Combinatorial Optimization at Work 2020

Traffic Optimization

Part I: Paths & Lagrange Relaxation

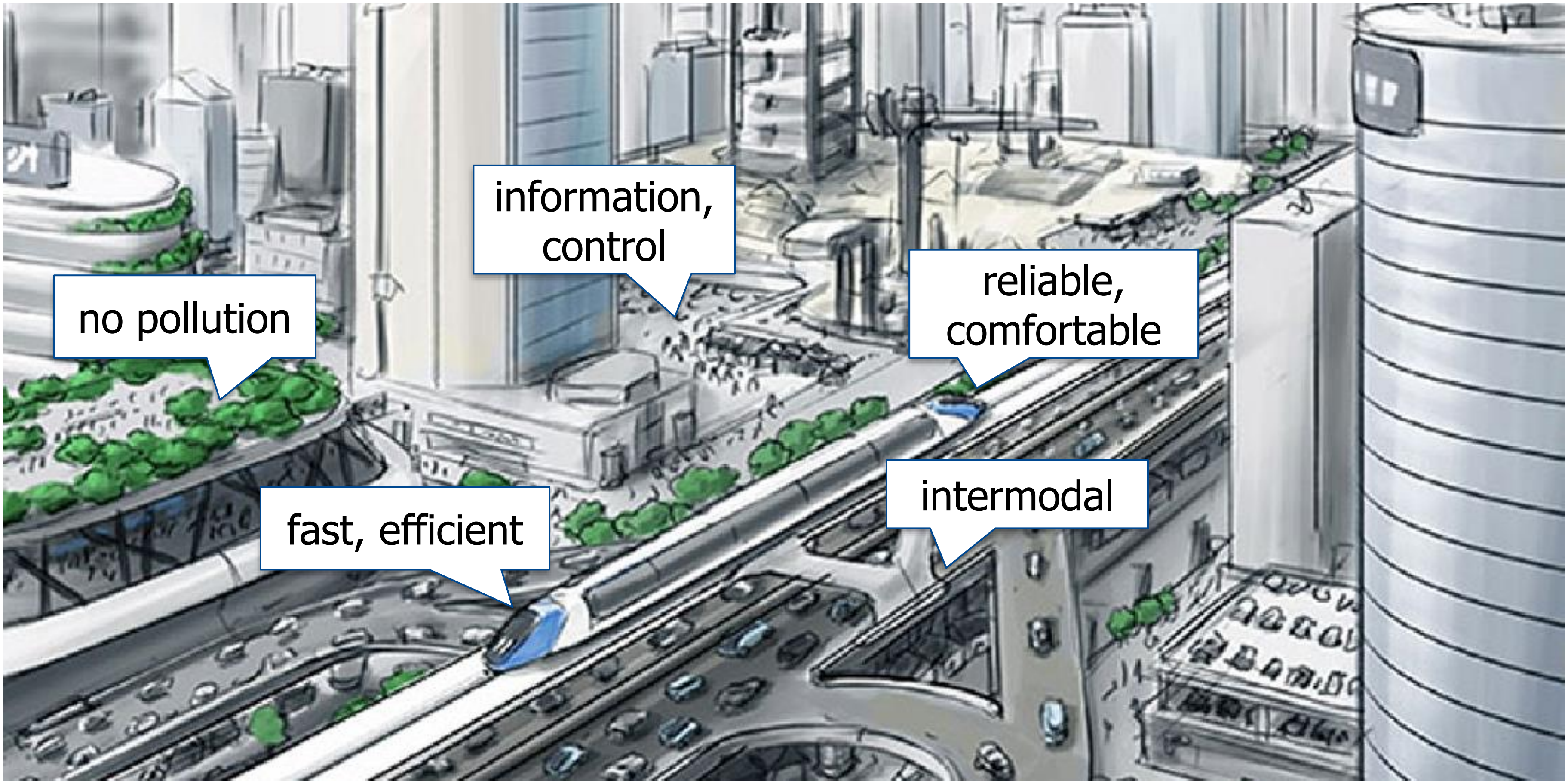
Part II: Vehicles & Crews

Part III: Pollsters & Vehicles

Zuse Institute Berlin, 22.09.2020

Ralf Borndörfer Güvenç Şahin Luis Torres

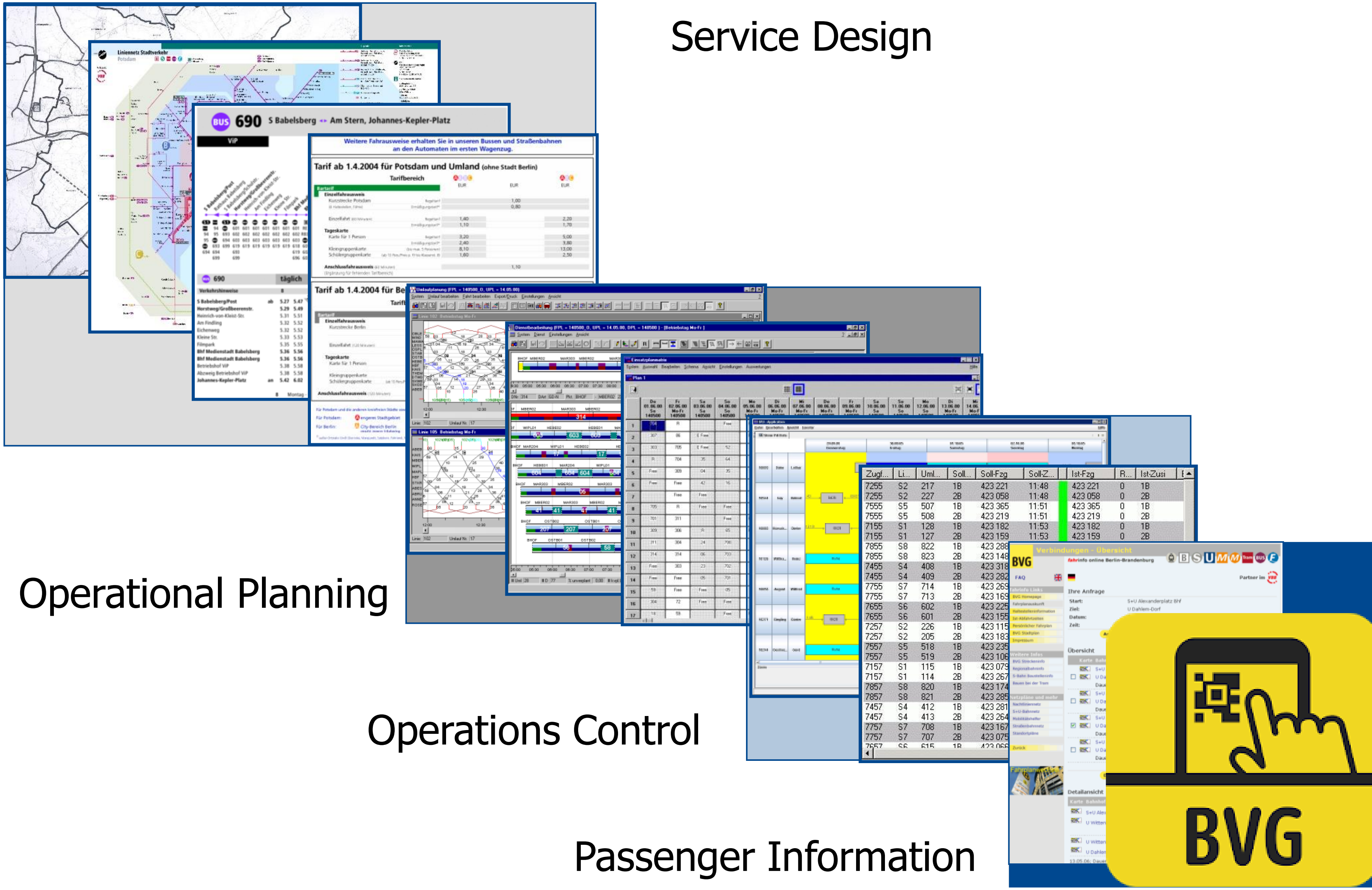
Traffic of the Future



- Needs data and communication to assess and control system status
- Needs mathematics to find smart solutions

Planning Problems in Public Transit

Service Design



Operational Planning

Operations Control

Passenger Information

Vehicle Rotation Planning is a Crucial Part of the Production Planning Process

Slide of DB

Customer

- Market research
- Demand forecast
- Passenger counting

Marketing
- Which connections correspond best to demand?
- What capacity is needed?
- What kind of product should be offered?

Train Path Planning
- How can customer-friendly time tables be realised
- How can conflicts with other trains be avoided?

Vehicle Rotation Planning
- **How can vehicles be employed efficiently?**
- **How can plans be made maintenance friendly and robust?**

Maintenance Planning
- How can maintenance capacity be used most efficiently
- How can the availability of vehicles be maximised?

Train Preparation Planning
- How can vehicles be stabled, shunted, catered, cleaned etc. between maintenance and commercial service?

Staff Planning
- How can shifts be planned staff friendly?
- How can high productivity of personnel be obtained?

