A Route Choice Model Based on Mental Representations

Evanthia Kazagli & Michel Bierlaire

Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
École Polytechnique Fédérale de Lausanne

May 28, 2015

Agenda

1. Introduction
2. Methodology
3. Case study
4. Conclusion

Agenda

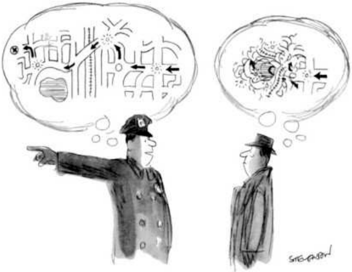
Motivation

Estimation of RUMs[1] with RP[2] data and path assumption is challenging

Operational limitations

- Data
- Choice set
- Structural correlation

Behavioral limitations



[1]Random Utility Models.
[2]Revealed Preferences.

Proposed framework

1. Simple model exploiting RP data
2. Not based on paths
3. Key feature: *mental representations*
4. The general framework may be network-free, yet applicable to traffic assignment

Agenda

Backbone of the framework

A *path* is solely the manifestation of the route choice –the way the traveler implements her decision to take a specific route.

How can we represent a route in a behaviorally realistic way without increasing the model complexity?

- Choice takes place at a higher conceptual level.

$\rightarrow$ Mental Representation Item (*MRI*) = *main modeling element*

Outline of the methodology

1. Definition of the *MRI*:
    1. Empirical evidence through simple qualitative analyzes
    2. Literature review in relevant fields
2. Definition of a RUM model based on *MRI*:
    1. Choice set $\mathcal{C}_n$
    2. Explanatory variables x_{in}, z_n
    3. Specification of the deterministic utility function V_{in}
    4. Assumption about the error terms ε_{in}

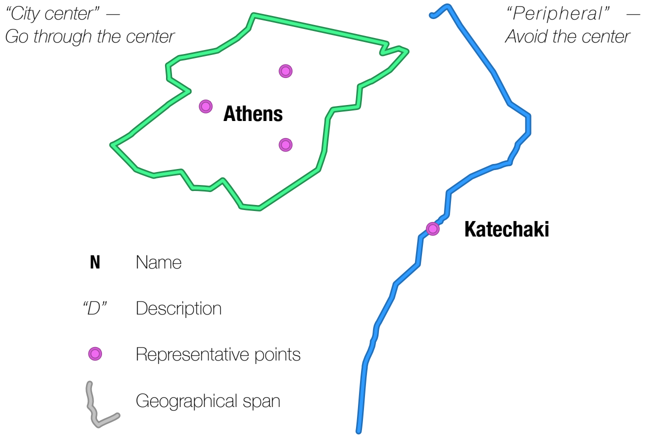
Mental Representation Item (*MRI*)

- *MRIs* are associated with mental representations used in daily language to describe a route.
- An *MRI* is an item characterising the mental representation of an itinerary:

  E.g. a highway, the city center or a bridge.
- Strategic decisions.

The *MRI* components

Perceptual: *a name and a description;* ***Tangible:*** *a point and a span*



Definition of the alternatives

A route is either one-*MRI* or a sequence-of-*MRIs*.

The number of *MRIs* should be kept low so that the number of sequences-of-*MRIs* is also low and can be enumerated.

Issues:

1. How to relate available data to *MRI* alternatives; and
2. How to specify the utility function for the abstract alternatives.

$\rightarrow$ Different heuristics can be considered and evaluated.

From data to *MRIs*

Geographical span

Maximum likelihood estimation:

Let i be an alternative of the *MRI* model, and y an observation, then:

$$\sum_i P(y|i) \cdot P(i|C, x_{in}, z_n)$$

where $P(y|i)$ is the measurement model, $P(i|C, x_{in}, z_n)$ is the choice model.

Associating each piece of data to a single alternative, so that $P(y|i)$ takes values 0 and 1 only, is convenient. For more complex measurement models, we refer to [**?**] and [**?**].

