Symulacje złożonych systemów społecznych
modelowanie zachowania ludzkiego

@ GMUM, Faculty of Mathematics and Computer Science, Jagiellonian University, Kraków
Rafał Kucharski
rafal.kucharski@uj.edu.pl
https://rafal-kucharski.u.matinf.uj.edu.pl/

Agenda

this talk

Idea

1. What computers (and Computer Scientists) usually do? optimization
2. What human individuals usually do? decisions
3. What groups of humans (society) usually do? society.

Cases

- discrete choice models - Multinomial Logit Model
- social networks - Behavioural Profiling
- traffic flow models - Traffic Microsimulation

Leitmotif

While CS is purely STEM, now the big-data is also in social science, opening widely for novel and exciting CS applications

European Research Council
Established by the European Commission

Agenda
this talk

Idea

1. What computers (and Computer Scientists) usually do? optimization
2. What human individuals usually do? decisions
3. What groups of humans (society) usually do? society.

Cases

- discrete choice models - Multinomial Logit Model
- social networks - Behavioural Profiling
- traffic flow models - Traffic Microsimulation

Leitmotif

While CS is purely STEM, now the big-data is also in social science, opening widely for novel and exciting CS applications

Agenda

this talk

Idea

1. What computers (and Computer Scientists) usually do? optimization
2. What human individuals usually do? decisions
3. What groups of humans (society) usually do? society.

Cases

- discrete choice models - Multinomial Logit Model
- social networks - Behavioural Profiling
- traffic flow models - Traffic Microsimulation

Leitmotif

While CS is purely STEM, now the big-data is also in social science, opening widely for novel and exciting CS applications

Agenda

this talk

Idea

1. What computers (and Computer Scientists) usually do? optimization
2. What human individuals usually do? decisions
3. What groups of humans (society) usually do? society.

Cases

- discrete choice models - Multinomial Logit Model
- social networks - Behavioural Profiling
- traffic flow models - Traffic Microsimulation

Leitmotif

While CS is purely STEM, now the big-data is also in social science, opening widely for novel and exciting CS applications

Agenda
this talk

Idea

1. What computers (and Computer Scientists) usually do? optimization
2. What human individuals usually do? decisions
3. What groups of humans (society) usually do? society.

Cases

- discrete choice models - Multinomial Logit Model
- social networks - Behavioural Profiling
- traffic flow models - Traffic Microsimulation

Leitmotif

While CS is purely STEM, now the big-data is also in social science, opening widely for novel and exciting CS applications

myself

Rafał Kucharski

now: associate. prof, Jagiellonian University, Faculty of Math. and CompSci, GMUM, prof. Jacek Tabor

2023-2028 ERC Starting Grant - COeXISTENCE 3 PhDs + PostDoc; Reinforcement Learning

2023-2026 Horizon Europe - SUM 2 PhDs + PostDoc; Transport Planning

2021-2024 NCN OPUS - Post-corona shared mobility 2 PhDs + PostDoc; Network Science+Optimisation

past: PostDoc @ TU Delft working in Critical MaaS ERC Starting Grant

- shared rides algorithms ExMAS
- agent based model MaasSim

past^2: assist. prof @ Politechnika Krakowska, prof. Andrzej Szarata

PhD: DTA, La Sapienza Rome, prof. Guido Gentile

outside academia:

- R&D software developer (PTV SISTeMA, Rome)
- transport modeller (models for Kraków, Warsaw and more)
- data scientist, ML engineer (NorthGravity)

Humans and computers

Optimisation

finding optimal value

Problem

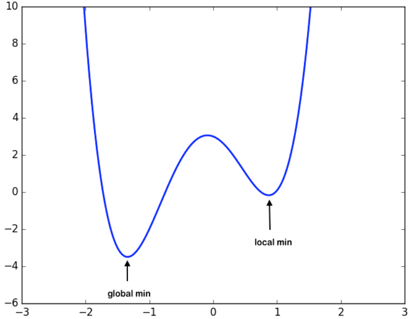
What is the minimal value of the function?

$$\underset{x}{\arg\min}\, f(x) \tag{1}$$

where

- x is the vector of decision variables,
- $f(x)$ is the objective function to be minimized.

both are deterministic.



Optimisation

finding optimal value

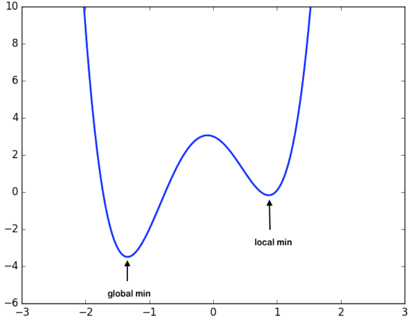
Optimisation

can be:

- policriterial: $x \longrightarrow x, y, \ldots,$
- stochastic: $f(x) \longrightarrow f(x+\epsilon)$,
- multivariate: $f(x) \longrightarrow f(x, y, \ldots)$,
- black-box (like neural network)

Computer Science

We can reduce bigger part of CS to solving optimization problems.