- Attractive timetable
- Reliability
- High product quality

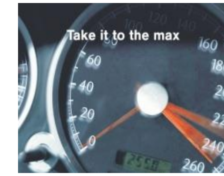
Customer

IVU.suite for Buses and Rail Transport

IVU TRAFFIC TECHNOLOGIES AG

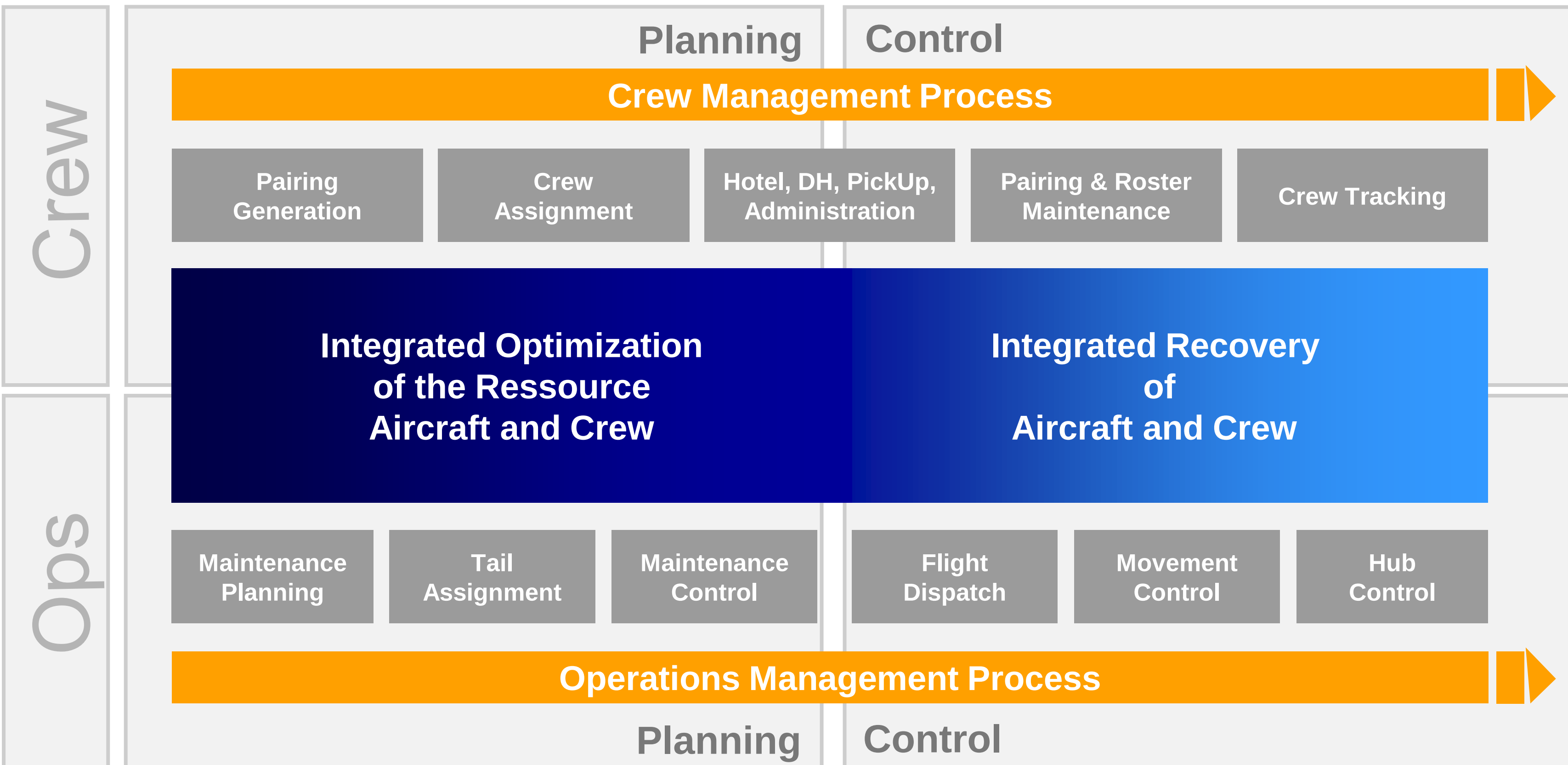
PLANNING	DISPATCH	FLEET MANAGEMENT	TICKETING	PASSENGER INFORMATION	ACCOUNTING
IVU.plan Timetable planning, vehicle working and duty scheduling	**IVU.**vehicle Vehicle dispatch	**IVU.**fleet Traffic control centre	**IVU.**fare Background system	**IVU.**realtime Dynamic passenger information	**IVU.**control Performance assessment and statistics
IVU.pool Data integration	**IVU.**crew Personnel dispatch	**IVU.**cockpit On-board computer software	**IVU.**ticket Ticket sales and e-ticketing	**IVU.**journey Consistent journey planning	
		IVU.box On-board computer hardware	**IVU.**validator E-ticketing terminal		

Slide of IVU



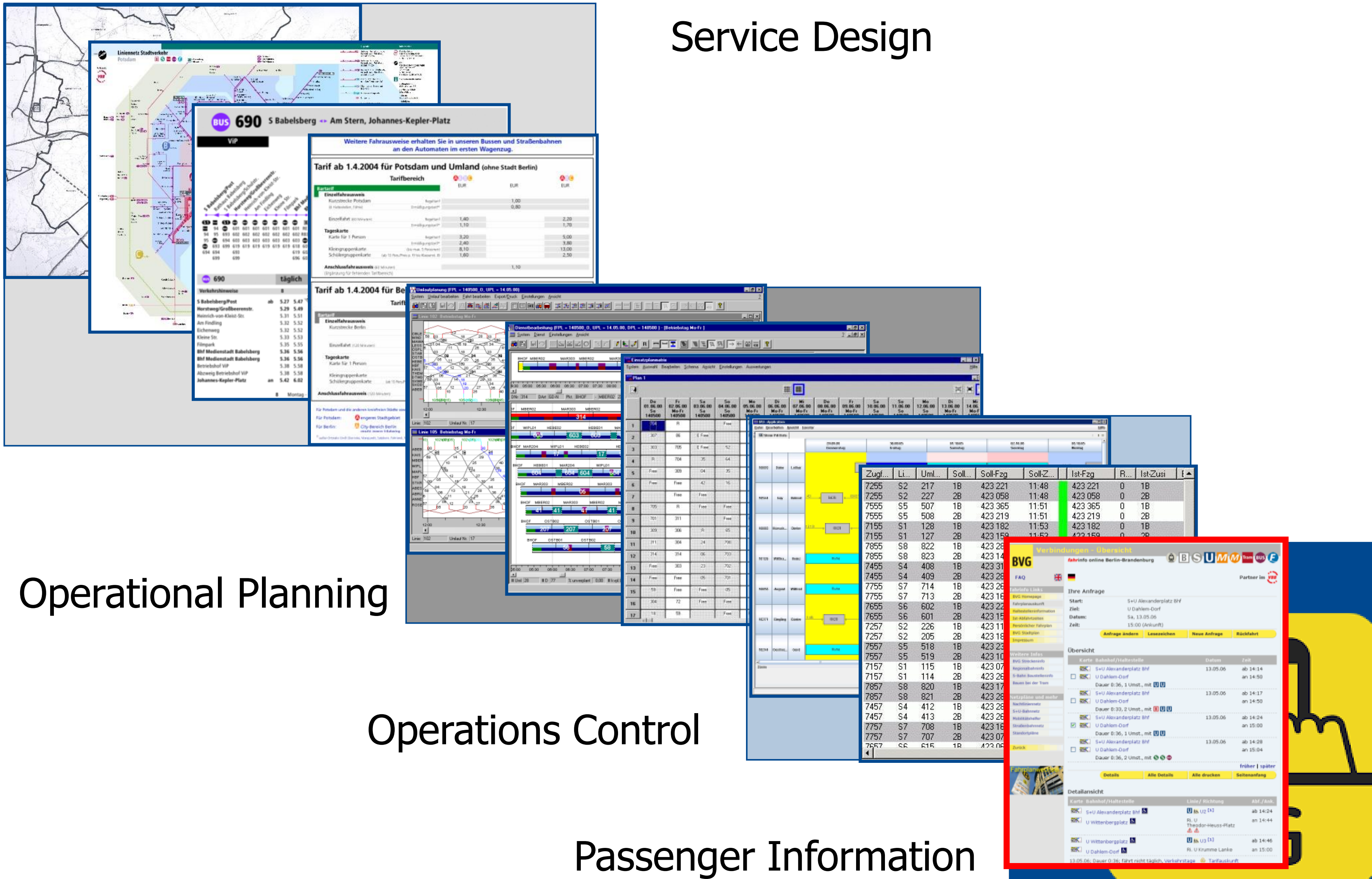
Workflow Oriented and Integrated Optimization: How fast business processes can follow IT?

Slide of LSB



Planning Problems in Public Transit

Sabancı Universitesi



Service Design

Operational Planning

Operations Control

Passenger Information

The Shortest Path Problem

Shortest Path Problem

Sabancı Universitesi

1 Def. (Single Source Shortest (st-)Path Problem (ShPP)): Let $G = (V, E, c)$ be a (un)directed graph on n nodes with edge weights $c \in \mathbb{R}^E_{\geq 0}$, $s, t \in V$ two nodes, $P^D_{st} = P_{st}$ the set of all st-paths in D.

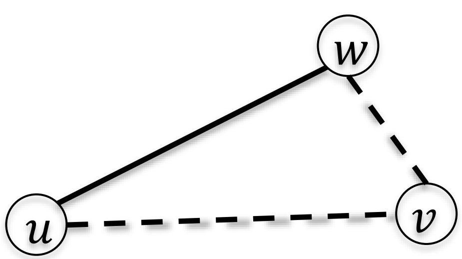
$\min c(P)\,, P \in P_{st}$ **shortet (st)-path problem (ShPP)**

a) ShPP **conservative** $:\Leftrightarrow c(C) \geq 0\ \forall$ (di)cycles $C \subseteq E$

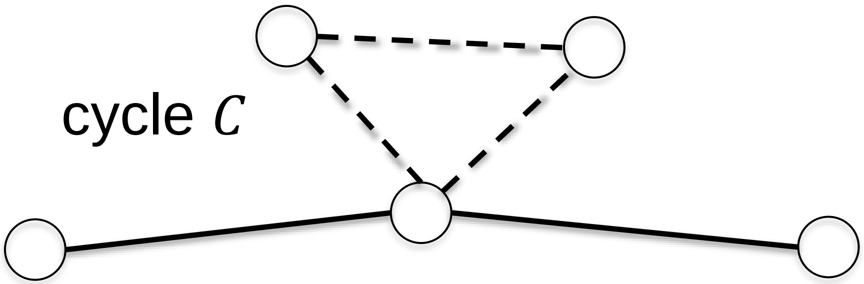
b) ShPP **metric** $:\Leftrightarrow c_{uv} + c_{vw} \geq c_{uw}\ \ \forall uv, vw, uw \in E$ (Δ-inequality)

c) ShPP **Euclidean** $:\Leftrightarrow c_{uv} = \left|\left|u - v\right|\right|_2\ \ \forall uv \in E \subseteq \mathbb{R}^2$

2 Obs. (Simplicity): A conservative ShPP has a simple optimal solution (no node repetitions).



Δ-inequality



simple/non-simple path

Dijkstra's Algorithm

Sabancı Üniversitesi

3 Alg. (Dijkstra's Algorithm, Dijkstra [1959]):

Input: $G = (V, E, c), s, t \in V, c \in \mathbb{R}^E_{\geq 0}$ **Output:** $P \in \operatorname{Argmin}_{P \in P_{st}} c(P)$

Data Structures: $d \in (\mathbb{R} \cup \{+\infty\})^V,\ \text{pred} \in (V \cup \{\text{nil}\})^V, R \subseteq V$ reached nodes

1. $d[v] \leftarrow +\infty,\ \text{pred}[v] \leftarrow \text{nil}\ \forall v \in V,\ d_s \leftarrow 0, R \leftarrow \{s\}$
2. while $R \neq \emptyset$ do
3. $\quad u \leftarrow \text{argmin}_{v \in R}\, d[v]$, $R \leftarrow R \setminus \{u\}$
4. $\quad$ forall $uv \in E$ do
5. $\quad\quad$ if $d[u] + c_{uv} < d[v]$ then
6. $\quad\quad\quad d[v] \leftarrow d[u] + c_{uv}, R \leftarrow R \cup \{v\}, \text{pred}[v] \leftarrow u$
7. $\quad\quad$ endif
8. endwhile
9. output $(t, \text{pred}[t], \text{pred}^2[t], \dots, s \text{ or nil})$

$O(n \times \log n)$ (lines 2–3)

