Behavior Modeling in Transportation Networks

Lecture Series #3-1 (16:00-16:30)

Reinforcement Learning and Network Design

Yuki Oyama

Shibaura Institute of Technology

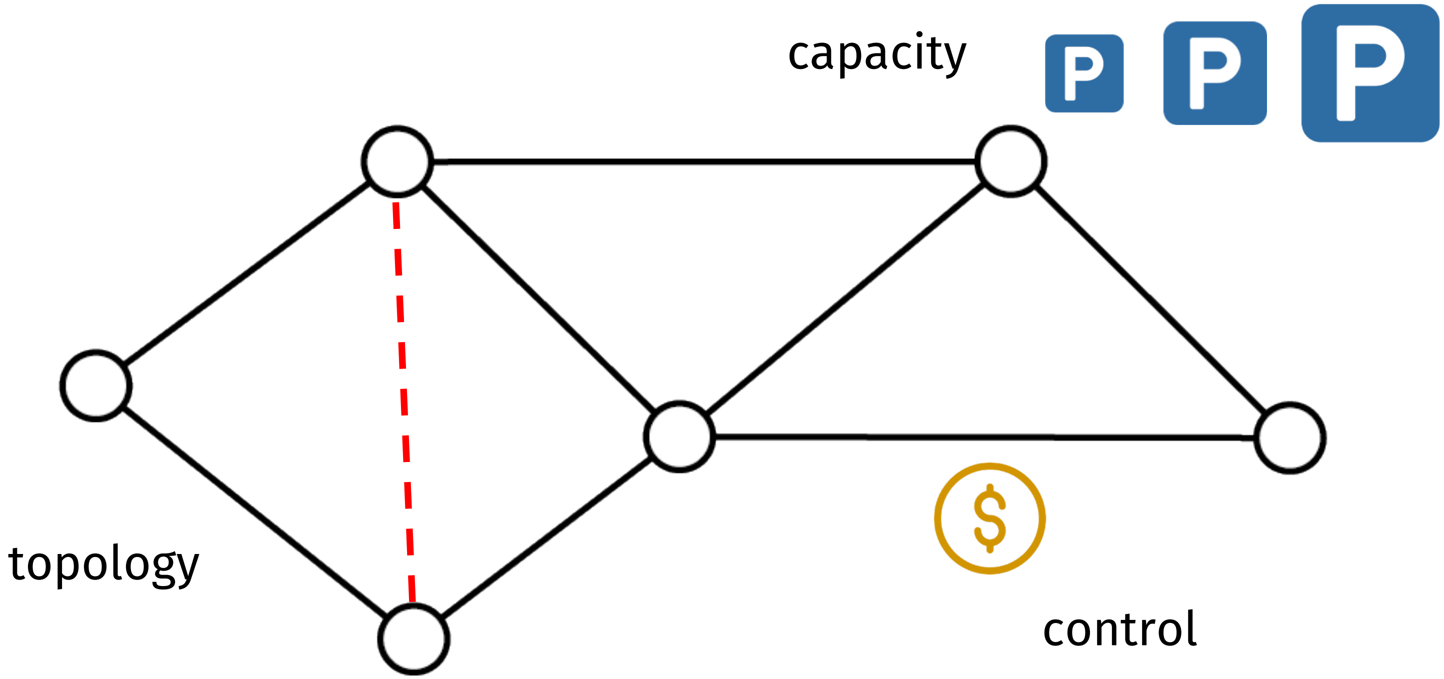
Activity Landscape Design Lab.

September 17, 2021

A road network example

The planner who aims **to maximize efficiency** wants to answer:

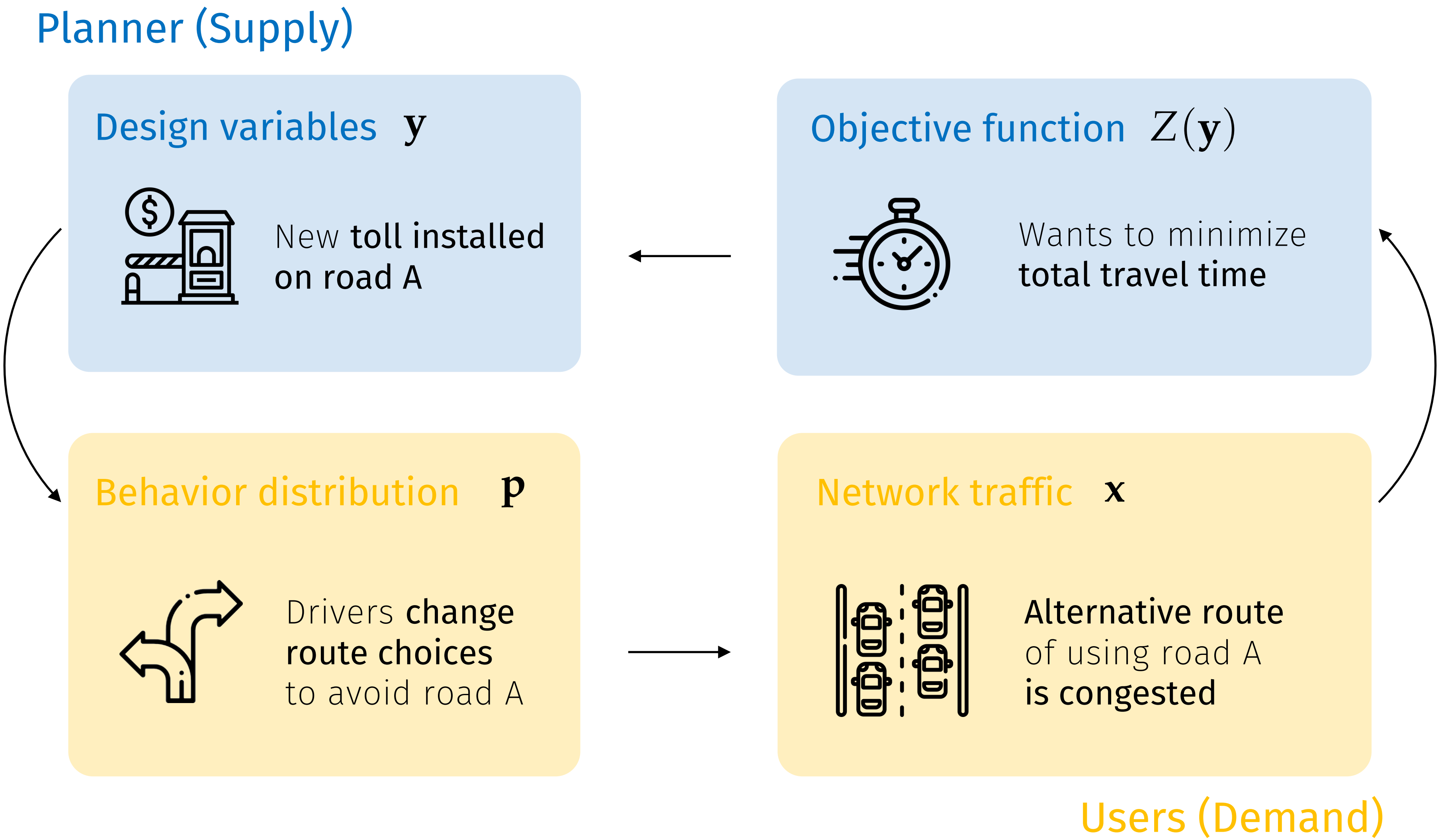
- **if a new road** should be constructed
- **where and how large parking spaces** should be placed
- **on which road and how much tolls** should be charged
- etc.