Optimisation

finding optimal value

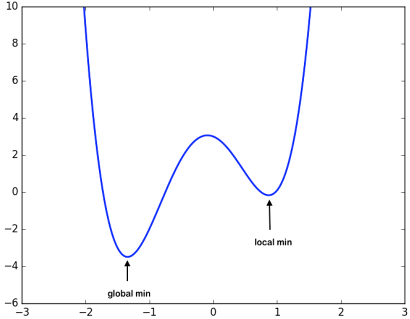
Optimisation

can be:

- policriterial: $x \to x, y, \ldots,$
- stochastic: $f(x) \to f(x + \epsilon)$,
- multivariate: $f(x) \to f(x, y, \ldots)$,
- black-box (like neural network)

Computer Science

We can reduce bigger part of CS to solving optimization problems.



Discrete Choice

Path choice

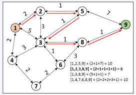
Problem

Given a weighted network $G(N, A)$ find a path (sequence of nodes $n \in N$) from origin o to destination d

Computers

Shortest Path Choice

Define objective function (e.g. distance or more generically a cost $c(a)$: $a \in A$) and propose an algorithm to find a solution.
e.g. Dijkstra - which deterministically and reliably outputs an optimal path.



Humans

Discrete Choice

Each agent i selects the optimal path k from her origin o_i to destination d_i at her departure time τ:

$$k_{od,i} = \arg\min_{k \in K_{od}} \sum_{a \in k} c_{a,i} \tag{2}$$

Discrete Choice

Path choice

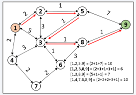
Problem

Given a weighted network $G(N, A)$ find a path (sequence of nodes $n \in N$) from origin o to destination d

Computers

Shortest Path Choice

Define objective function (e.g. distance or more generically a cost $c(a) : a \in A$) and propose an algorithm to find a solution.
e.g. Dijkstra - which deterministically and reliably outputs an optimal path.



Humans

Discrete Choice

Each agent i selects the optimal path k from her origin o_i to destination d_i at her departure time τ:

$$k_{od,i} = \underset{k \in K_{od}}{\arg\min} \sum_{a \in k} c_{a,i} \tag{2}$$

Behaviour

Discrete Choice

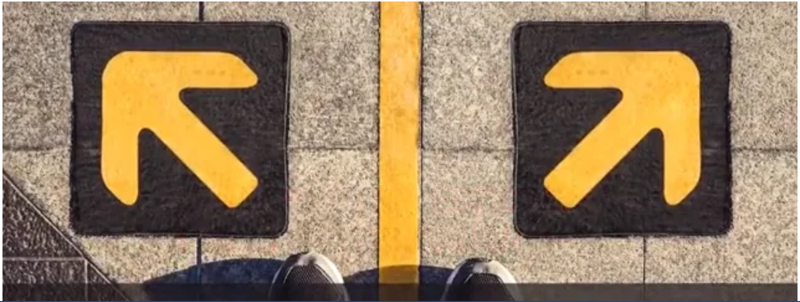
Example

Problem

There are two products.

1. cheap, nice and low quality
2. expensive, ugly and high quality

which is optimal?



Rational utility maximisers

in path choice

Rational

Let's assume all humans are rational:

$$\Pr(k|od, i) = \Pr\left(c_{k,i} = \min_{k' \in K_{od}} c_{k',i}\right)$$

i.e. we take the best option.

Costs

Each path candidate has a given:

- length
- travel time
- cost (fare)
- comfort factor
- ...

Perceived costs

utility

length and travel time are physical
cost is subjective, in discrete choice called Utility

$$U_{k,i} = \beta_{0,i} + \beta_{t,i} t_k + \beta_{c,i} c_k + \cdots + \varepsilon$$

- β_0 alternative-specific constant, i.e. taste variation, i.e. sentiment
- ε random term
- β_t value of time (10€/h)
- β_c value of money

Rational utility maximisers

in path choice

Rational

Let's assume all humans are rational:

$$\mathtt{Pr}(k|od, i) = \mathtt{Pr}\left(c_{k,i} = \min_{k' \in K_{od}} c_{k',i}\right)$$

i.e. we take the best option.

Costs

Each path candidate has a given:

- length
- travel time
- cost (fare)
- comfort factor
- ...