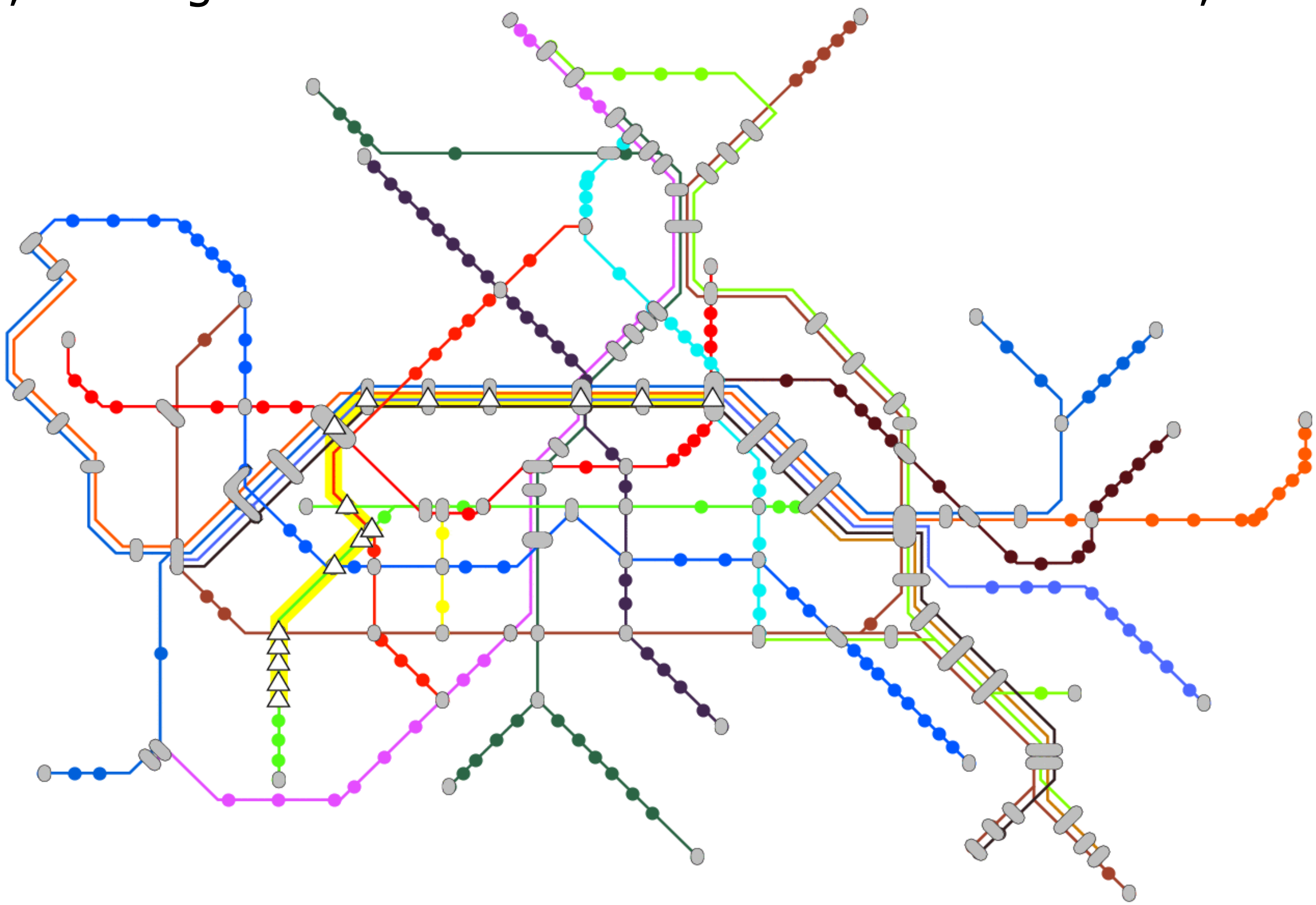
$O(m)$ (amortized) (lines 4–6)

4 Prop. (Correctness and Run Time of Dijkstra's Algorithm):
Alg. 3 is correct and runs in $O(n \log n + m)$.

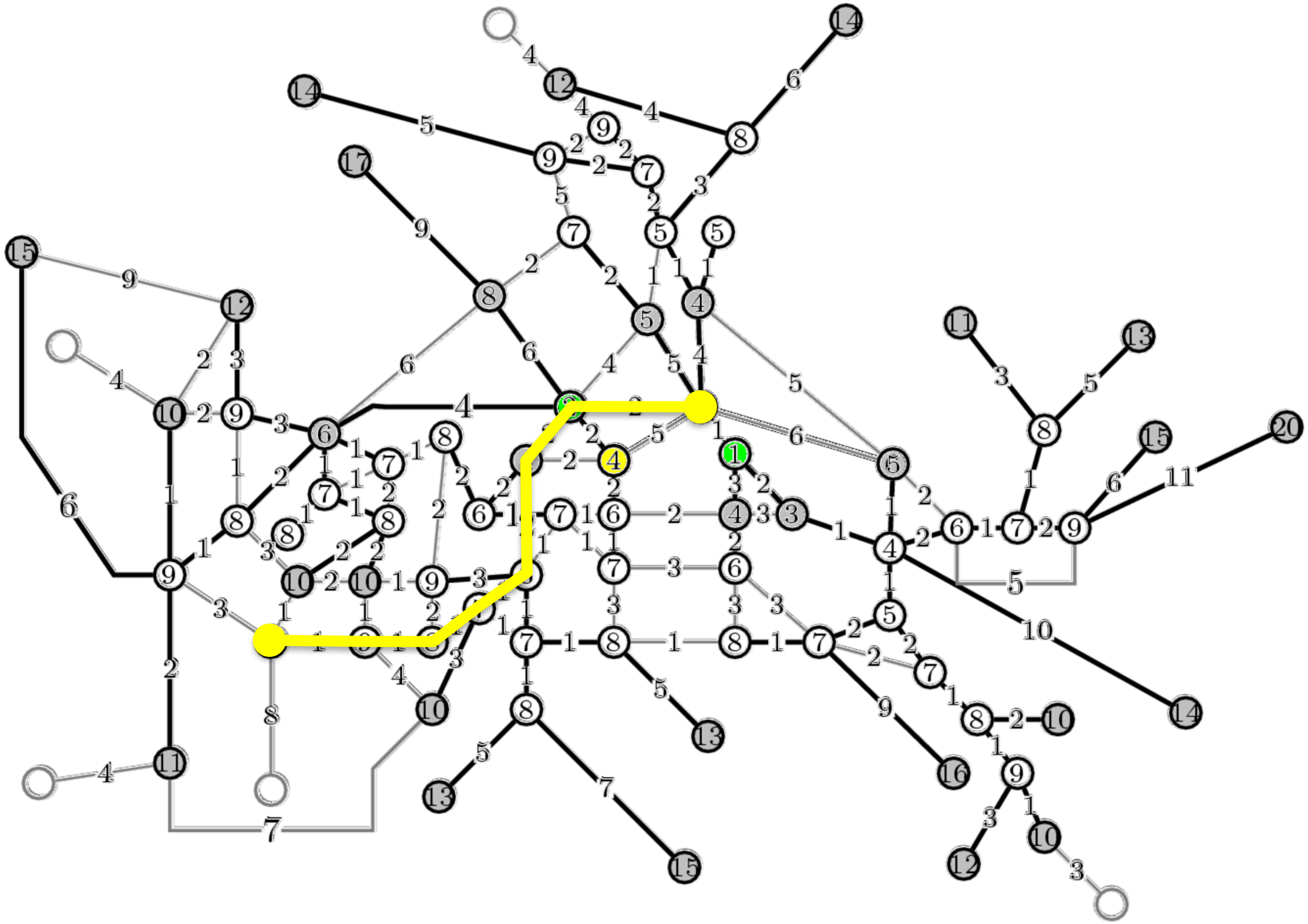
Shortest Path Problem

306 nodes, 445 edges

80 nodes, 122 edges



Dijkstra's Algorithm



IP Formulation of the ShPP

5 Def. (IP Formulation of the ShPP): Let $D = (V, A, c)$ be a **directed** graph, with arc weights $c \in \mathbb{R}^A_{\geq 0}$, $s, t \in V$ two nodes.

(ShPP)	$\min c^T x$		objective
(i)	$x\big(\delta^+(v)\big) - x\big(\delta^-(v)\big) = 0$	$\forall v \neq s, t$	flow conservation
(ii)	$x\big(\delta^+(s)\big) = 1$		flow constraint
(iii)	$0 \leq x \leq 1$		bounds
(iv)	x integer		integrality

a) $P^{ShPP} := \text{conv}\,\{\chi_P : P \in P_{st}\}$ **<u>st-path polytope</u>**

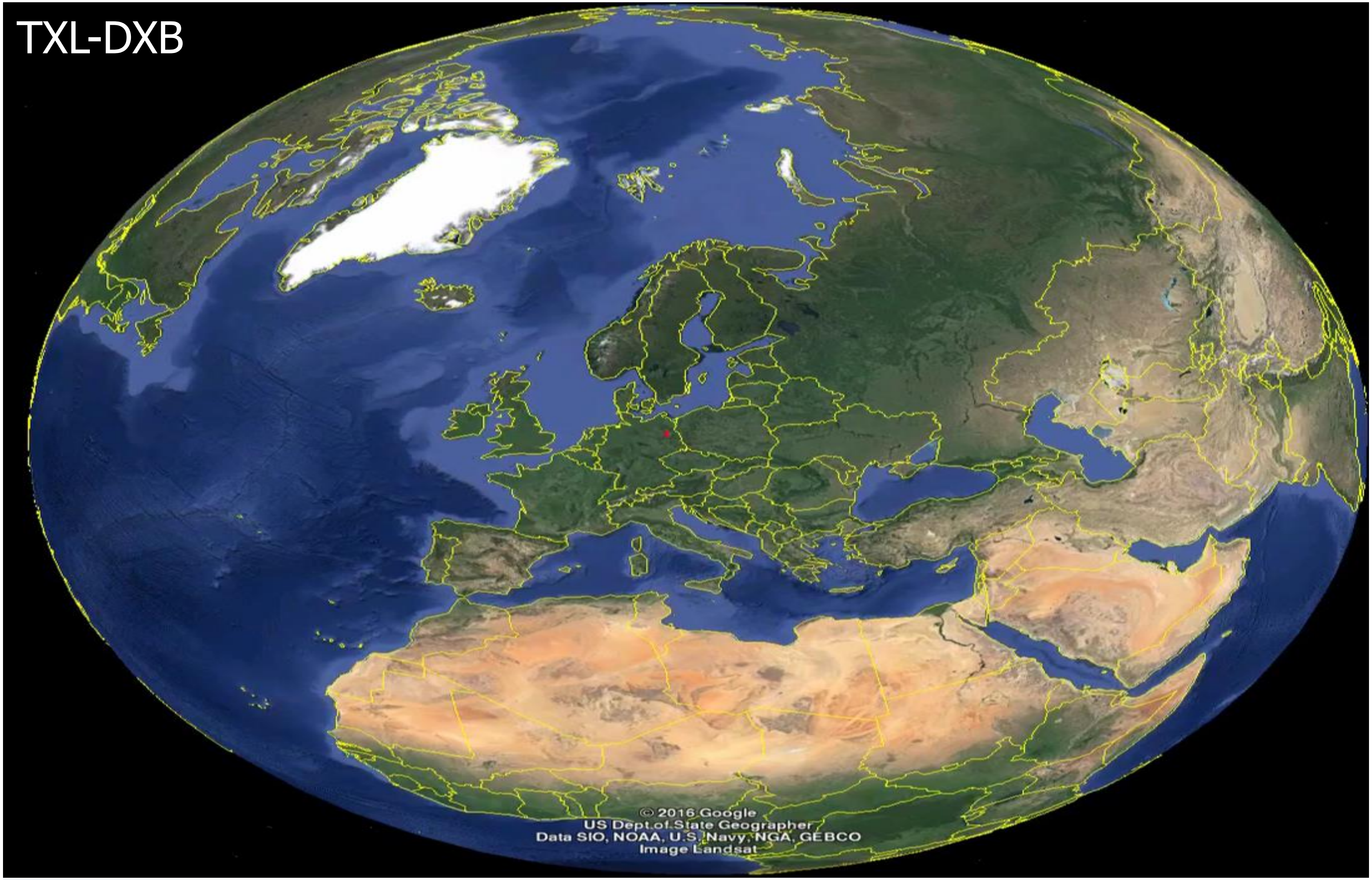
b) $P_I^{ShPP} := \text{conv}\,\{x \in \mathbb{R}^E : (\text{ShPP})\ (\text{i}) - (\text{iv})\}$ **<u>ShPP polytope</u>**