Specification of the utility function

Probably the most complex part

The main modeling element is a mental representation. This has implications for the specification of the utility functions:

! The attributes are fuzzy and based on perceptions rather than objective measurements.

✓ Possibilities to investigate the impact of perception on behavior:

1. Model perceptions –e.g. using latent variables;
2. Network-free approach –e.g. using the level of service of the *MRIs*;
3. Use network data to generate attributes for each *MRI* and specify the utility functions.

Specification of utility functions

Deterministic approach

1. For each *MRI* determine a representative node m (OD dependent).
2. Calculate the fastest path from O to m.
3. Calculate the fastest path from m to D.
4. Use the attributes of the generated path for the *MRI*.

Agenda

Borlänge data

✓ GPS data → `map-matched trajectories`

✓ Borlänge road network:

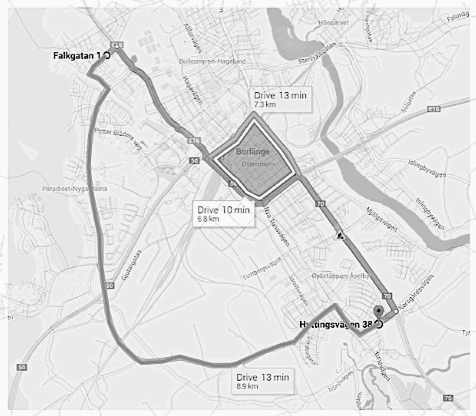
1. 3077 nodes and 7459 unidirectional links
2. Link travel times
3. Clear choices

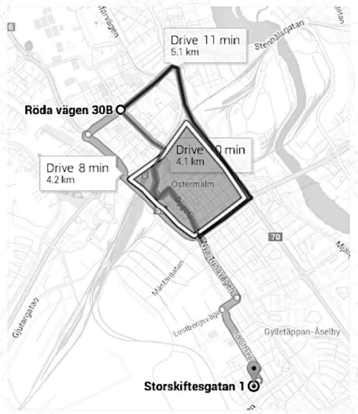
- We use a sample of 139 observations.
- We present one possible way to operationalize the model.

Borlänge road network

Borlänge *MRI* CS

$\mathcal{C}$ ={*1: through the city center (CC),*
2: clockwise movement around the CC,
3: counter-clockwise movement around the CC,
4: avoid the CC}





Example of observed routes (1)

Around the CC movements

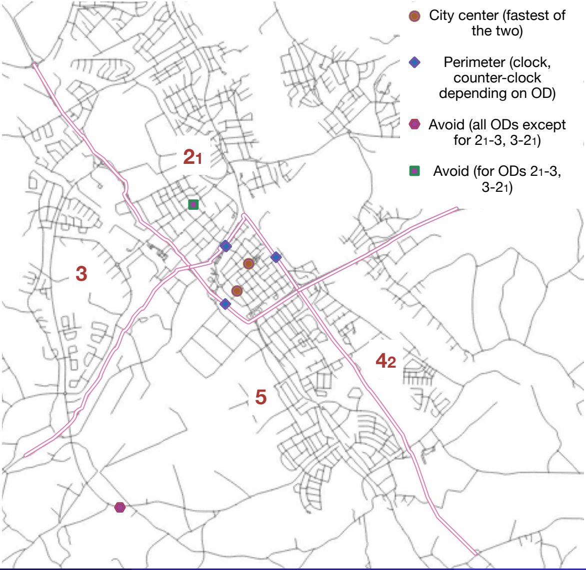
Example of observed routes (2)

Avoid the CC alternatives

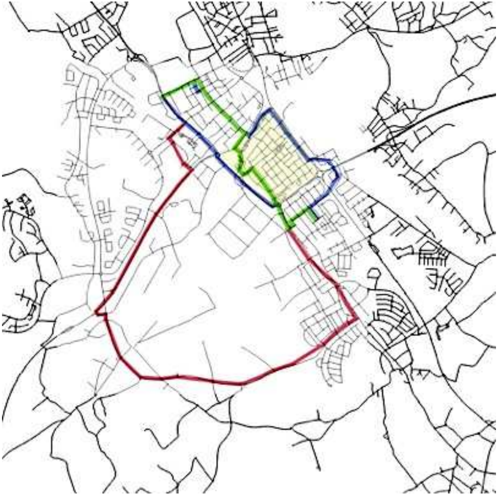
Example of observed routes (3)