These decisions will impact on travelers' behavior

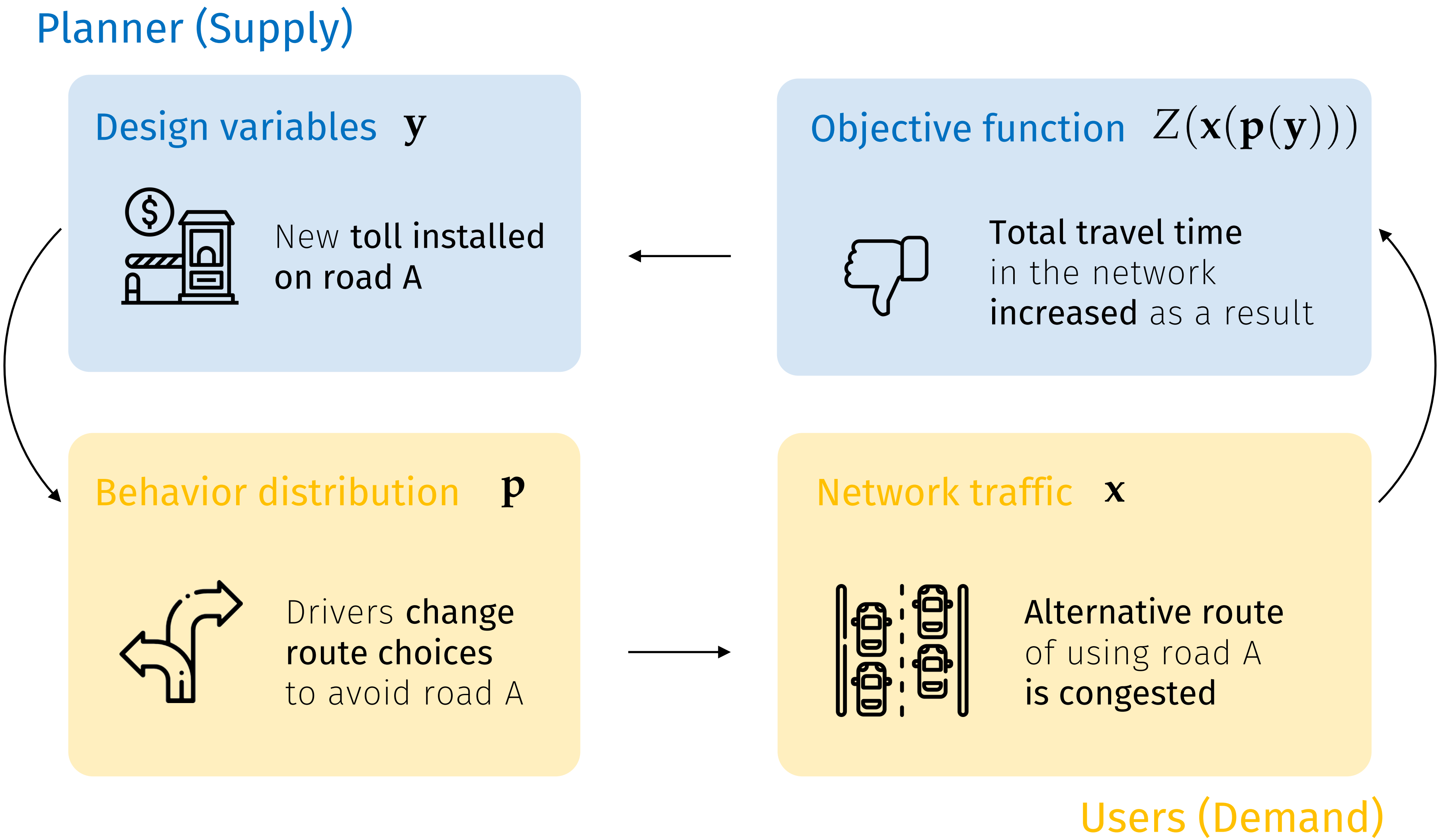
Let's generalize the framework

An example of pricing (on which road a toll is installed)



Let’s generalize the framework

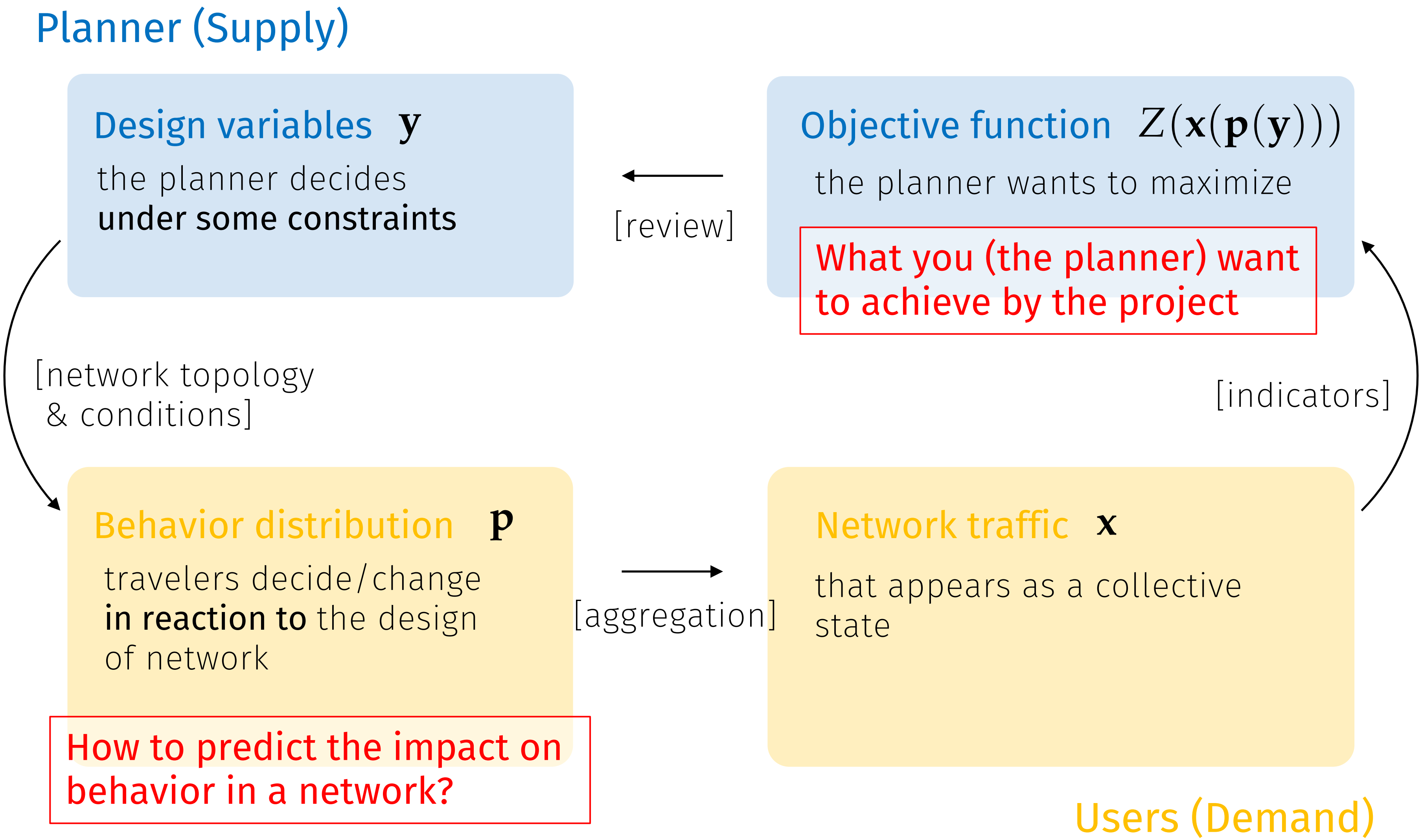
An example of pricing (on which road a toll is installed)



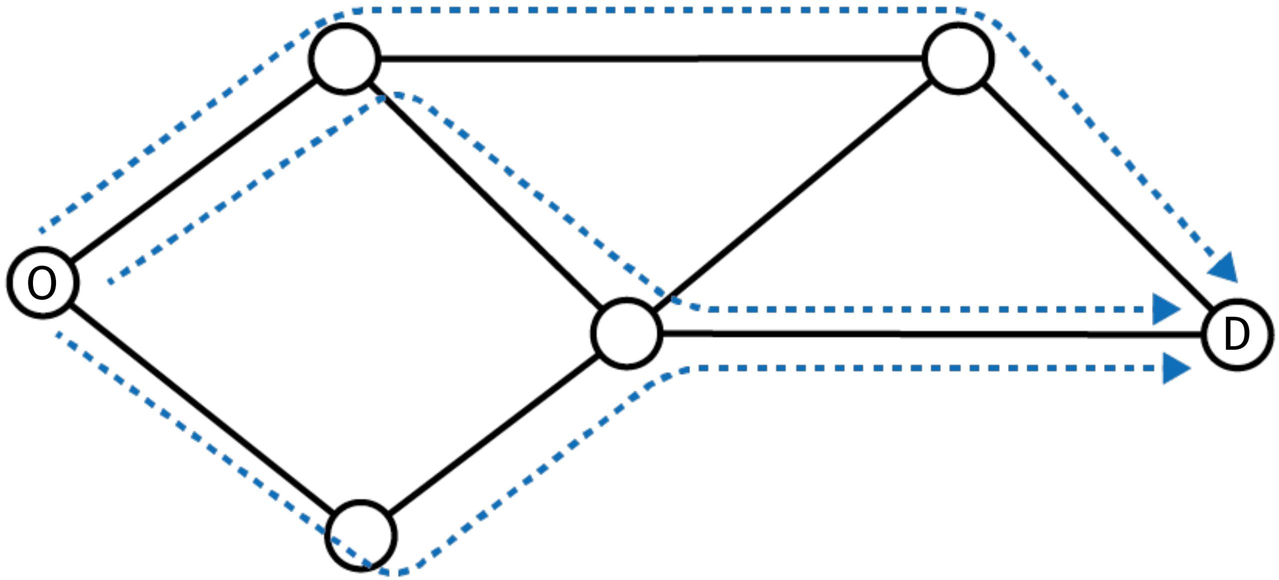
Network design

is a demand-based planning of network topology & systems

Follows Magnanti and Wong (1984); Farahani et al. (2013)



Modeling behavior in a network



Path choice model (logit)

$$P(r) = \frac{e^{\mu v_r}}{\sum_{r' \in \mathcal{R}} e^{\mu v_{r'}}}$$

$\mathcal{R}$: choice set (set of paths)

⟶

Traffic flow on paths
(in the static case)

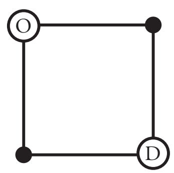
$$x_r = q_{od} P(r) \quad \forall r \in \mathcal{R}$$

Not as easy as it looks...