c) $P_{LP}^{ShPP} := \text{conv}\,\{x \in \mathbb{R}^E : (\text{ShPP})(\text{i}) - (\text{iii})\}$ **<u>ShPP LP-relaxation</u>**

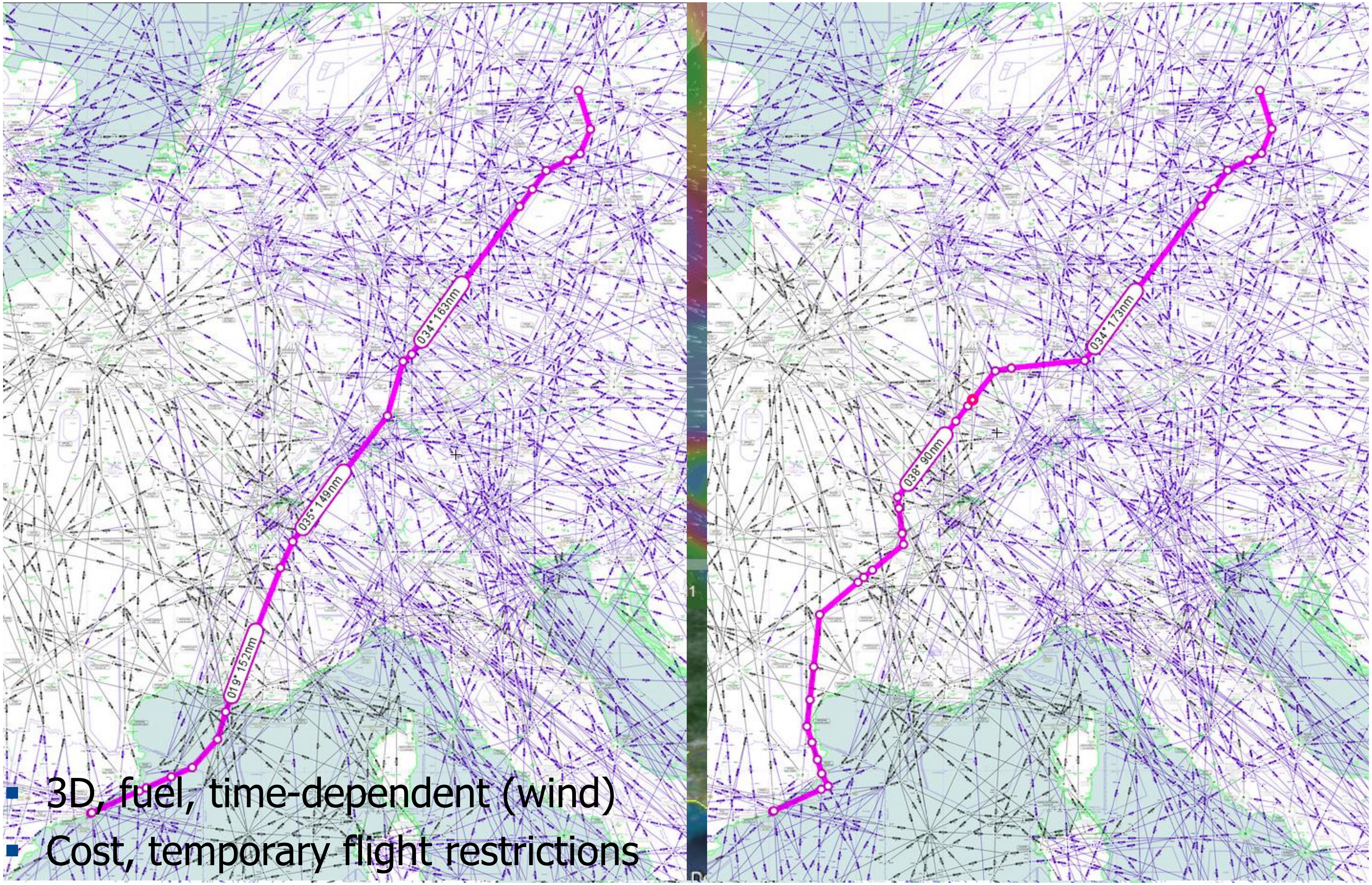
6 Prop. (Path Polytopes): $P^{ShPP} = P_I^{ShPP}$ is in general not true, but $P_I^{ShPP} = P_{LP}^{ShPP}$; $\underset{x \in P_{LP}^{ShPP}}{\text{Argmin}}\, c^T x$ contains a path for conservative c. **Proof:** P_I^{ShPP} allows subtours, P_{LP}^{ShPP} describes a flow. □

Dijkstra's Algorithm – A* – Superoptimal Wind

Sabancı Üniversitesi



Flight Planning



- 3D, fuel, time-dependent (wind)
- Cost, temporary flight restrictions

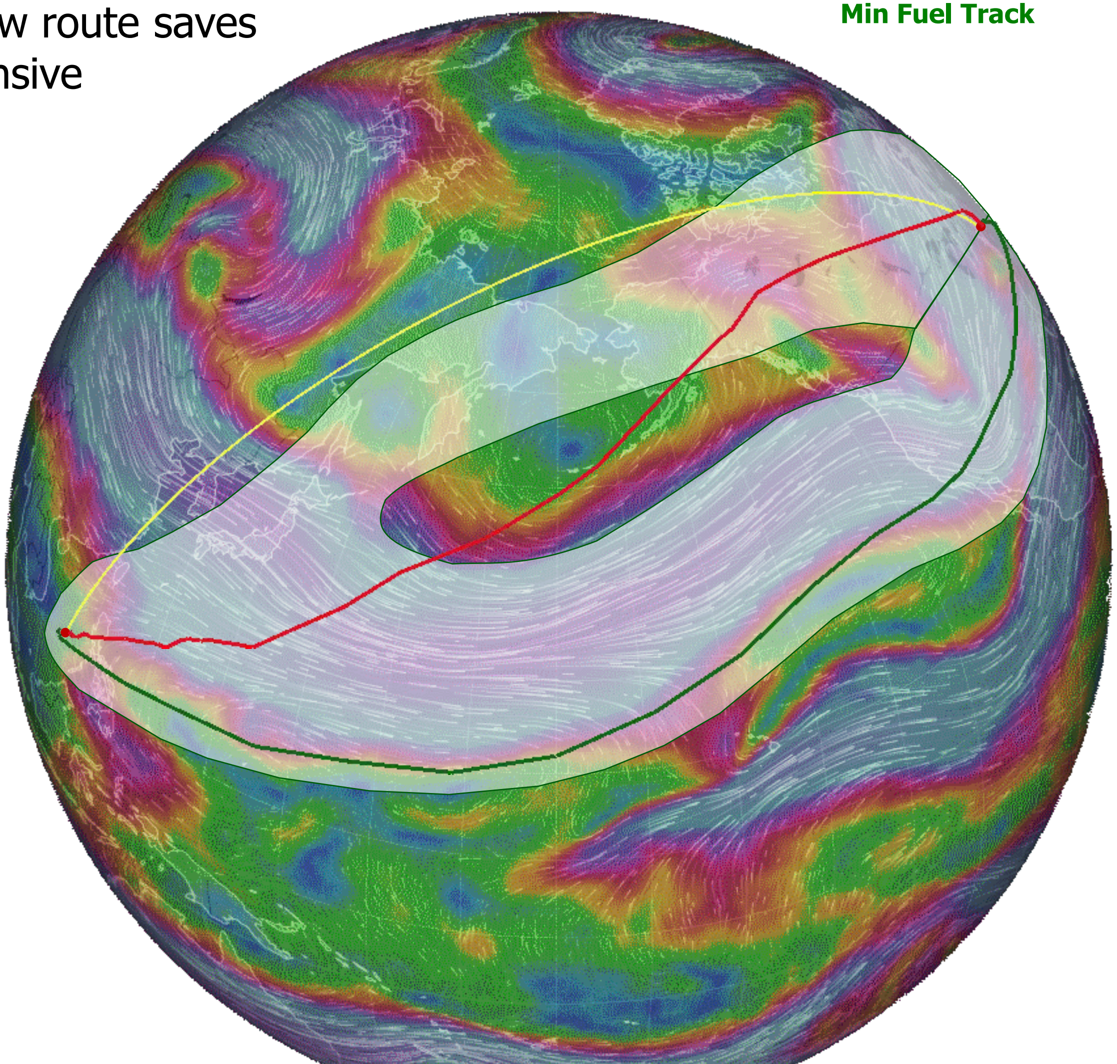
Taipei – New York

Boeing B777-300ER, 25 April 2017, great circle distance 12.565 km

- The green trajectory takes better advantage of the strong jet stream (~300 km/h).
- It is worth to take a long detour.
- Besides saving fuel and time, the new route saves overflight fees by avoiding the expensive airspaces of Canada and Japan.

	using old heuristic search space reduction	using new dynamic search space reduction	GAIN
distance flown (km)	13.385	14.635	-1250
flight time (hours)	14:40	13:55	0:45
fuel burn (kg)	95.524	89.859	5665 = 17,8 t CO_2
overflight fees (USD)	2291	1139	1152
total cost (USD)*	76.453	71.118	5335

* based on: fuel price 500 USD / ton, flight time costs: 1400 USD / hour



Constrained Shortest Path Problem

5 Def. (IP Formulation of the Constrained ShPP (CSP)): Let $D=(V,A,c)$ be a **directed** graph, with arc weights $c\in\mathbb{R}^A_{\geq 0}$, $s,t\in V$ two nodes, $Ax\leq b$ some linear constraints.