Through the CC movements

Representative nodes



Example of *MRI* choice set



——— chosen alternative (through CC)

——— around CC alternatives (clock and counter-clockwise)

——— avoid CC alternative

Estimation results

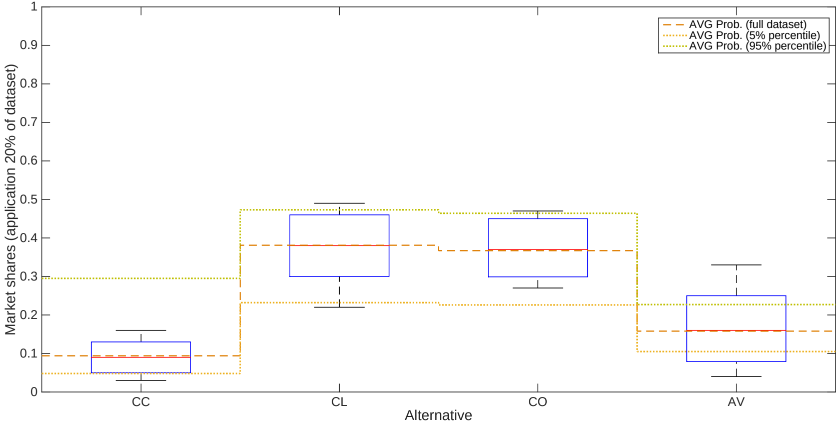
Parameters	**Model** 1 Parameter value; Rob. Std (Rob. *t*-test 0)	**Model 2** Parameter value; Rob. Std (Rob. *t*-test 0)
ASC_{AROUND}	-**2.11**; 1.44; **(-1.47)**	-**0.975**; 1.67; **(-0.58)**
ASC_{AVOID}	**1.87**; 2.09; **(0.89)**	**0.307**; 1.70; **(0.18)**
$\beta\, TIME_{CC}$	-**0.772**; 0.274; **(-2.82)**	
$\beta\, TIME_{AROUND}^{(0-10min)}$	-**0.286**; 0.165; **(-1.74)**	
$\beta\, TIME_{AROUND}^{(>10min)}$	-**0.616**; 0.216; **(-2.86)**	
$\beta\, TIME_{AVOID}$	-**0.583**; 0.187; **(-3.11)**	
$\beta\, LENGTH$		-**0.871**; 0.173; **(-5.03)**
$\beta\, LENGTH_{CC}$		-**1.48**; 0.493; **(-2.99)**
$\beta\, LEFT$	-**0.288**; 0.130; **(2.22)**	-**0.270**; 0.143; **(-1.89)**
$\beta\, IS$	-**0.0474**; 0.022; **(-2.16)**	-**0.063**; 0.018; **(-3.42)**
Number of observations	139	139
Number of parameters	8	6
$\overline{\rho}$	0.375	0.416
$\mathcal{L}(0)$	-183.201	-183.201
$\mathcal{L}(\hat{\beta})$	-106.563	-101.064

Forecasting results (Model 1)

1. Randomly select 80% of the data for estimation.
2. Apply the model in the rest 20%.
3. Repeat 100 times.

$\rightarrow$ Check market shares (MS), predicted probabilities, elasticities.

Boxplot of MS from the application in 20% of the data and CI from the estimation with the full dataset



Agenda

Conclusion

It is possible to have a meaningful model using simple heuristics.

Achievements

- Simple and flexible.
- Behaviorally realistic.

Challenges

- Involved modeling.
- Data processing.