Networks are generally complex…

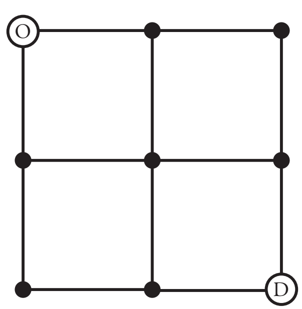
The path set $\mathcal{R}$ is almost impossible to define !!

$k = 1$



$|\mathcal{R}| = 2$

$k = 2$



$|\mathcal{R}| = 12$

k	simple paths
1	2
2	12
3	184
4	8,512
5	1,262,816
6	575,780,564
7	789,360,053,252
8	3,266,598,486,981,640
9	41,044,208,702,632,496,804
10	1,568,758,030,464,750,013,214,100
11	182,413,291,514,248,049,241,470,885,236
12	64,528,039,343,270,018,963,357,185,158,482,118
13	69,450,664,761,521,361,664,274,701,548,907,358,996,488
14	227,449,714,676,812,739,631,826,459,327,989,863,387,613,323,440
15	2,266,745,568,862,672,746,374,567,396,713,098,934,866,324,885,408,319,028

This is because **a path is a combination of links** in the network

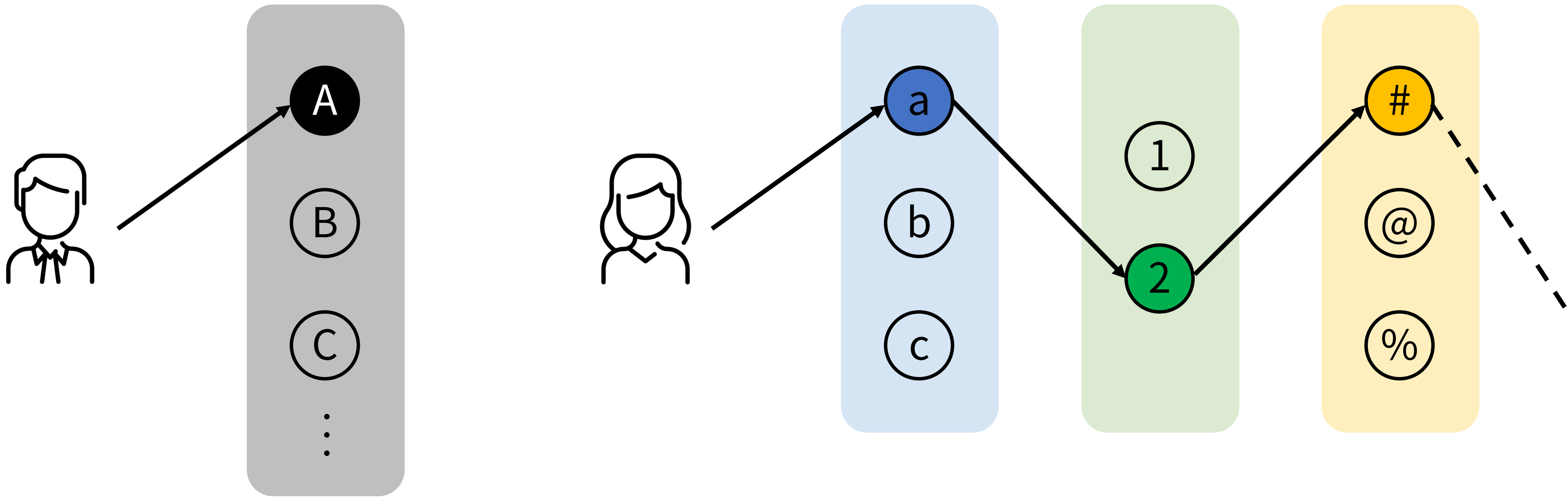
* A description of more complex choices (e.g., time) needs additional dimensions of network, which further increases the network size.

Not **combination** but **SEQUENCE**

An approach is modeling based on Reinforcement Learning (RL)

that models sequential decisions of agents.

$$A = [a, 2, \#, \ldots]$$



Choice

Sequence of choices

This presentation shows a special case of RL **for network path choice modeling**

How to model a sequence ?

A path $\mathbf{r}$ can be described as:

$$r = \underbrace{[a_1, a_2, \ldots, a_J]}$$

a sequence of links

Path choice probability:

$$P(r) = \prod_{j=1}^{J-1} p(a_{j+1}|a_j)$$

$p(a_{j+1}|a_j)$: Link choice probability **conditional on the previous link**

$\Rightarrow$ **what is link choice probability** exactly?

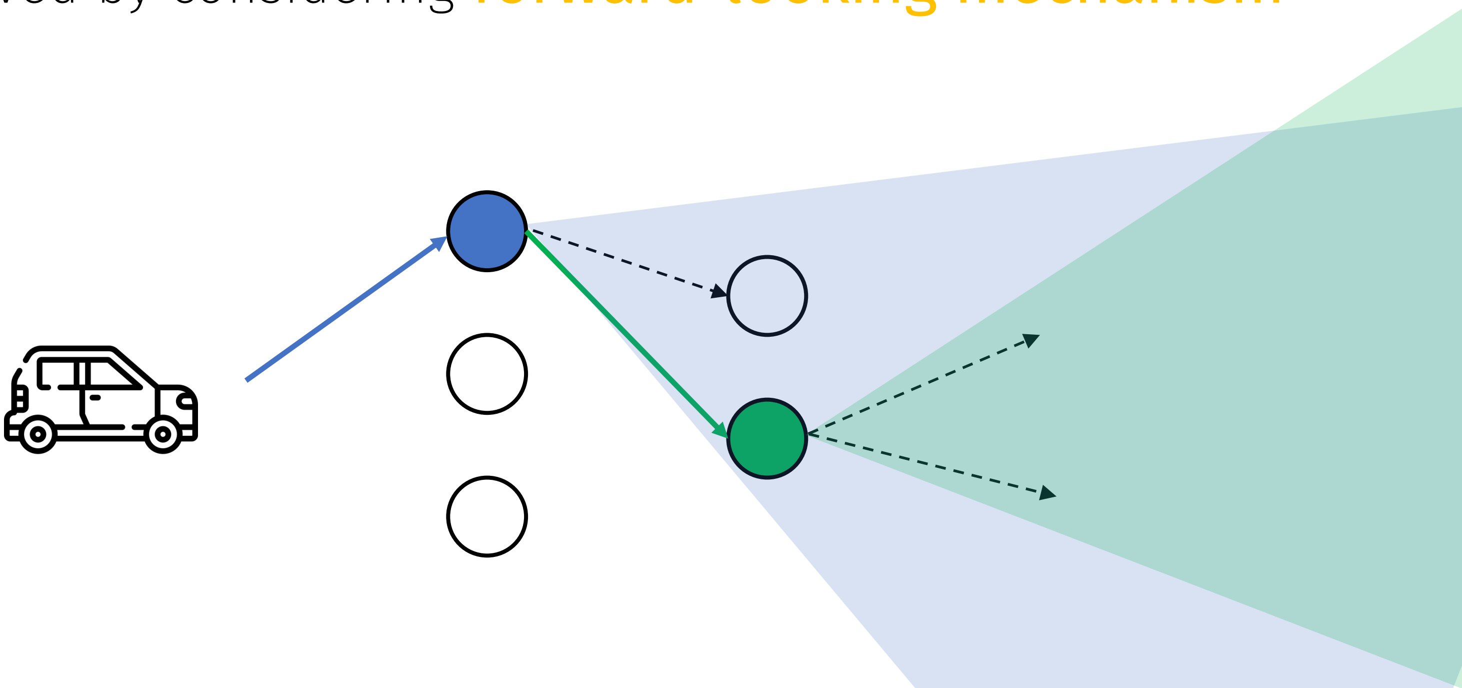
What should be considered is …

the outcome given by the product of link choice probabilities to be **consistent with the original model**, i.e.,