(CSP)	$\min c^T x$		objective
(i)	$x\big(\delta^+(v)\big)-x\big(\delta^-(v)\big)=0$	$\forall v\neq s,t$	flow conservation
(ii)	$x\big(\delta^+(s)\big)=1$		flow constraint
(iii)	$0\leq x\leq 1$		bounds
(iv)	$Ax\leq b$		path constraints
(v)	x integer		integrality

a) $P^{CSP} := \operatorname{conv}\{\chi_P : P\in P_{st}, A\chi_P\leq b\}$ **<u>const. st-path polytope</u>**

b) $P_I^{CSP} := \operatorname{conv}\{x\in\mathbb{R}^E : \text{(ShPP) (i)} - \text{(v)}\}$ **<u>CSP polytope</u>**

c) $P_{LP}^{CSP} := \operatorname{conv}\{x\in\mathbb{R}^E : \text{(ShPP)(i)} - (\mathrm{i}v)\}$ **<u>CSP LP-relaxation</u>**

6 Obs. (CSP): $P^{CShP}\subseteq P_I^{CSP}\subseteq P_{LP}^{CSP}$; equality does in general not hold. The CSP is NP-hard.

IP Formulation of the Constrained ShPP

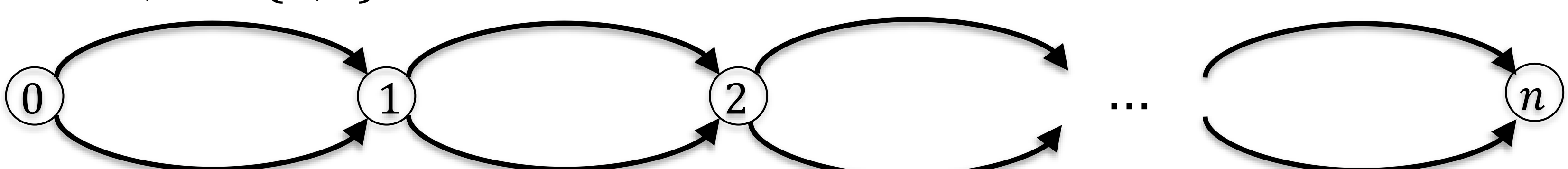
Sabancı Üniversitesi

5 Def. (IP Formulation of the Constrained ShPP (CSP)): Let $D = (V, A, c)$ be a **directed** graph, with arc weights $c \in \mathbb{R}^A_{\geq 0}$, $s, t \in V$ two nodes, $Ax \leq b$ some linear constraints.

(CSP)	$\min c^T x$		objective
(i)	$x(\delta^+(v)) - x(\delta^-(v)) = 0$	$\forall v \neq s, t$	flow conservation
(ii)	$x(\delta^+(s)) = 1$		flow constraint
(iii)	$0 \leq x \leq 1$		bounds
(iv)	$Ax \leq b$		path constraints
(v)	x integer		integrality

6 Obs. (CSP): $P^{CShP} \subseteq P_I^{CSP} \subseteq P_{LP}^{CSP}$; equality does in general not hold. The CSP is NP-hard. **Proof:** Solves knapsack problem $\min c^T x$, $a^T x \leq b, x \in \{0,1\}^n$:

$c_{uv} = 0, a_{uv} = 0$



$c_{uv} = c_i, a_{uv} = a_i$

□

Constrained Shortest Path Problem

5 Def. (IP Formulation of the Constrained ShPP (CSP)): Let $D = (V, A, c)$ be a **directed** graph, with arc weights $c \in \mathbb{R}^A_{\geq 0}$, $s, t \in V$ two nodes, $Ax \leq b$ some linear constraints.

(CSP)	$\min c^T x$		objective
(i)	$x(\delta^+(v)) - x(\delta^-(v)) = 0$	$\forall v \neq s, t$	flow conservation
(ii)	$x(\delta^+(s)) = 1$		flow constraint
(iii)	$0 \leq x \leq 1$		bounds
(iv)	$Ax \leq b$		path constraints
(v)	x integer		integrality

a) $P^{CSP} := \mathrm{conv}\,\{\chi_P : P \in P_{st}, A\chi_P \leq b\}$ **<u>const. st-path polytope</u>**

b) $P_I^{CSP} := \mathrm{conv}\,\{x \in \mathbb{R}^E : \text{(ShPP) (i)} - \text{(v)}\}$ **<u>CSP polytope</u>**

c) $P_{LP}^{CSP} := \mathrm{conv}\,\{x \in \mathbb{R}^E : \text{(ShPP)(i)} - (\mathrm{i}v)\}$ **<u>CSP LP-relaxation</u>**

6 Obs. (CSP): $P^{CShP} \subseteq P_I^{CSP} \subseteq P_{LP}^{CSP}$; equality does in general not hold. The CSP is NP-hard, even for acyclic digraphs and $c \geq 0$.

(Acyclic) Constrained Shortest Path Problem

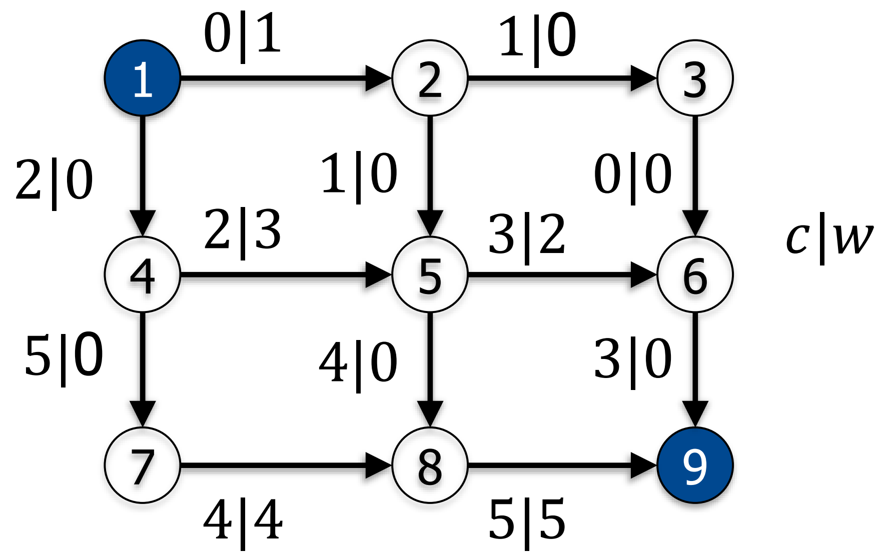
Sabancı Universitesi

8 Def. (Acyclic Constrained Shortest Path Problem (ACSP)): A CSP on a acyclic digraph is acylic.

9 Obs. (Topological Sorting): The nodes of an acyclic digraph $D = (V, A)$ can be topologically sorted s.t. $uv \in A \Longrightarrow u < v$.

7 Ex. (ACSP):

$$\text{(P)} \quad \min_{p \in P_{19}} c^T \chi(p), \ \underbrace{a^T \chi(p) \geq 6}_{-Ax \leq -b}$$

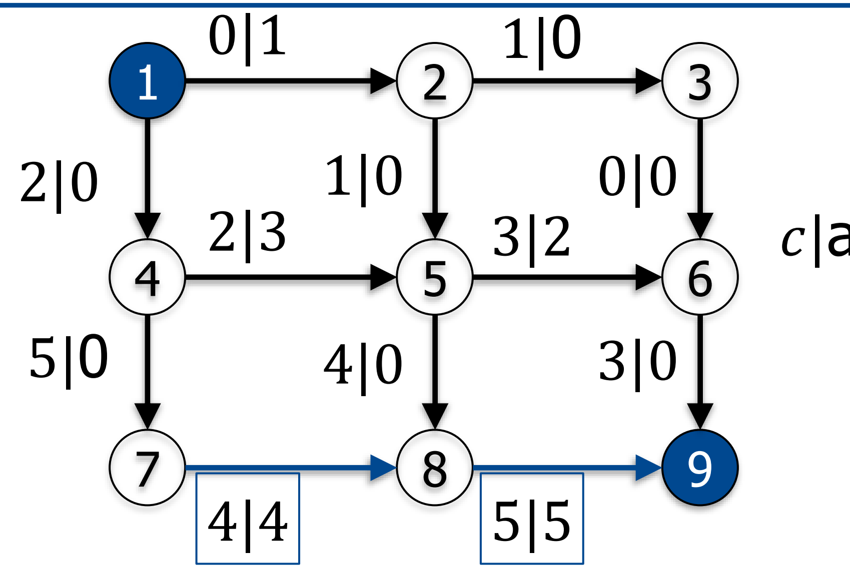


$1 < \cdots < 9$ sorts $V = [9]$ topologically.

Sabancı Üniversitesi

7 Ex. (ACSP):

$$(\mathrm{P})\quad \min_{p\in P_{19}} c^T\chi(p),\ \underbrace{a^T\chi(p)\geq 6}_{-Ax\leq -b}$$



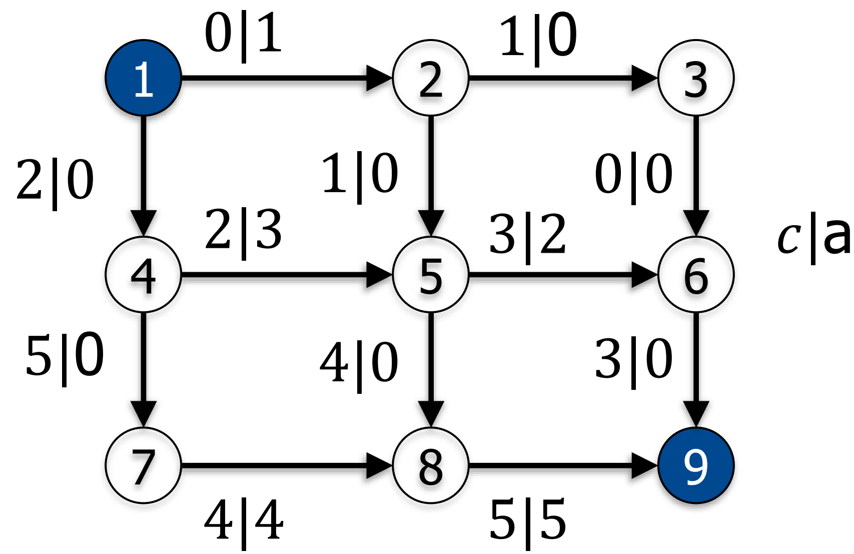
$^v/_a$	0	1	2	3	4	5	6	7	8	9	10
1	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$
2	$-\infty$	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$
3	$-\infty$	1	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$
4	2	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$
5	$-\infty$	1	$-\infty$	4	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$
6	$-\infty$	1	$-\infty$	4	$-\infty$	7	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$
7	7	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$
8	$-\infty$	5	$-\infty$	8	11	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$
9	$-\infty$	4	$-\infty$	7	$-\infty$	10	10	$-\infty$	13	16	$-\infty$

(Acyclic) Constrained Shortest Path Problem

Sabancı Üniversitesi

7 Ex. (ACSP):

$$\text{(P)} \quad \min_{p \in P_{19}} c^T \chi(p), \; \underbrace{a^T \chi(p) \geq 6}_{-Ax \leq -b}$$



Cor. (Pseudopolynomial Solution of the ACSP): The ACSP can be solved in pseudopolynomial time of $O(\prod_i |A_{i\cdot}|_1 \max \delta^+(v) + m)$, and if $A, b \geq 0$, in $O(\prod |b_j| \max \delta^+(v) + m)$.

