct to take an estimated utility function and then change the list of attributes. For example, if walking access/egress to/from PT is not included in the estimation, then its effect may be partially be included in the alternative-specific constant, or in the population density (which may serve as a proxy for the density of PT access points). Just adding the effect of walking access/egress from some other study is thus incorrect.
- Even when MATSim is able *in principle* to add these elements, doing so in practice poses a considerable statistical challenge. For example, one may assume that households inside a zone self-select their precise residential location based on the PT accessibility of their regular destinations. In contrast, a typical MATSim initial demand generation process will first assign residential locations, then generate their destinations, e.g., their workplaces. Thus, persons who might reach their destinations easily by PT might have their MATSim residences far away from the relevant PT stop.

Therefore, it is necessary to estimate the scoring function with exactly those attributes available in the simulation with sufficient precision. Kickhöfer (2009) has, in consequence, re-estimated his scoring function based on data from Vrtic et al. (2008). For the same reasons, it is not possible to combine functions independently estimated for different choice dimensions. This is not even possible when they all contain monetary units. For example, assume that one has

$$\ldots + \beta_t \, t_{trav} + \beta_m \, \Delta m + \ldots$$

for mode choice, and

$$\ldots + \beta_r \, \rho + \tilde{\beta}_m \, \Delta m + \ldots$$

for parking, where t_{trav} is the travel time, ρ is congestion in a parking lot, and Δm is, in all cases, the change in the monetary budget, e.g., cost for gas, PT fare, or parking. Even then, it is not possible to say

$$\ldots + \beta_t \, t_{trav} + \beta_m \, \Delta m + \beta_r \, \frac{\beta_m}{\tilde{\beta}_m} \, \rho + \ldots \, ,$$

since that confuses the scale parameters of the two separate estimations.[5] If only travel time is available as common attribute, the situation deteriorates, since time valuation in MATSim is non-linear; thus, operating points for linearization need to be defined, or found by iterative procedures (Horni, 2013, p.75ff).

As a long-term perspective, one could also imagine estimating choice models directly inside MATSim, possibly taking hints from UrbanSim which has such an approach at its core. An early step in this direction within MATSim using Cadyts (see Chapter 32), is described by Flötteröd et al. (2012).

97.4.5 Agent-Specific Preferences

MATSim scenarios so far consider a relatively small set of agent attributes, essentially because of missing data suitable for deriving detailed large population attributes (Müller and Flötteröd, 2014). Some studies, however, used larger sets of attributes. Grether et al. (2010); Kickhöfer et al. (2011) estimated individual income-contingent utility functions. Horni and Axhausen (2012b,a)

[5] Realistically, combining separate estimations via their conversion in monetary terms may be the best one can do in many situations.

incorporated agent-specific travel preferences and individual income-dependent marginal utilities of money; preference values, however, were assigned randomly. Because consideration of agent-specific preferences is one of the cornerstones of agent-based microsimulations, future work should exploit this avenue.

97.4.6 Frozen Randomness for Choice Dimensions Other Than Destination Choice

For destination choice, an iteration-stable random error term has been successfully applied to incorporate unobserved heterogeneity not included by the stochasticity of the co-evolutionary process (see Chapter 27). Other choice dimensions might also benefit from explicit agent-specific error terms. This could incorporate a mechanism to generate the error terms with the correct correlation structures.

More formally: The current MATSim choice process can be interpreted as maximizing, for each agent n,

$$U_{ni} = V_i + \tilde{V}_{ni} + \boldsymbol{\beta}^T \boldsymbol{\eta}_{\mathbf{ni}} + \tilde{\varepsilon}_{ni} \,, \tag{97.7}$$

where V_i is the systematic utility of alternative i, $\tilde{V}_{ni}$ is an agent-specific addition, $\boldsymbol{\beta}^T \boldsymbol{\eta}_{\mathbf{ni}}$ describes randomness inserted by the network loading model (see Equations (49.4) and (49.5)), and $\tilde{\varepsilon}_{ni}$ is remaining (unexplained) noise. Two challenges are:

- $\tilde{V}_{ni}$ denotes aspects often assumed as random in choice models, but fixed in typical MATSim runs. An example is walking distance to the next PT stop, which may have to be assumed as random in an estimation context based on travel analysis zones, but which is fixed in the context of a MATSim run.

  $\rightarrow$ To be consistent, a choice model and a MATSim implementation used together should use exactly the same disaggregated attributes.
- In most MATSim runs, the $\tilde{\varepsilon}_{ni}$ are either assumed as zero (`BestScore`), or are parameterized by the MATSim choice model (`ChangeExpBeta` or `SelectExpBeta`), which can be interpreted as that the $\tilde{\epsilon}_{ni}$ are re-drawn from the distribution every time a choice is made. This leads, for example, to purely random "logit" switchers between a base and a policy case (e.g., Grether et al., 2010).