$$P(r) = \prod_{j=1}^{J-1} p(a_{j+1}|a_j) = P_{\text{Logit}}(r)$$

*when assuming logit model

This is achieved by considering forward-looking mechanism

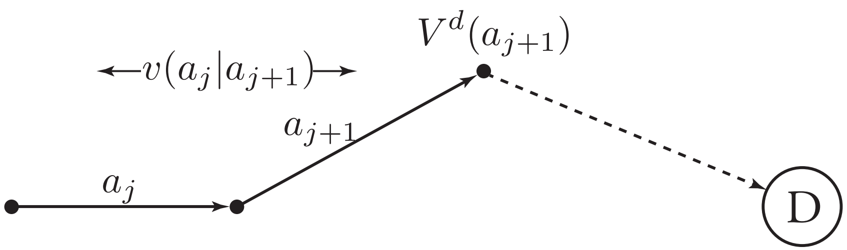


Value function

Goal is modeling

1. Myopic
2. Forward-looking

mechanisms of behavior



v : Link choice utility

V^d : Value function

$$V^d(a_j) = \mathbb{E}\left[\max_{a_{j+1}\in\mathcal{A}(a_j)} \{v(a_{j+1}|a_j) + \varepsilon(a_{j+1}|a_j) + V^d(a_{j+1})\}\right]$$

Random utility

Generalization

c.f. **Shortest Path (SP) problem**:

$$V^d(a_j) = \max_{a_{j+1}\in\mathcal{A}(a_j)} \{v(a_{j+1}|a_j) + V^d(a_{j+1})\}$$

Value function is the **SP cost** from a_j to destination

Gumbel distribution has a nice property:

$$\varepsilon_k \overset{\text{i.i.d.}}{\sim} \text{Gumbel}(0,\mu), \forall k \quad \Rightarrow \quad \max_k \{\eta_k + \varepsilon_k\} \sim \text{Gumbel}(\frac{1}{\mu} \ln \sum_k \mu\eta_k, \mu)$$

Value function is the solution to:

$$V^d(a_j) = \mathbb{E}\left[\max_{a_{j+1}\in\mathcal{A}(a_j)} \{v(a_{j+1}|a_j) + \varepsilon(a_{j+1}|a_j) + V^d(a_{j+1})\}\right]$$

$$= \frac{1}{\mu} \ln \sum_{a_{j+1}\in\mathcal{A}(a_j)} e^{\mu\{v(a_{j+1}|a_j)+V^d(a_{j+1})\}}$$

$$\Leftrightarrow \quad e^{\mu V^d(a_j)} = \sum_{a_{j+1}\in\mathcal{A}(a_j)} e^{\mu v(a_{j+1}|a_j)} e^{\mu V^d(a_{j+1})}$$

a system of linear equations.

(Recurrence relation)

$$\Rightarrow \quad \mathbf{z}^d = \mathbf{W}\mathbf{z}^d + \mathbf{e}^d$$

$\mathbf{z}^d \equiv [e^{\mu V_k^d}]_{k\in\mathcal{L}}$	$\mathbf{W} \equiv [e^{\mu v(l\|k)}]_{k,l\in\mathcal{L}}$	$\mathbf{e}^d \equiv [\delta_k^d]_{k\in\mathcal{L}}$
Value function	Weight incidence matrix	Unit vector

Let's check the consistency!

Link choice probability is given by:

$$p^d(a_{j+1}|a_j) = \frac{e^{\mu\{v(a_{j+1}|a_j)+V^d(a_{j+1})\}}}{\sum_{a_{j+1}\in\mathcal{A}(a_j)} e^{\mu\{v(a_{j+1}|a_j)+V^d(a_{j+1})\}}} = \frac{W(a_{j+1}|a_j)z^d(a_{j+1})}{z^d(a_j)}$$

*like logit by assuming $U(a_{j+1}|a_j) = \underbrace{v(a_{j+1}|a_j) + V^d(a_{j+1})}_{\text{New deterministic utility}} + \varepsilon(a_{j+1}|a_j)$

Then we have:

$$P^{od}(r) = \frac{W(a_1|o)z^d(a_1)}{z^d(o)} \cdot \frac{W(a_2|a_1)z^d(a_2)}{z^d(a_1)} \cdot \dots \cdot \frac{W(d|a_J)z^d(d)}{z^d(a_J)}$$

$z^d(d) = 1$

$$= \frac{\prod_{j=0}^{J} W(a_{j+1}|a_j)}{z^d(o)} = \frac{e^{\mu\sum_{j=0}^{J} v(a_{j+1}|a_j)}}{e^{\mu V^d(o)}} = \frac{e^{\mu v_r}}{\sum_{r'\in\mathcal{R}^{od}} e^{\mu v_{r'}}}$$

Path utility is sum of link utilities

Exp. Max. of **all possible paths**

$$= P_{\text{Logit}}(r|\mathcal{R}^{od})$$

⇒ **Consistent with logit** model with the *universal* path set

What's the point ?

Now you can model path choice behavior
without explicitly defining choice set

k	simple paths
1	2
2	12
3	184
4	8,512
5	1,262,816
6	575,780,564
7	789,360,053,252
8	3,266,598,486,981,640
9	41,044,208,702,632,496,804
10	1,568,758,030,464,750,013,214,100
11	182,413,291,514,248,049,241,470,885,236
12	64,528,039,343,270,018,963,357,185,158,482,118
13	69,450,664,761,521,361,664,274,701,548,907,358,996,488
14	227,449,714,676,812,739,631,826,459,327,989,863,387,613,323,440
15	2,266,745,568,862,672,746,374,567,396,713,098,934,866,324,885,408,319,028

No longer needed!

1. Decompose path choice into **sequential link choices:**

$$P(r) = \prod_{j=1}^{J-1} p(a_{j+1}|a_j)$$

2. Describe forward-looking behavioral mechanism by **value function:**

$$V^d(a_j) = \frac{1}{\mu} \ln \sum_{a_{j+1} \in \mathcal{A}(a_j)} e^{\mu\{v(a_{j+1}|a_j) + V^d(a_{j+1})\}}$$

Recursively computed

This (efficient) computational method of modeling is called:

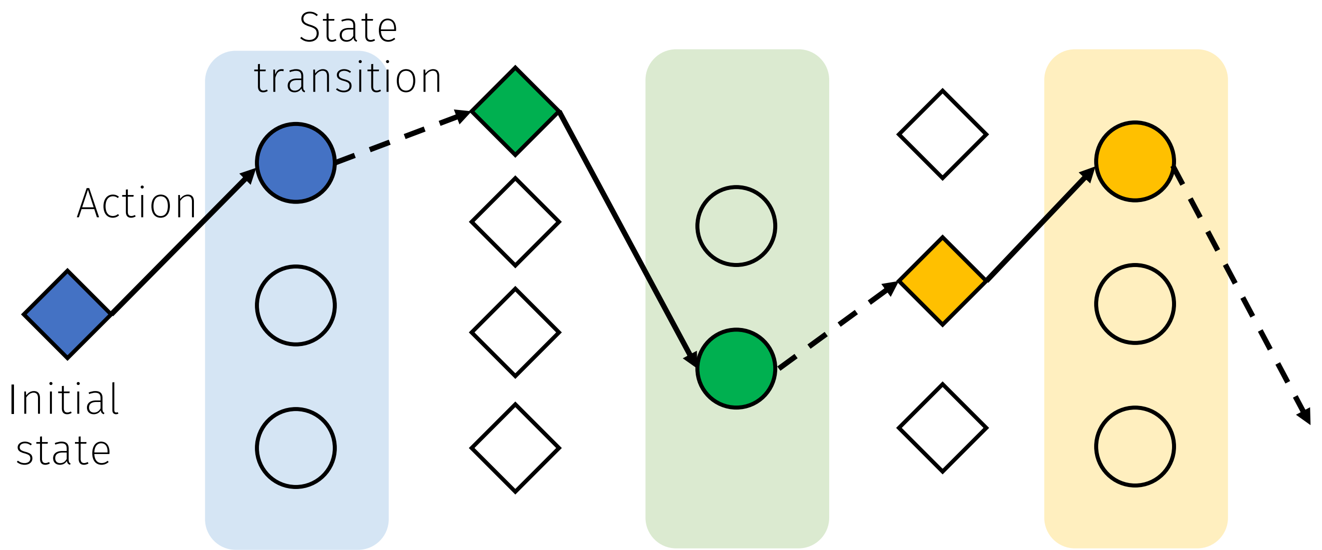
"Recursive Logit (RL) model"