Future steps

1. Traffic assignment.
2. *MRI* sequences and additional complexity → `Quebec GPS dataset`
3. Extention using a multiple-level representation.
4. Other model specifications.

THANK YOU!

Descriptive statistics of the main variables

	mean	median	min	max	std.dev
TT_CC (min)	10.18	8.38	3.88	38.03	6.41
TT_CL (min)	9.98	8.18	2.86	38.93	6.32
TT_CO (min)	10.21	8.37	3.81	36.47	6.23
TT_AV (min)	11.80	13.12	2.66	38.58	11.81
L_CC (km)	7.65	5.21	1.88	42.91	7.39
L_CL (km)	7.84	5.47	1.57	43.82	7.30
L_CO (km)	7.95	5.48	2.33	42.62	7.23
L_AV (km)	9.18	9.04	1.54	42.29	8.90

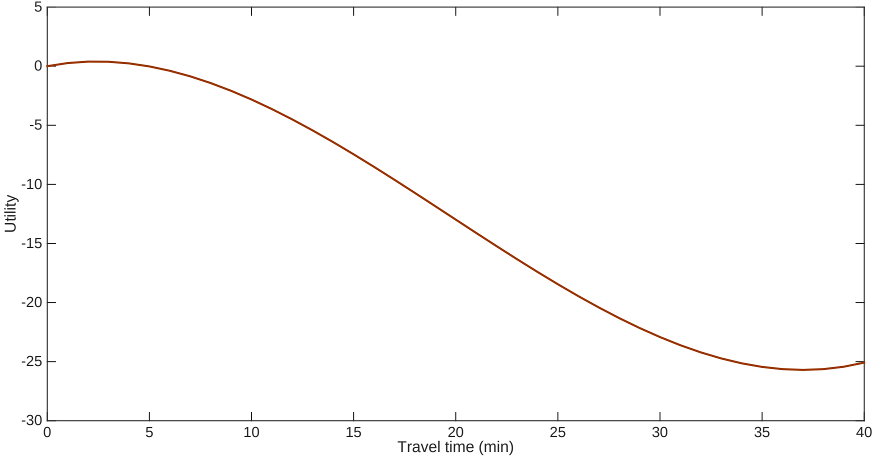
alternative	# times chosen
Through CC	13
Clockwise	53
Counter-clockwise	51
Avoid CC	22

Specification table of model 1

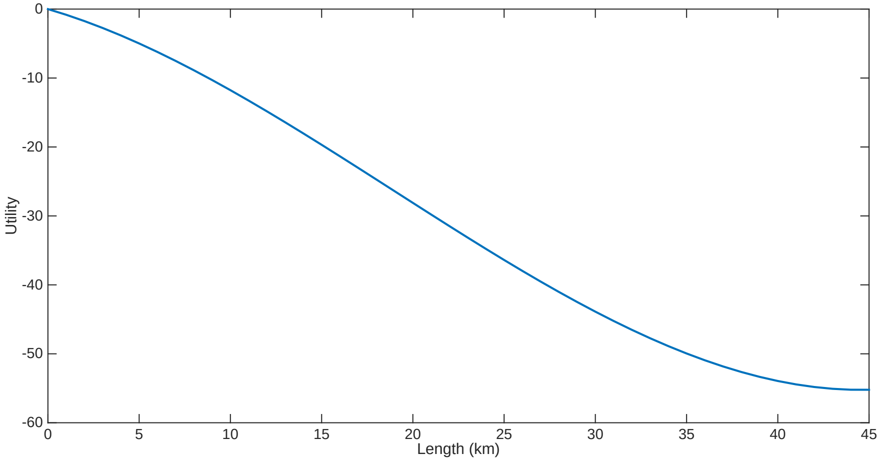
Piecewise linear travel time for the around alternatives