Proof: Sort D topologically in linear time of $O(m)$, then fill the dynamic programming table in $O(\prod_i |A_{i\cdot}|_1 \max \delta^+(v))$.

Lagrange Relaxation

3.1 Def. (Lagrange(an) Relaxation): Let $c \in \mathbb{R}^n$, $D \in \mathbb{R}^{m \times n}, d \in \mathbb{R}^m$, $X \subseteq \mathbb{R}^n$ closed, and consider the optimization problem

$$\begin{array}{llll} \text{(P)} & \min & c^T x & \\ & & Dx = d & \text{complicated/ing} \\ & & x \in X & \text{tractable.} \end{array}$$

Let $\lambda \in \mathbb{R}^m$ a vector of **Lagrange multipliers**.

a) $L_P^{Dx=d}(\lambda) = \min_{x \in X} c^T x - \lambda^T (Dx - d) \in \mathbb{R} \cup \{\pm\infty\}$

Lagrange relaxation (of (P)) (w.r.t. $Dx = d$) at λ

b) $L_P^{Dx=d}\colon \mathbb{R}^m \to \mathbb{R} \cup \{\pm\infty\},\ \lambda \mapsto \min_{x \in X} c^T x - \lambda^T (Dx - d)$

Lagrange function (of (P)) (w.r.t. $Dx = d$)

Notation: If (P), $Dx = d$ are clear, we write L instead of $L_P^{Dx=d}$ etc.

Lagrange Relaxation

Sabancı Üniversitesi

3.2 Thm. (Properties of the Lagrange Relaxation, Geoffrion [1974]) Let (P) as in Def. 3.1 and

$$\nu(P) = \min c^T x,\ Dx = d,\ x \in X \quad \in \mathbb{R} \cup \{\pm\infty\}.$$

a) $\sup_{\lambda} L(\lambda) \leq \nu(P)$

b) Let $X = \{Ax \geq b\}, X \cap \{Dx = d\} \neq \emptyset$ for $A \in \mathbb{R}^{k\times n}, b \in \mathbb{R}^n$. Then

$$\max_{\lambda} L_P^{Dx=d}(\lambda) = \nu(P)$$

c) Let $X = \{Ax \geq b\} \cap \mathbb{Z}^n, X \cap \{Dx = d\} \neq \emptyset$ for $A \in \mathbb{R}^{k\times n}, b \in \mathbb{R}^n$.

Then $$\min_{Dx=d, Ax\geq b} c^T x \leq \max_{\lambda} L_P^{Dx=d}(\lambda) \leq \nu(P)$$

d) Let X be a $\begin{Bmatrix}\text{finite set}\\ \text{polytope}\end{Bmatrix}, X \cap \{Dx = d\} \neq \emptyset$.

Then L is i) concave, ii) piecewise affine, iii) bounded from above.

Lagrange Relaxation

3.2 Thm. (Properties of the Lagrange Relaxation, Geoffrion [1974]) Let (P) as in Def. 3.1 and

$$\nu(P) = \min c^T x,\ Dx = d,\ x \in X \quad \in \mathbb{R} \cup \{\pm\infty\}.$$

a) $\sup_{\lambda} L(\lambda) \leq \nu(P)$

Proof:

a) $L(\lambda) = \min_{x \in X}(c^T - \lambda^T D)x + \lambda^T d \leq \min_{\substack{x \in X \\ Dx = d}} c^T x - \lambda^T \underbrace{(Dx - d)}_{= 0} = \nu(P)$

3.2 Thm. (Properties of the Lagrange Relaxation, Geoffrion [1974]) b) Let $X = \{Ax \geq b\}, X \cap \{Dx = d\} \neq \emptyset$ for $A \in \mathbb{R}^{k \times n}, b \in \mathbb{R}^n$.

$$\max_{\lambda} L(\lambda) = \nu(P).$$

Proof: b)

$$\begin{array}{cccccc}
\nu(P) & = & \min & c^T x & = & \max & \lambda^T d + \mu^T b \\
 & & & Dx = d & \text{Duality} & & \lambda^T D + \mu^T A = c^T \\
 & & & Ax \geq b & \text{Thm.} & & \mu \geq 0
\end{array}$$

$$\begin{array}{cccccc}
= & \max & \lambda^T d & + & \max & \mu^T b \\
 & \lambda & & & \mu & \mu^T A = c^T - \lambda^T D \\
 & & & & & \mu \geq 0
\end{array}$$

$$\begin{array}{cccccc}
= & \max & \lambda^T d & + & \min & (c^T - \lambda^T D)x \\
\text{DT} & \lambda & & & & Ax \geq b
\end{array}$$

Lagrange Relaxation

3.2 Thm. (Properties of the Lagrange Relaxation, Geoffrion [1974]) Let (P) as in Def. 3.1 and

$$\nu(P) = \min c^T x,\ Dx = d,\ x \in X \quad \in \mathbb{R} \cup \{\pm\infty\}.$$

c) Let $X = \{Ax \geq b\} \cap \mathbb{Z}^n, X \cap \{Dx = d\} \neq \emptyset$ for $A \in \mathbb{R}^{k \times n}, b \in \mathbb{R}^n$.

Then $$\min_{Dx=d, Ax \geq b} c^T x \leq \max_{\lambda} L_P^{Dx=d}(\lambda) \leq \nu(P)$$

Proof:

c) Follows from a) and b).

3.2 Thm. (Properties of the Lagrange Relaxation, Geoffrion [1974]) Let (P) as in Def. 3.1 and

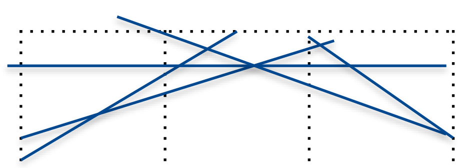
$$v(P) = \min c^T x,\ Dx = d,\ x \in X \quad \in \mathbb{R} \cup \{\pm\infty\}.$$

d) Let X be a $\left\{\begin{matrix}\text{finite set}\\ \text{polytope}\end{matrix}\right\}$, $X \cap \{Dx = d\} \neq \emptyset$. Then L is i) concave, ii) piecewise affine, iii) bounded from above.

Proof:

d) Let X be a $\left\{\begin{matrix}x_1, \dots, x_k\\ \text{conv}\,\{x_1, \dots, x_k\}\end{matrix}\right\}$. Then

$$L(\lambda) = \min_{x \in X}(c^T x - \lambda^T D)x + \lambda^T d = \underbrace{\min_{i=1,\dots,k}(\underbrace{c^T x_i - \lambda^T D)x_i + \lambda^T d}_{\text{affine in } \lambda}}_{\text{concave}}$$



$\leq v(P) < \infty$ as X is bounded and $\{Dx = d\} \cap X \neq \emptyset$. $\square$

Lagrange Relaxation (Inequality Version)

Sabancı Üniversitesi

3.3 Cor. (Lagrange(an) Relaxation): Let $c \in \mathbb{R}^n$, $D \in \mathbb{R}^{m\times n}, d \in \mathbb{R}^m$, $X \subseteq \mathbb{R}^n$ closed, and consider the optimization problem

$$\begin{array}{llll} \text{(P)} & \min & c^T x & \\ & & Dx \geq d & \text{complicated/ing} \quad \text{(standard form)} \\ & & x \in X & \text{tractable.} \end{array}$$

Let $\lambda \in \mathbb{R}^m_{\geq 0}$ a vector of **Lagrange multipliers**.

a) $L_P^{Dx=d}(\lambda) = \min_{x \in X}(c^T - \lambda^T D)x + \lambda^T d \ \in \mathbb{R} \cup \{\pm\infty\}$

Lagrange relaxation (of (P)) (w.r.t. $\boldsymbol{Dx \geq d}$) at $\boldsymbol{\lambda}$

b) $L_P^{Dx=d}\colon \mathbb{R}^m_{\geq 0} \to \mathbb{R} \cup \{\pm\infty\},\ \lambda \mapsto \min_{x \in X} (c^T - \lambda^T D)x + \lambda^T d$

Lagrange function (of (P)) (w.r.t. $\boldsymbol{Dx \geq d}$)

Proof: Ex. □

Subgradient Optimization

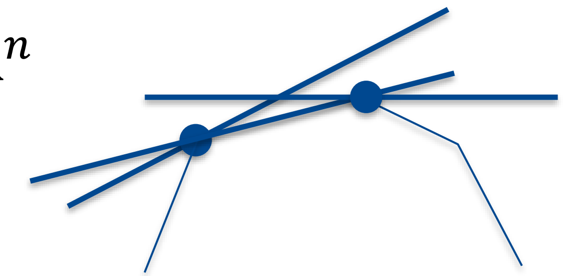
3.4 Def. (Subgradient, Subdifferential): Let $f: \mathbb{R}^n \to \mathbb{R}$ be concave, $\lambda_0 \in \mathbb{R}^n$.

a) $u \in \mathbb{R}^n$: $f(\lambda) \leq f(\lambda_0) + u^T(\lambda - \lambda_0) \quad \forall \lambda \in \mathbb{R}^n$

<u>u subgradient of f at λ_0</u>

b) $\partial f(\lambda_0) := \{u \in \mathbb{R}^n : u$ subgradient of f at $\lambda_0\}$

<u>subdifferential of f at λ_0</u>



3.5 Prop. (Sufficient Optimality Condition): Let $f: \mathbb{R}^n \to \mathbb{R}$ be concave, $\lambda_0 \in \mathbb{R}^m$. Then $0 \in \partial f(\lambda_0) \Longrightarrow f(\lambda_0) = \max f$.

Proof: Ex. ☐

3.6 Prop. (Diff'able Case): Let $f: \mathbb{R}^n \to \mathbb{R}$ be concave and diff'able at $\lambda_0 \in \mathbb{R}^n$. Then $\partial f(\lambda_0) := \{f'(\lambda_0)\}$.