Named by Fosgerau et al. (2013)

Markov Decision Process (MDP)

To more generalize, define

- **Action**: choice behavior (what agent does)
- **State**: situation (where agent is) that changes as result of action



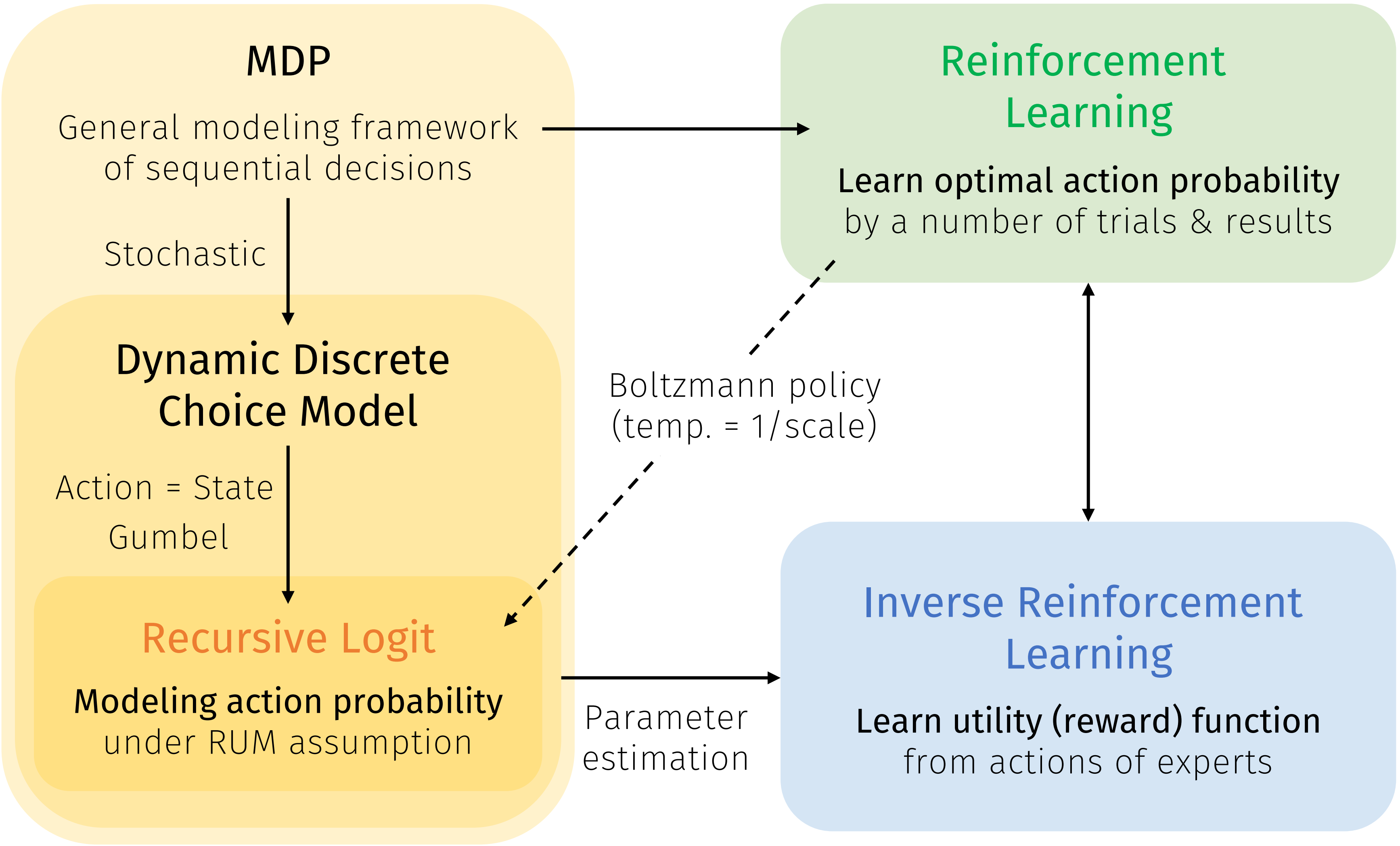
$$V(s) = \max_a \left\{ \sum_{s'} \underline{P(s'|s,a)} \{ v(s,a,s') + \gamma V(s') \} \right\}$$

Discount factor

State transition probability

*In path choice (recursive) modeling: **Action** is directly choice of **State**

Reinforcement Learning approaches



See also: Mai and Jaillet (2020)

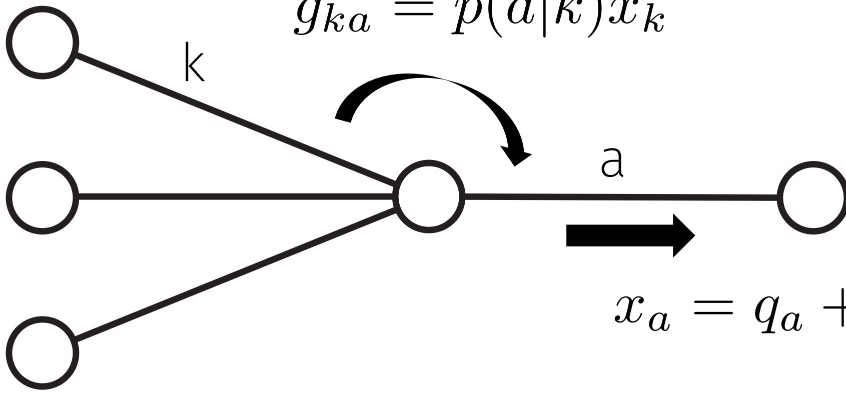
Now, we have **link transition probabilities** $\{p(a|k)\}_{k,a\in\mathcal{L}}$

Given OD demand $\mathbf{q}$, we compute network **traffic flows**

$\{x_a\}$: link flow (on a)

$\{g_{ka}\}$: **transition flow** (from link k to link a)

$$g_{ka} = p(a|k)x_k \quad [1]$$



$$x_a = q_a + \sum_{k\in\mathcal{B}(a)} g_{ka} \quad [2]$$

[1] and [2] reduces to:

$$x_a = q_a + \sum_{k\in\mathcal{B}(a)} p(a|k)x_k \quad \Leftrightarrow \quad \mathbf{x} = \mathbf{P}^\top\mathbf{x} + \mathbf{q}$$

Can be efficiently computed!

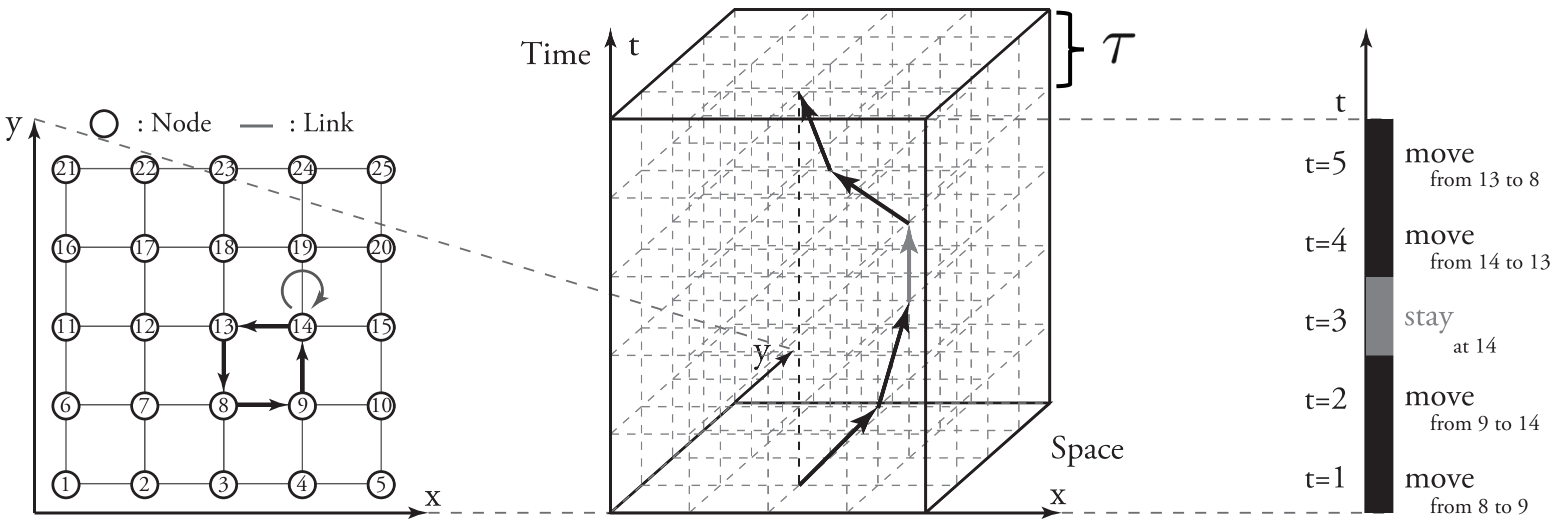
Another dimension may be needed

e.g., a planner may expect changes of visitors' **time-use** in a city center

Time-structured network

allows for **integrated modeling of route, activity place and duration**.