Parameter name	Through CC	Around clock CC	Around counter CC	Avoid CC
ASC_{CC}	0	0	0	0
ASC_{AROUND}	0	1	1	0
ASC_{AVOID}	0	0	0	1
$\beta TIME_{CC}$	TT (min)	0	0	0
$\beta TIME_{AROUND}^{(0-10min)}$	0	TT (min) $\leq$ 10	TT (min) $\leq$ 10	0
$\beta TIME_{AROUND}^{(>10min)}$	0	TT (min) $>$ 10	TT (min) $>$ 10	0
$\beta TIME_{AVOID}$	0	0	0	TT (min)
$\beta LEFT$	# left turns	# left turns	# left turns	# left turns
βIS	# intersections	# intersections	# intersections	# intersections

Power series of degree 3 for the travel time



Power series of degree 3 for the length



Specification table of model 2

Length

Parameter name	Through CC	Around clock CC	Around counter CC	Avoid CC
ASC_{CC}	0	0	0	0
ASC_{AROUND}	0	1	1	0
ASC_{AVOID}	0	0	0	1
$\beta LENGTH_{CC}$	Length (km)	0	0	0
$\beta LENGTH$	0	Length (km)	Length (km)	Length (km)
$\beta LEFT$	# left turns	# left turns	# left turns	# left turns
βIS	# intersections	# intersections	# intersections	# intersections

Application

Traffic assignment

1. Metropolis-Hastings (MH) algorithm [?] to sample paths given the OD and $\mathcal{C}$.

2. The probability of each *path* p to be selected, given the OD and $\mathcal{C}$, is:

   $P(p|\mathcal{C}) = \sum_i P(p|i) \cdot P(i|\mathcal{C})$

   where the sum spans the alternatives in the *MRI* models, $P(i|\mathcal{C})$ is the *MRI*-choice model, and $P(p|i)$ is the probability of path p to be actually used by a traveler who has chosen the sequence of *MRIs* i.

Application

Route guidance

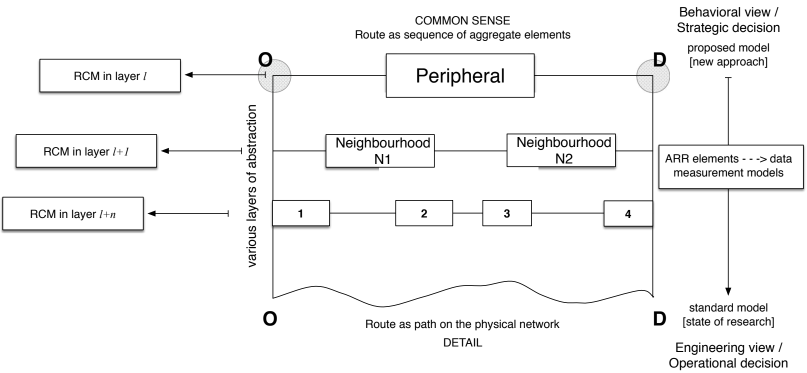
Provision of information in an aggregate manner:

1. Guidance on VMS[3]
2. Radio announcements
3. Oral instructions in in-vehicle navigation systems

[3] Variable message signs.

Hierarchical ordering of the decision process

Multi-level hierarchical structure ∼Normative Pedestrian Flow Theory [?]



Model structure

Layer ℓ

- Choice set: list of *MRIs* $\mathcal{C}_\ell$.
- Choice model:

$$P_\ell(i|\mathcal{C}_\ell; \beta^\ell)$$

Layer $\ell + 1$

- Choice set: list of *MRIs* $\mathcal{C}_{\ell+1}$.
- Choice model:

$$P_{\ell+1}(i|\mathcal{C}_{\ell+1}; \beta^{\ell+1})$$

Behavioral consistency

- All layers refer to the same choice.
- Level of granularity varies.
- Analysis can be performed in any layer.

Structural consistency

$$\bar{P}_\ell(i|\mathcal{C}_\ell; \beta^\ell) = \sum_{j \in \mathcal{C}_{\ell+1}} P(i|j, \mathcal{C}_\ell; \beta^\ell) P(j|\mathcal{C}_{\ell+1}; \beta^{\ell+1})$$