Proof: Ex. ☐

Sabancı Üniversitesi

3.7 Prop. (Polyhedral Case): Let $f = L$ be a Lagrange function (as in Def. 3.1) and let $X = \begin{cases} x_1, \dots, x_k \\ \text{conv}\{x_1, \dots, x_k\} \end{cases}$. Then

$$\partial f(\lambda_0) = \text{conv}\{-(Dx_i - d) : x_i \in \text{Argmin}\, f(\lambda_0)\}.$$

Proof:

Let $\lambda_0 \in \mathbb{R}^n$, $X(\lambda_0) := \text{Argmin}\, f(\lambda_0)$, $u_i := -(Dx_i - d), i = 1, \dots, k$.

"$\supseteq$": $\forall x_j \in X(\lambda_0)$ holds

$$f(\lambda_0) + u_j^T(\lambda - \lambda_0) = \underbrace{c^T x_j - \lambda_0^T(Dx_j - d)}_{= f(\lambda_0)} - \underbrace{(Dx_j - d)^T}_{= u_j^T}(\lambda - \cancel{\lambda_0})$$

$$= c^T x_j - \lambda^T(Dx_j - d)$$

$$\geq \min_{i=1,\dots,k} c^T x_i - \lambda^T(Dx_i - d)$$

$$= f(\lambda_0).$$

3.7 Prop. (Polyhedral Case): Let $f = L$ be a Lagrange function (as in Def. 3.1) and let $X = \begin{Bmatrix} x_1, \ldots, x_k \\ \mathrm{conv}\,\{x_1, \ldots, x_k\} \end{Bmatrix}$. Then

$$\partial f(\lambda_0) = \mathrm{conv}\,\{-(Dx_i - d) : x_i \in \mathrm{Argmin}\, f(\lambda_0)\}.$$

Proof:

Let $\lambda_0 \in \mathbb{R}^n$, $X(\lambda_0) := \mathrm{Argmin}\, f(\lambda_0)$, $u_i := -(Dx_i - d), i = 1, \ldots, k$.

"$\subseteq$":
$$\min_{x_i \notin X(\lambda_0)} c^T x_i - \lambda_0^T (Dx_i - d) > f(\lambda_0)$$

$$\Longrightarrow \exists \epsilon > 0: \min_{x_i \notin X(\lambda_0)} c^T x_i - \lambda^T (Dx_i - d) > f(\lambda) \quad \forall \lambda \in U_\epsilon(\lambda_0)$$

$$\Longrightarrow X(\lambda) \subseteq X(\lambda_0) \quad \forall \lambda \in U_\epsilon(\lambda_0).$$

Subgradient Optimization

3.7 Prop. (Polyhedral Case): Let $f = L$ be a Lagrange function (as in Def. 3.1) and let $X = \begin{Bmatrix} x_1, \dots, x_k \\ \text{conv}\,\{x_1, \dots, x_k\} \end{Bmatrix}$. Then

$$\partial f(\lambda_0) = \text{conv}\,\{-(Dx_i - d) : x_i \in \text{Argmin}\, f(\lambda_0)\}.$$

Proof:

Let $\lambda_0 \in \mathbb{R}^n$, $X(\lambda_0) := \text{Argmin}\, f(\lambda_0)$, $u_i := -(Dx_i - d), i = 1, \dots, k$.

"$\subseteq$": $\exists \epsilon > 0 : X(\lambda) \subseteq X(\lambda_0) \forall \lambda \in U_\epsilon(\lambda_0)$. Let $\text{u} \notin \text{conv}\,\{u_i : x_i \in X(\lambda_0)\}$

$\Longrightarrow \exists \pi \in \mathbb{R}^m : \pi^T u < \pi^T u_i \quad \forall x_i \in X(\lambda_0)$ (by sep. hyperplane Thm.)

$$\Longrightarrow f(\underbrace{\lambda_0 + \epsilon\pi}_{=\lambda}) \geq \min_{x_i \in \underbrace{X(\lambda_0)}_{\supseteq X(\lambda_0 + \epsilon\pi)}} c^T x_i \underbrace{- (\lambda_0 + \epsilon\pi)^T (Dx_i - d)}_{= -\lambda_0^T (Dx_i - d) + \epsilon\pi^T u_i}$$

$$= \text{f}(\lambda_0) + \min_{x_i \in X(\lambda_0)} \epsilon\pi^T u_i$$

$$> \text{f}(\lambda_0) + \epsilon\pi^T u$$

$$= \text{f}(\lambda_0) + u^T (\overbrace{\lambda_0 + \epsilon\pi}^{=\lambda} - \lambda_0) \lightning \quad \square$$

Subgradient Algorithm

3.8 Alg. (Subgradient Algorithm):

Input:

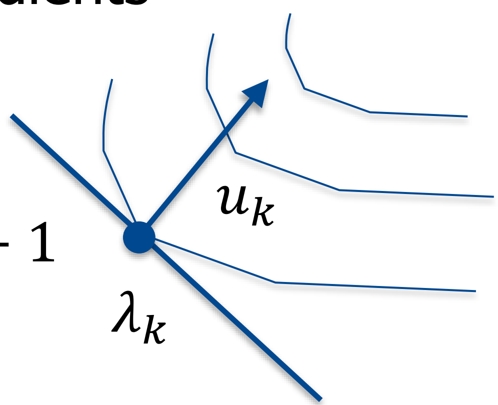
$f\colon \mathbb{R}^m \to \mathbb{R}$	concave (by func. & subgrad. oracle)
$\lambda_0 \in \mathbb{R}^m$	starting point
$(\alpha_k)_{k=1}^{\infty} > 0^*$	sequence of step lengths

Output: $(\lambda_k)_{k=1}^{\infty} \in (\mathbb{R}^m)^*$ iterates

Data St.:

$k \in \mathbb{N}_0$	iteration counter
$(u_k)_{k=1}^{\infty} \in (\mathbb{R}^m)^*$	sequence of subgradients

1. $k \leftarrow 0,\ u_0 \underset{\in}{\leftarrow} \partial f(\lambda_0)$
2. $\lambda_{k+1} \leftarrow \lambda_k + \alpha_k u_k,\ \ u_{k+1} \underset{\in}{\leftarrow} \partial f(\lambda_{k+1}),\ \ k \leftarrow k+1$
3. goto 2



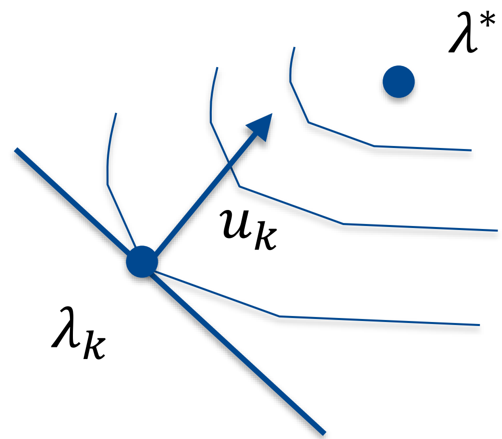
Convergence of the Subgradient Algorithm

3.9 Thm. (Convergence of the Subgradient Algorithm): Let $f\colon \mathbb{R}^m \to \mathbb{R}$ be concave and

a) $f^* = \max f < \infty$ (in particular, the maximum exists)

b) $||u||_2 \le L \quad \forall u \in \partial f$ for some $L \in \mathbb{R}$

c) $\sum_{k=0}^{\infty} \alpha_k^2 < \infty, \quad \sum_{k=0}^{\infty} \alpha_k \to \infty$.



Then $$\lim_{k\to\infty} \max_{j=1,\dots,k} f(\lambda_j) = f^*.$$

Proof: Let $\lambda^* \in \operatorname{Argmin} f$. Then

$$\begin{aligned} ||\lambda_{k+1} - \lambda^*||_2^2 &= ||\lambda_k + \alpha_k u_k - \lambda^*||_2^2 \\ &= ||\lambda_k - \lambda^*||_2^2 + 2\alpha_k \overbrace{u_k(\lambda_k - \lambda^*)}^{\le f(\lambda_k) - f^*} + \alpha_k^2 ||u_k||_2^2 \\ &\le ||\lambda_0 - \lambda^*||_2^2 - \sum_{j=1}^{k} 2\alpha_k \big(f^* - f(\lambda_j)\big) + \sum_{j=1}^{k} \alpha_k^2 ||u_k||_2^2 \end{aligned}$$

Convergence of the Subgradient Algorithm

3.9 Thm. (Convergence of the Subgradient Algorithm): Let $f: \mathbb{R}^m \to \mathbb{R}$ be concave s.t. a) $f^* = \max f < \infty$ b) $\exists L \in \mathbb{R}: ||u||_2 \leq L$ $\forall u \in \partial f$ c) $\sum_{k=0}^{\infty} \alpha_k^2 < \infty$, $\sum_{k=0}^{\infty} \alpha_k \to \infty$. Then $\max_{j=1,\ldots,k} f(\lambda_j) \xrightarrow[k\to\infty]{} f^*$.

Proof: Let $\lambda^* \in \operatorname{Argmin} f$. Then

$$
\begin{aligned}
||\lambda_{k+1} - \lambda^*||_2^2 &= ||\lambda_k + \alpha_k u_k - \lambda^*||_2^2 \\
&= ||\lambda_k - \lambda^*||_2^2 + \overbrace{2\alpha_k u_k(\lambda_k - \lambda^*)}^{\leq f(\lambda_k) - f^*} + \alpha_k^2 ||u_k||_2^2 \\
&\leq ||\lambda_0 - \lambda^*||_2^2 - \sum_{j=1}^{k} 2\alpha_j \big(f^* - f(\lambda_j)\big) + \sum_{j=1}^{k} \alpha_k^2 ||u_k||_2^2
\end{aligned}
$$

$$\Longrightarrow ||\lambda_{k+1} - \lambda^*||_2^2 + \sum_{j=1}^{k} 2\alpha_k \left(f^* - f(\lambda_j)\right) \leq ||\lambda_0 - \lambda^*||_2^2 + \sum_{j=1}^{k} \alpha_k^2 ||u_k||_2^2$$

$$\Longrightarrow \cancel{||\lambda_{k+1} - \lambda^*||_2^2} + \sum_{j=1}^{k} 2\alpha_j \left(f^* - f(\lambda_j)\right) \leq ||\lambda_0 - \lambda^*||_2^2 + \sum_{j=1}^{k} \alpha_k^2 ||u_k||_2^2$$