A path $r = [s_1, s_2, \ldots, s_J]$ represents **multiple activities**



Network traffic:

$$x_a = \sum_t x_{ta}$$: no. people who visited space a

$$T_a = x_a \tau$$: **total time spent** at space a

Calculate **indicators** based on traffic

Examples:

$$\sum_a (x_a \times \mathrm{Time}_a)$$: **total travel time** experienced [min.]

$$\sum_a (x_a \times \mathrm{Price}_a)$$: **total revenue** the manager gains [JPY]

$$C \sum_a (x_a \times \mathrm{Length}_a)$$: **total CO_2 emission** [g CO_2]

$$\sum_{od} q_{od} V^d(o)$$: **consumer surplus** (welfare)

Remark (again):

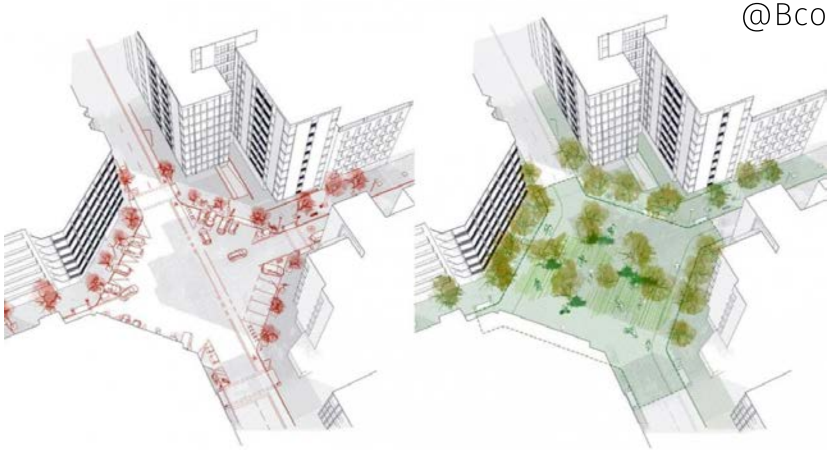
The choice of objective reflects
what you (the planner) want to achieve through the project

Minimizing negative indicators is enough? What is a better/ideal city you think?

A public project entails **trade-offs of goals**

A road closure may **increase travel time** of the network.

Barcelona superblock
@Bcomu Global



But the space can be utilized as a park that is **good for activities, health and environment**.

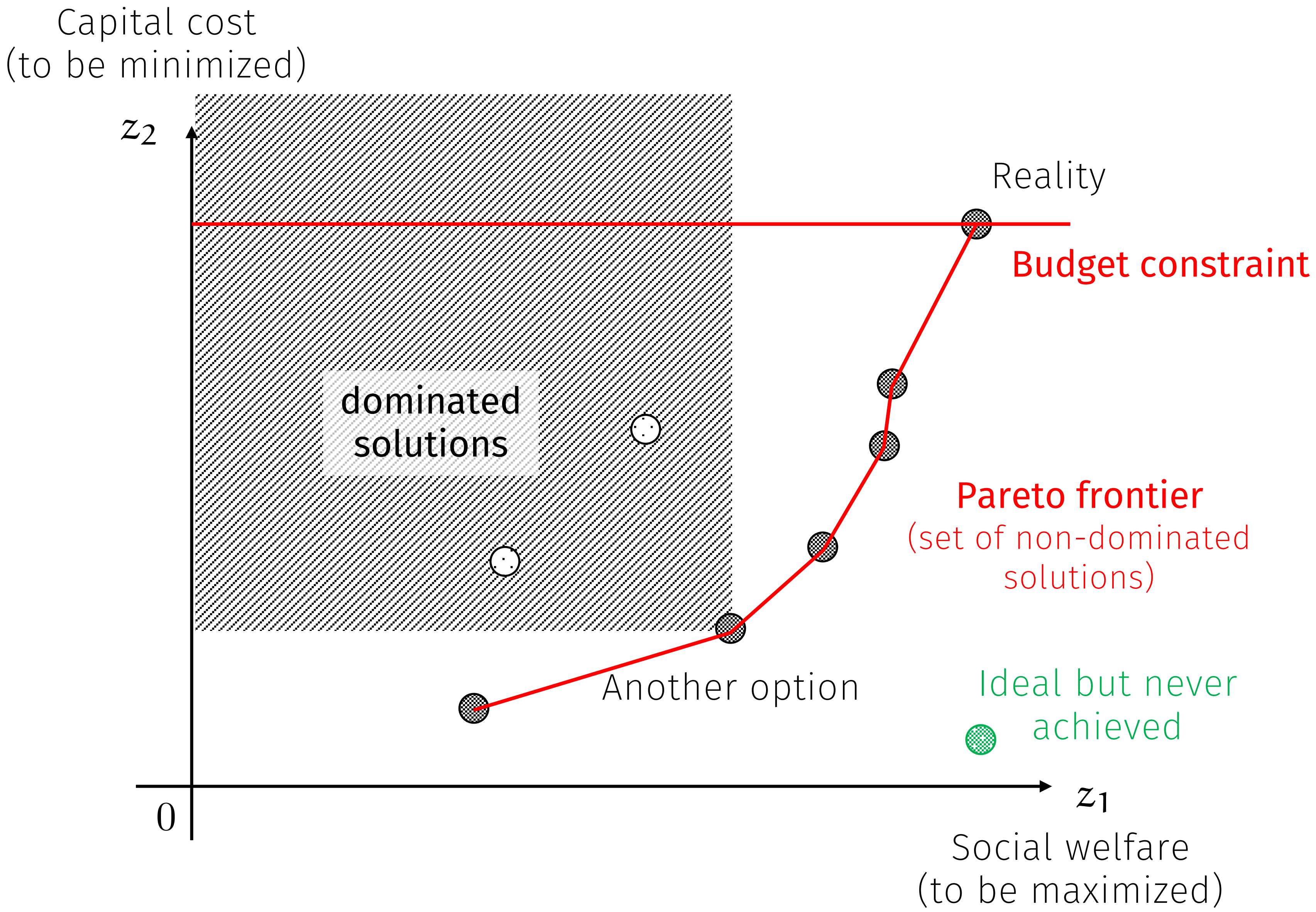
Of course, it requires **a large capital cost**, and the budget is limited.

Weighted sum is enough ?

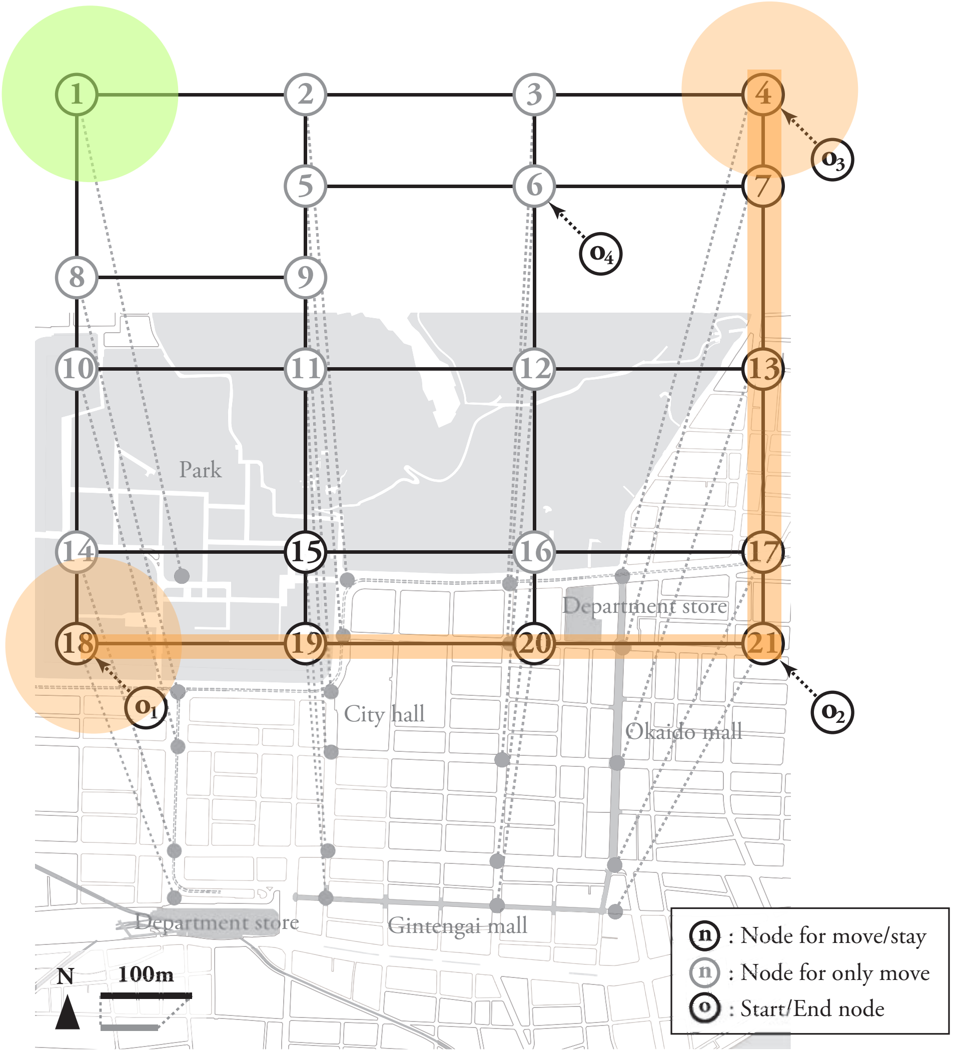
$$Z = \alpha_1 Z_1 + \alpha_2 Z_2 + \alpha_3 Z_3 + \cdots$$