  Moreover, the default plans removal (Sections 4.5.1.4 and 97.3.2) has a tendency to remove all alternatives except the best, effectively setting the $\tilde{\varepsilon}_{ni}$ to zero for *all* typical MATSim configurations when run for sufficiently many iterations.

  This is acceptable in situations where most of the noise can be assumed to be in the $\tilde{V}_{ni}$ and/or the $\boldsymbol{\beta}^T \boldsymbol{\eta}_{\mathbf{ni}}$ (and thus generated with hopefully plausible structure by the MATSim dynamics); this may be the case for the choice dimensions of route, mode, and time. It is clearly wrong for locations where $\tilde{\varepsilon}_{ni}$ subsumes preferences that are specific to each person-alternative-pair and that often cannot be included into the $\tilde{V}_{ni}$. For example, a person may have a strong preference for "swimming" in a situation where the data only knows about "leisure" facilities. In this situation, a possible approach is to generate random but "frozen" $\tilde{\varepsilon}_{ni}$, as described in Chapter 27.

  $\rightarrow$ One should thus evaluate how far, and how, a similar approach could be introduced for choice dimensions beyond destination choice.

97.4.7 Economic evaluation

As the above Section 97.4.6 already indicates, further work is desirable to better understand the connection between MATSim scores and utility from consumer theory. At face value, Equation (97.7) could be taken as each agent's utility. As also discussed in Section 51.2.5, problems arise when the $\tilde{\epsilon}_{ni}$ are not explicitly known for each person-alternative-pair *ni*.

In many past MATSim studies, their effect has therefore been parametrized by a logit choice model (with the use of `Change/SelectExpBeta`). In these cases, the logsum of all plans' scores of an agent is each agent's correct utility measure. See Section 51.2.5.1 for details.

In other MATSim studies, the $\tilde{\epsilon}_{ni}$ were effectively assumed to be zero (with the use of `BestScore`). In these cases, the highest of all plans' scores of an agent is each agent's correct utility measure. Because of the `BestScore` plan selection model, the plan with the highest score will at the same time also be the selected plan. See Section 51.2.5.2 for details.

For some of us, it seems attractive to move into the direction of working with frozen randomness, as discussed in Section 97.4.6. That approach would combine the advantages of the two approaches from above: It would inherit the parsimonious interpretation of the `BestScore` approach, where only the plan with highest score (= the selected plan) of each agent needs to be considered, and at the same time include the idea of random utility theory, and, hence, the effect of the ϵ_{ni} on individual choices.

Another avenue of research is to further push the understanding of the econometric and statistical properties of the MATSim choice modeling, cf. Section 51.2.5.5.

97.5 Double-Queue Mobsim

The standard MATSim mobsim QSim implements a single-queue model as described in Chapter 50. The associated FD (flow vs. density) is horizontal for medium densities, and falls to zero very steeply at very high densities. This is consistent with the fact that a vehicle leaving a link opens up its space already in the next time step; jam patterns thus have a backwards traveling speed of $L/1\,s$ (L is the length of respective link) rather than the conventional approx. 15 kilometers per hour (see also Charypar et al., 2009).

The JDEQSim (Section 4.3.2) and the deprecated DEQSim (Section 43.1) implement a double-queue model with backward traveling gaps. Recently, the QSim has also implemented a double-queue variant (Agarwal et al., 2015a), switched on by using a "holes" option in the config; it is, however, not yet thoroughly tested.

97.6 Choice Dimensions, in particular, Expenditure Division

As shown in Section 46.2.2.3 and pictured in Figure 46.1, the Zürich group targets a fuller scheduling model. In addition to standard choice dimensions (printed in red in the cited figure), numerous choices are subject to ongoing research. In particular, "expenditure division" is unexplored not only in MATSim, but in transport planning in general; studies have focused on single-travelers or household-based groups. The field's understanding of both expenditure patterns and allocation styles inside a household are poor, which is no surprise since relevant questions are missing in surveys. First tests for necessary survey works are currently in process and will lead to a better understanding of activity participation and time values that travelers bring to their decisions.

97.7 Considering Social Contacts

Apparently, social contacts, within households as well as within extended social networks, have a substantial influence on travel decisions, particular for social activities in leisure time (Kowald et al., 2009). An early social networks study, in context, but not based on MATSim, is by Marchal and Nagel (2005). Further work based on or, again, in context of MATSim was undertaken by Hackney (2009); Illenberger (2012); Illenberger et al. (2011); Kowald et al. (2009). The most recent work on joint trips is reported in Chapter 28. Despite this range of valuable work, future research is required on this topic, especially for leisure destination choice (Horni, 2013).