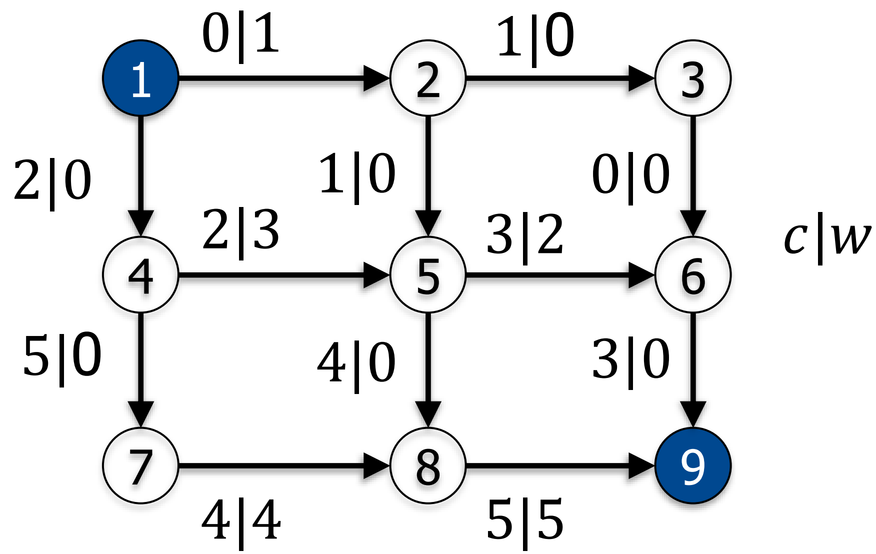
$$\Longrightarrow f^* - \max_{j=1,\ldots,k} f(\lambda_j) \leq \left(||\lambda_0 - \lambda^*||_2^2 + \sum_{j=1}^{k} \alpha_k^2 ||u_k||_2^2\right) / 2\sum_{j=1}^{k} \alpha_j \to 0. \ \square$$

Lagrange Relaxation

Sabancı Üniversitesi

3.10 Ex. (Lagrange Relaxation)

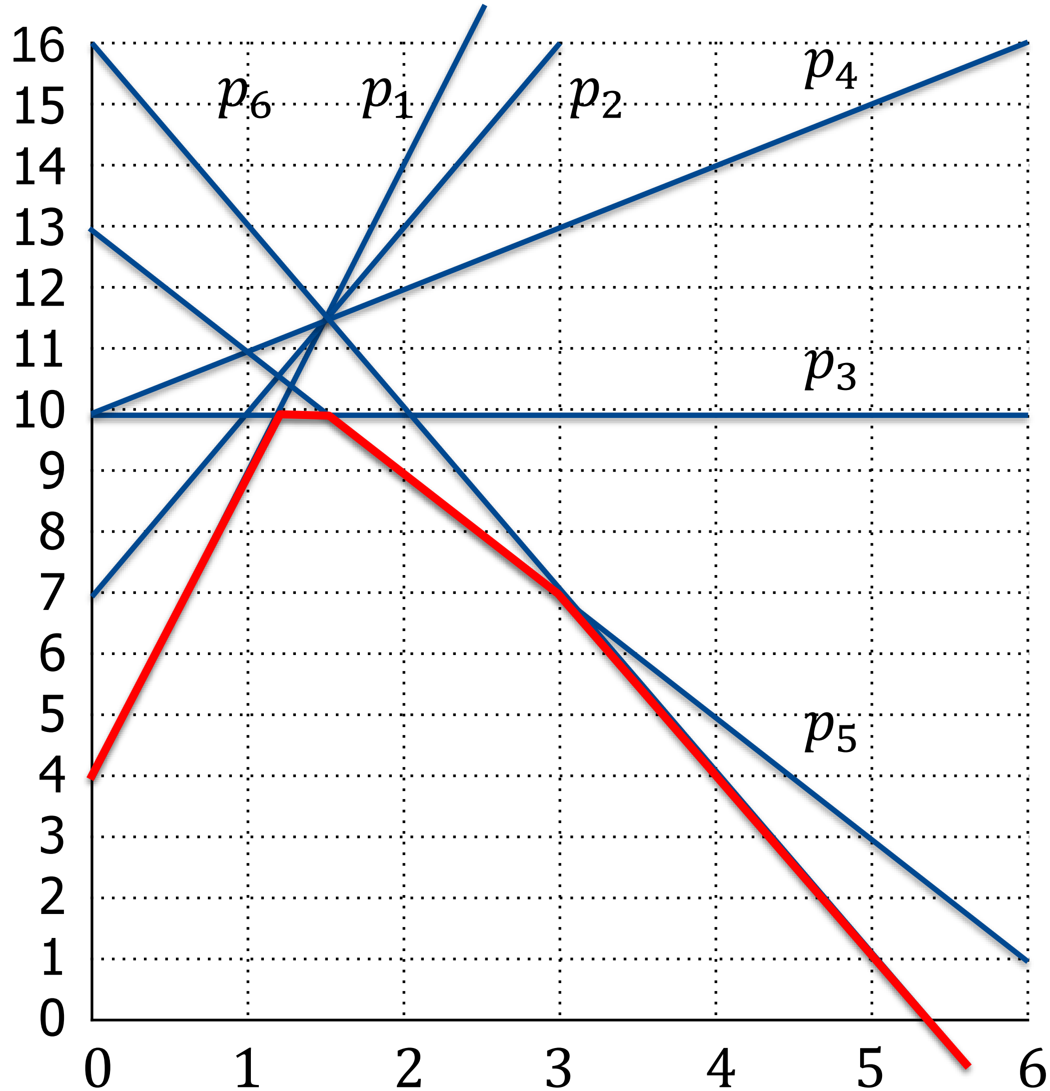
$$(\mathrm{P}) \quad \underbrace{\min_{p \in P_{19}}}_{X} c^T \chi(p), \quad \underbrace{w^T \chi(p) \geq 6}_{Dx \geq d}$$



p	nodes	$c(p)$	$w(p)$	$w(p)-6$	$c(p)-\lambda(w(p)-6)$	
p_1	12369	4	1	-5	$4-\lambda(-5)$	$=4+5\lambda$
p_2	12569	7	3	-3	$7-\lambda(-3)$	$=7+3\lambda$
p_3	12589	10	6	0	$10-\lambda 0$	$=10$
p_4	14569	10	5	-1	$10-\lambda(-1)$	$=10+\lambda$
p_5	14589	13	8	2	$13-\lambda(-1)$	$=13-2\lambda$
p_6	14789	16	9	3	$16-\lambda(3)$	$=16-3\lambda$

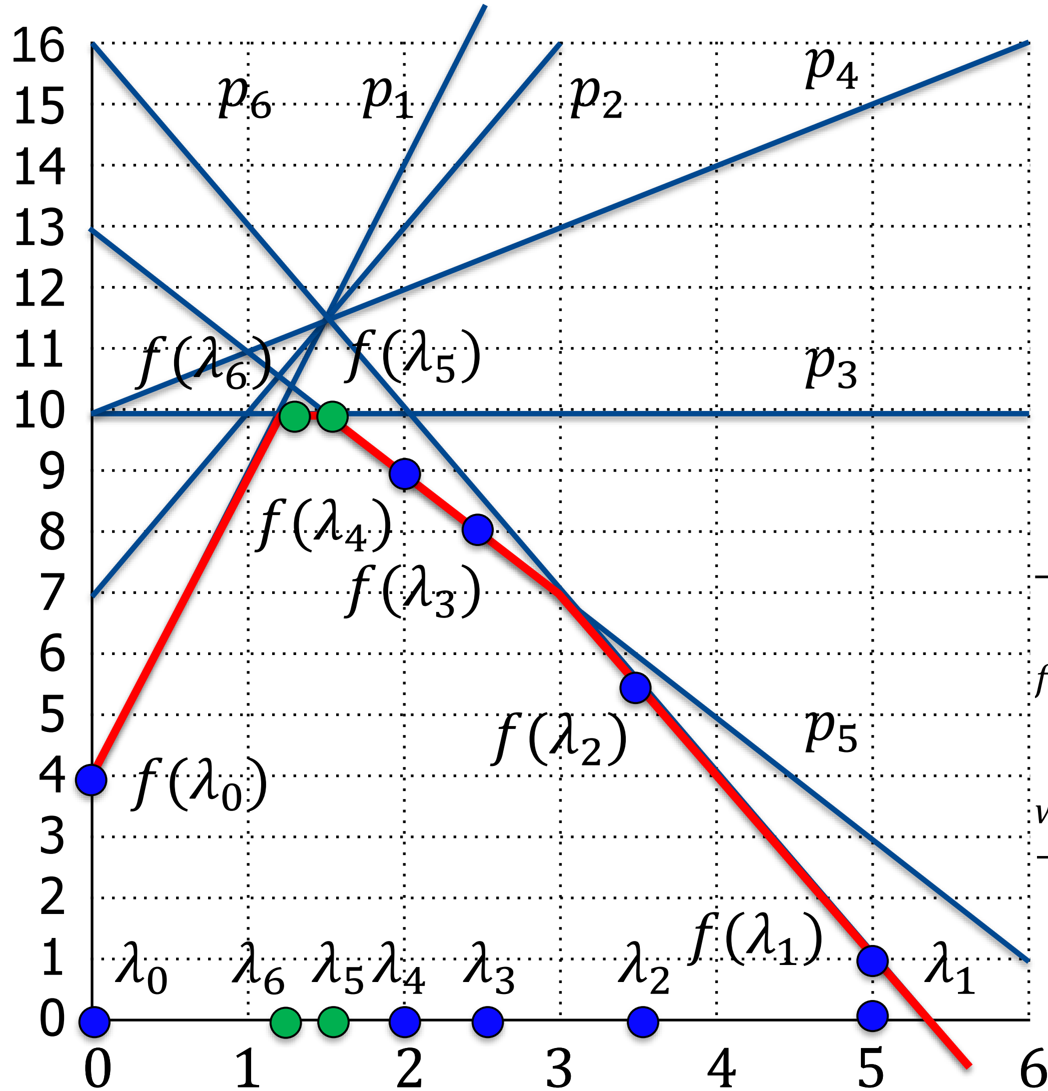
Sabancı Üniversitesi

3.10 Ex. (Lagrange Relaxation)



$\boldsymbol{p}$	$\boldsymbol{c(p) - \lambda(w(p) - 6)}$
p_1	$= 4 + 5\lambda$
p_2	$= 7 + 3\lambda$
p_3	$= 10$
p_4	$= 10 + \lambda$
p_5	$= 13 - 2\lambda$
p_6	$= 16 - 3\lambda$

3.10 Ex. (Lagrange Relaxation)



p	$c(p)$	$w(p)$
p_1	4	1
p_2	7	3
p_3	9	6
p_4	10	5
p_5	13	8
p_6	16	9

k	0	1	2	3	4	5	6
λ_k	0	5	$^7/_2$	$^5/_2$	2	$^8/_5$	$^{19}/_{15}$
$f(\lambda_k)$	4	1	$^{11}/_2$	8	9	10	10
p_k	p_1	p_6	p_6	p_5	p_5	p_5, p_3	p_3
$w(p_k) - 6$	-5	3	3	2	2	-2, 0	0
u_k	5	-3	-3	-2	-2	2,0	0
α_k	1	½	$^1/_3$	¼	$^1/_5$	$^1/_6$	$^1/_7$

Thank you for your attention

Prof. Dr. Güvenç Şahin
Sabancı University
Orta Mahalle
Üniversite Caddesi
No:27 Tuzla
34956 İstanbul
Turkey

guvencs@sabanciuniv.edu

myweb.sabanciuniv.edu/guvencs/