- Often, there is a **clear trade-off** between two objectives
- Weight selection may lead to a **biased policy decision**

Multi-objective design

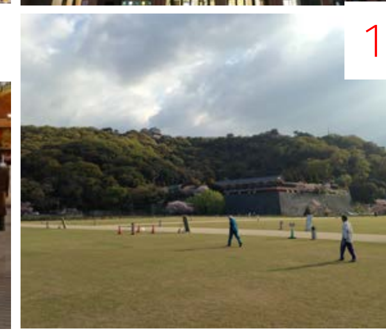


Case study | A pedestrian activity network design



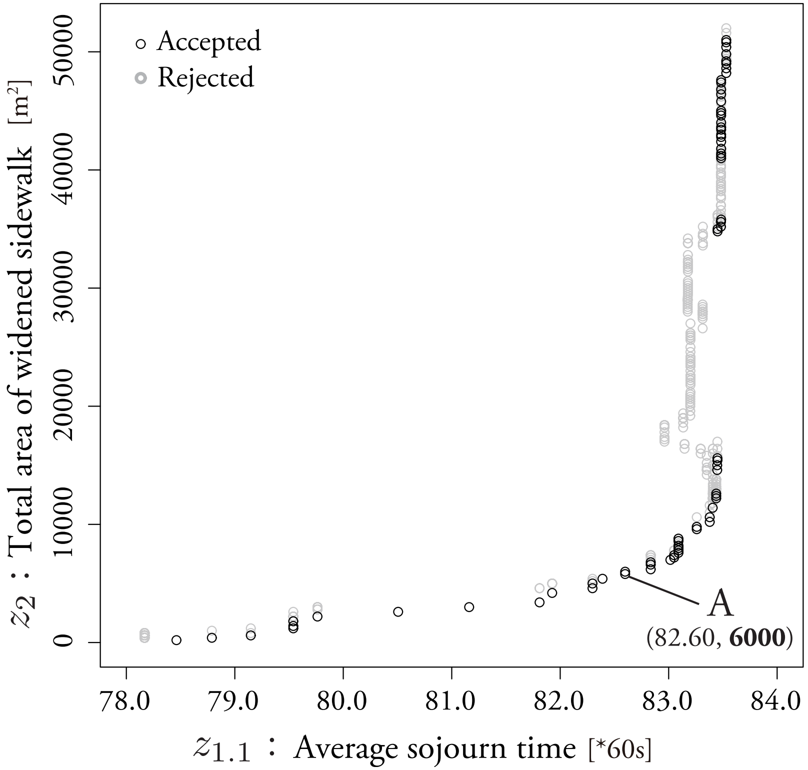
City center of Matsuyama city

- **Design**: expansion of walking space on each street [m.]
- **Expectation**: resistance decreases, and more places are visited