Perceived costs

utility

length and travel time are physical
cost is subjective, in discrete choice called Utility

$$U_{k,i} = \beta_{0,i} + \beta_{t,i} t_k + \beta_{c,i} c_k + \cdots + \varepsilon$$

- β_0 alternative-specific constant, i.e. taste variation, i.e. sentiment
- ε random term
- β_t value of time (10€/h)
- β_c value of money

Discrete choice theory

Key concepts

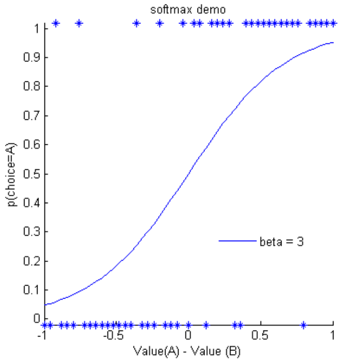
Non-determinism

we can reasonably well predict the probability of selecting an option a by individual i, yet there is always non-determinism.
Probabilities only asymptotically approach to 0 and 1.

Heterogeneity

We are different, each of us has its' own:

- $\beta_{0,i}$ alternative-specific constant, i.e. taste variation, i.e. sentiment
- ε random term
- $\beta_{t,i}$ value of time
- $\beta_{c,i}$ value of money



Discrete choice theory

Nobel prize

Daniel McFadden won the Nobel prize in 2000 for his pioneering work in developing the theoretical basis for discrete choice.

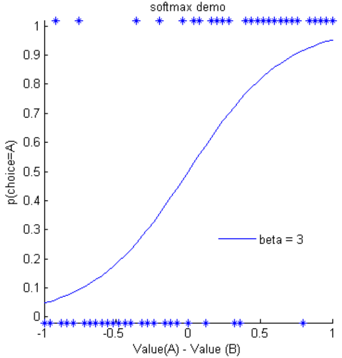
Discrete choice theory

Discrete choice models statistically relate the choice made by each person to the attributes of the person and the attributes of the alternatives available to the person.

Logit model

assumption:
$\varepsilon \approx \text{Gumbel}(0, \sigma)$, yields
Probability of selecting option a in the choice set C by individual i is:

$$p_{a,i} = \frac{\exp \mu U_{a,i}}{\sum_{a' \in C} \exp \mu U_{a',i}}$$



Estimation

Bigdata

Datasets

obsID	personID	panelID	choice	ivttWalk	ivttBike	ivttCar	ivttTransit	ovttWalk	ovttBike	ovttCar	ovttTransit	costWalk	costBike	costCar	costTransit	betaIvtt	betaOvtt
1	1	1	3	0	0	58	72	96	109	10	12	0	0	5	1	-0.373273721	-1.066419469
2	1	2	3	0	0	38	42	166	55	0	15	0	0	7	3	-0.62497503	-1.119626123
3	1	3	3	0	0	56	65	145	63	1	16	0	0	3	1	-0.643188316	-1.174059586
4	1	4	3	0	0	19	20	106	37	9	15	0	0	4	1	-0.438671827	-1.24832442
5	1	5	3	0	0	54	81	185	41	2	19	0	0	5	3	-0.287124529	-0.844659563
6	1	6	3	0	0	41	35	48	30	8	22	0	0	3	3	-0.257752721	-0.55035108
7	1	7	3	0	0	27	33	106	25	3	13	0	0	2	0	-0.569873118	-0.764597732
8	1	8	4	0	0	18	21	163	41	8	12	0	0	6	1	-0.369689557	-1.261403142
9	1	9	3	0	0	24	22	66	42	10	16	0	0	3	3	-0.096837917	-1.025072048
10	1	10	3	0	0	14	17	35	27	5	24	0	0	3	2	-0.191661813	-0.63550085
11	2	1	1	0	0	13	11	0	0	4	19	0	0	1	2	-0.02980754	-0.397523895
12	2	2	3	0	0	43	49	135	31	4	14	0	0	4	4	-0.22930545	-0.915244977
13	2	3	4	0	0	50	42	1142	84	9	11	0	0	6	1	-0.491108147	-0.627612196
14	2	4	1	0	0	22	23	18	20	6	21	0	0	5	2	-0.484219256	-0.889211994
15	2	5	3	0	0	18	18	61	29	2	13	0	0	4	0	-0.617233817	-1.441170191
16	2	6	2	0	0	17	21	167	21	10	15	0	0	3	1	-0.136576508	-0.828500671
17	2	7	3	0	0	34	42	179	63	3	19	0	0	5	2	-0.427847708	-1.014582039
18	2	8	3	0	0	51	51	84	37	0	16	0	0	6	1	-0.433200047	-1.418900935
19	2	9	4	0	0	44	35	531	177	5	15	0	0	9	3	-0.524877465	-0.971677976