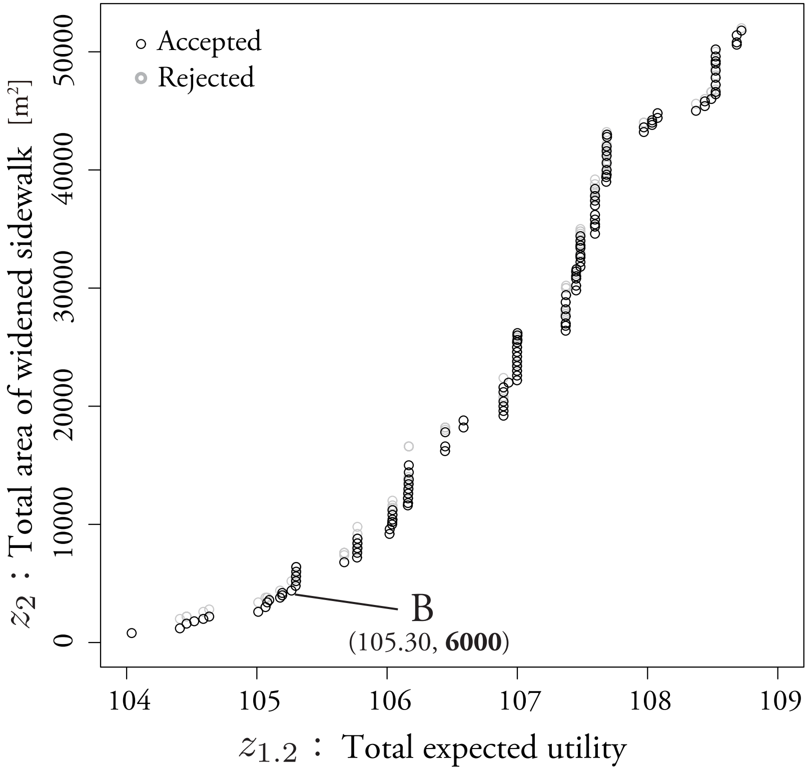


Case study | A pedestrian activity network design

Goal I:
Sojourn time maximization

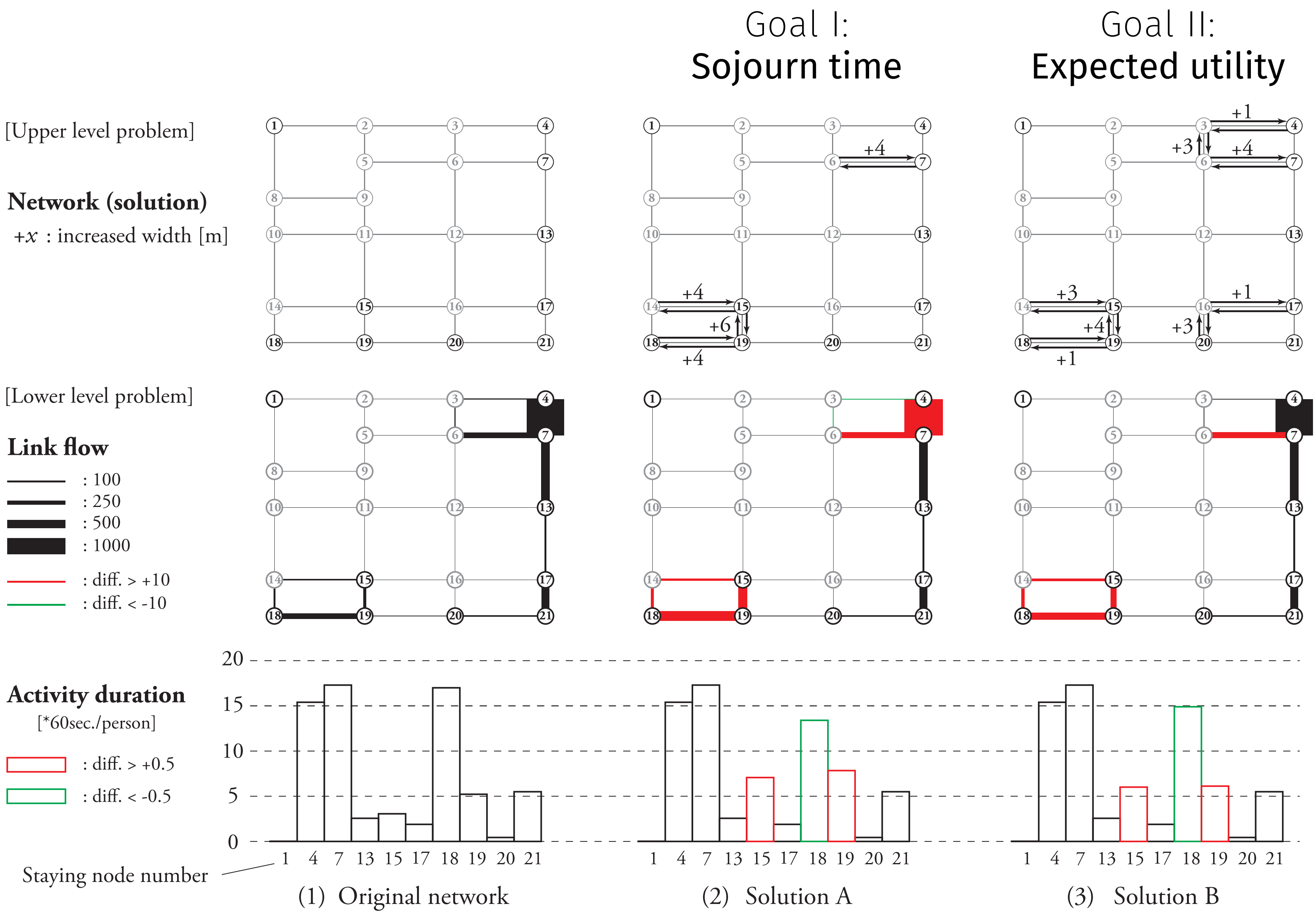


Goal II:
Expected utility maximization



- **Clear trade-offs between goals and budget** are observed.
- Pareto frontier **offers a variety of policies** based on the investment level

Case study | A pedestrian activity network design



Summary & Remarks

- Reinforcement Learning is a general framework of modeling **sequential decisions in networks**.
  - You can model any "state-action network"
  - "State = action = space" is just an example
- Network design is a mathematical problem of **behavior (in a network) based planning**
  - Be thoughtful when you set an objective
  - Multi-objective design may fit in public projects

Questions ?

oyama@shibaura-it.ac.jp

References

- Magnanti, T.L., Wong, R.T., 1984. Network design and transportation planning: models and algorithms. Transportation Science 18 (1), 1–55.
- Farahani, R. Z., Miandoabchi, E., Szeto, W. Y., & Rashidi, H. 2013. A review of urban transportation network design problems. European Journal of Operational Research, 229(2), 281-302.
- Fosgerau, M., Frejinger, E., Karlstrom, A., 2013. A link based network route choice model with unrestricted choice set. Transportation Research Part B: Methodological 56, 70–80.
- Mai, T., & Jaillet, P., 2020. A Relation Analysis of Markov Decision Process Frameworks. arXiv:2008.07820.
- Oyama, Y., 2017. A Markovian route choice analysis for trajectory-based urban planning. PhD thesis, The University of Tokyo.
- Oyama, Y., Hato, E., 2019. Prism-based path set restriction for solving Markovian traffic assignment problem. Transportation Research Part B: Methodological 122, 528–546.
- 大山雄己, 羽藤英二, 2017. 多目的最適化に基づく歩行者の活動ネットワークデザイン. 都市計画論文集 52(3): 810-817.

Appendix | Design levels and examples

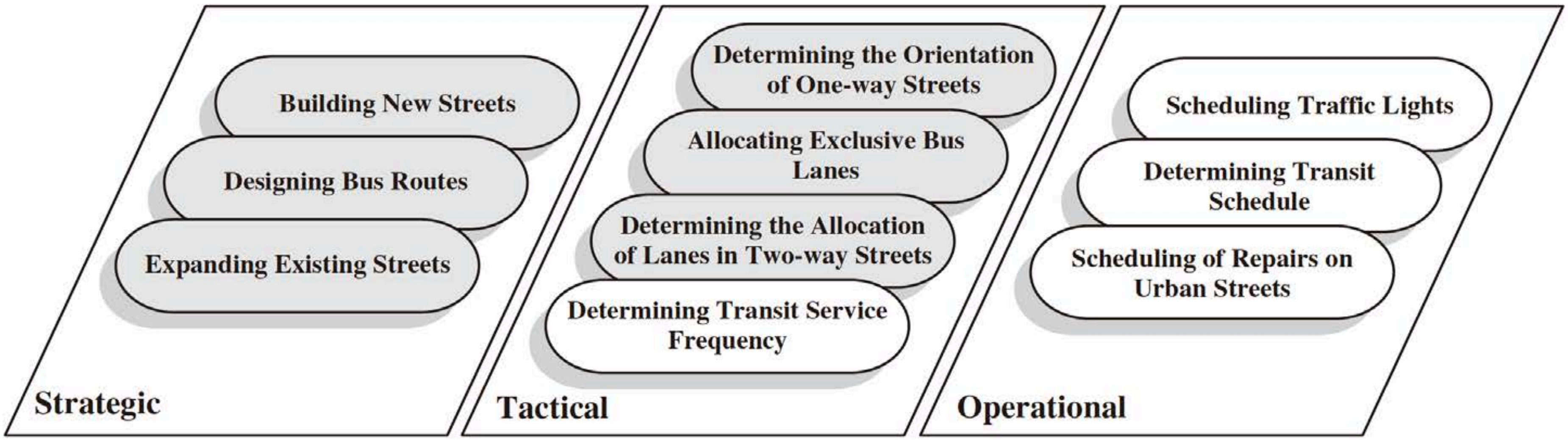


Figure 1 in Farahani et al. (2013)

Appendix | Solution algorithms (metaheuristics)

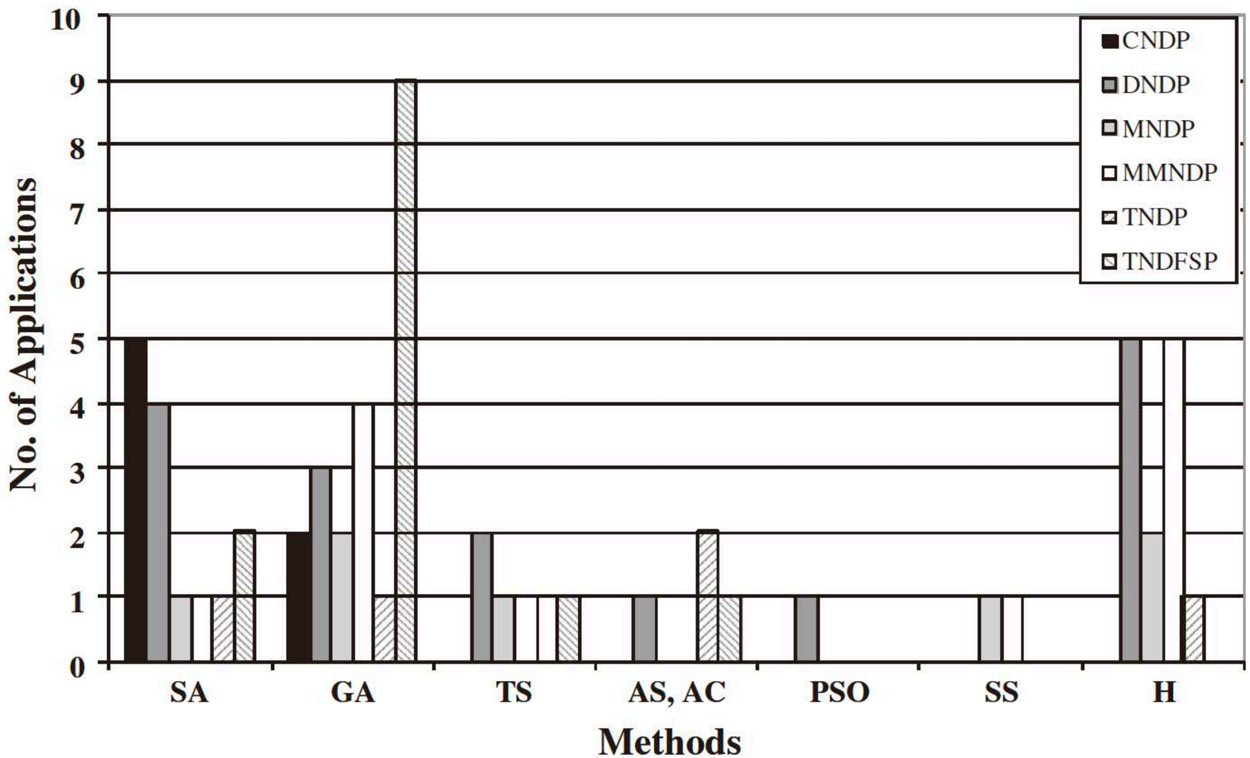


Figure2 in Farahani et al. (2013)

SA: Simulated Annealing; GA: Genetic Algorithm; TS: Tabu Search; AC: Ant Colony; PSO: Particle Swarm Optimization; SS: Scatter Search; H: Hybrid metaheuristics