Binary classifier

Predict the binary (0/1 value)
d6922a778401

Machine Learning

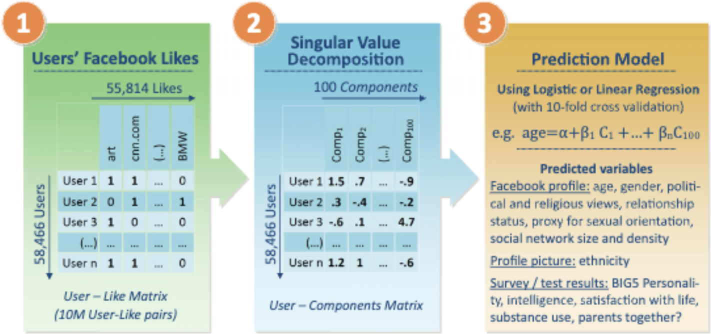
Lately, instead of classical methods (like `BIOGEME's` max log-likelihood) neural networks are used to classify choices - still in infancy.

European Research Council
Established by the European Commission

Predict behaviour from digital traces

Internet privacy

What Facebook likes tell us about us?[1]



[1] Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behaviour. Proceedings of the national academy of sciences, 110(15), 5802-5805.

Internet privacy

What Facebook likes tell about us?[2]

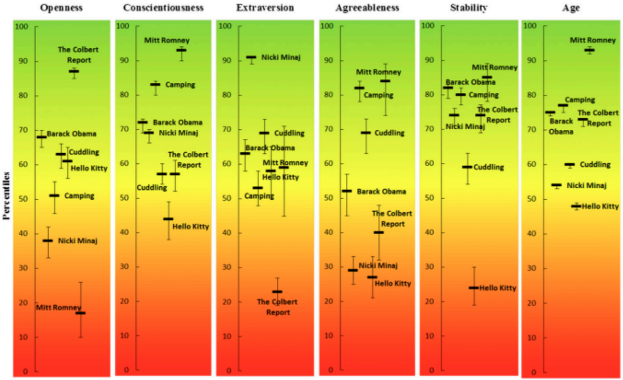


Fig. S1. Average levels of five personality traits and age of the users associated with selected Likes presented on the percentile scale. For example, the average extraversion of users associated with "The Colbert Report" was relatively low: it was lower only for 23% of other Likes in the sample. Error bars signify 95% confidence intervals of the mean.

[2] Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behaviour. Proceedings of the national academy of sciences, 110(15), 5802-5805.

European Research Council
Established by the European Commission

Traffic flow

Phantom jam

Let's drive around the circle at constant speed

Phantom jam

Let's drive around the circle at constant speed

Video

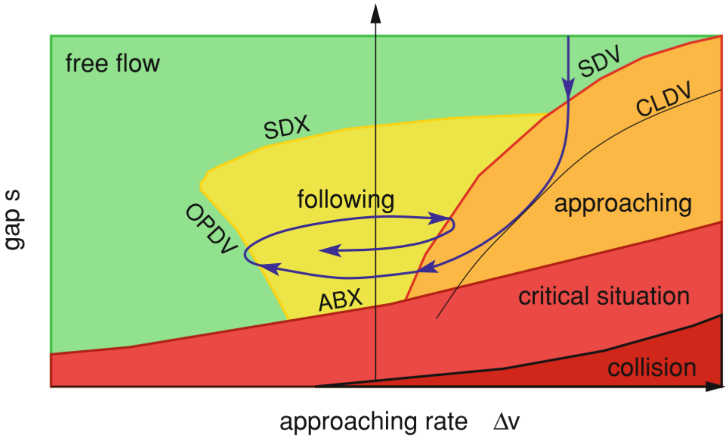
https://youtu.be/FW9VkoibWDw?si=a0qexb-zSMxPxLwY&t=25

What can go wrong?

Why there was a traffic breakdown?
Why we couldn't do such an easy task and led to the phantom jam?

Car following Model

Wiedemann



Microsimulation

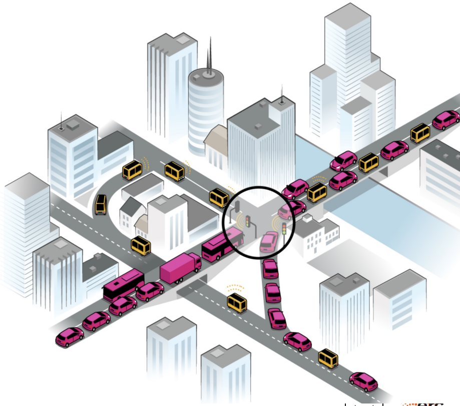
PTV Vissim, SUMO, Aimsun, ...

https://www.youtube.com/watch?v=bqF-Hyovg9E&t=3s

Thank you!

Thank you for your attention,

welcome to discuss



Rafał Kucharski

rafal.kucharski@uj.edu.pl

Acknowledgements

This research was funded by the COeXISTENCE project (grant number 101075838), which is financed by the European Research Council.