

Robust Artificial Intelligence: beyond pure worst-case optimization

Fabio D'Andreagiovanni

UNIMORE
UNIVERSITÀ DEGLI STUDI DI
MODENA E REGGIO EMILIA



Ciclo di seminari sull'AI,
Ordine degli Ingegneri della Provincia di Roma, 11 aprile 2024

About me



SAPIENZA
UNIVERSITÀ DI ROMA



- ✚ **2004: Laurea Triennale in Ingegneria Gestionale (summa cum laude)**
B.Sc. Thesis: “Assegnazione ottima di frequenze per tecniche di Frequency Hopping”
- ✚ **2006: Laurea Magistrale in Ingegneria Gestionale (summa cum laude)**
M.Sc. Thesis: "Instradamento ottimo con vincoli di QoS in reti di telecomunicazioni"
Awarded with:
 - Accenture Italia Master Thesis Award 2004/05
- ✚ **2010: Ph.D. in Operations Research (with excellence)**
Ph.D. Thesis: “Pure 0-1 Programming approaches to Wireless Network Design”
Awarded with:
 - INFORMS Doctoral Dissertation Award for Operations Research in Telecommunications 2010
 - Best Ph.D. Thesis - Sapienza Università di Roma 2010-2012

About me

Esperienza di **più di 15+ anni** su **teoria ed applicazioni di modelli ed algoritmi di Ottimizzazione Matematica ed Intelligenza Artificiale**



2006 - 2024

5 SOTTOPERIODI PRINCIPALI

Esperienza



SAPIENZA
UNIVERSITÀ DI ROMA



2006 - 2010

2010 - 2016

2016 - 2023

✚ **2006 - 2009: Research Assistant and Ph.D. Student**

Dipartimento di Informatica e Sistemistica "A. Ruberti", [Sapienza Università di Roma](#)

✚ **2009 – 2010: (Post-doctoral) Research Assistant**

Dipartimento di Informatica e Sistemistica "A. Ruberti", [Sapienza Università di Roma](#)

Esperienza



 **COLUMBIA UNIVERSITY**
IN THE CITY OF NEW YORK



**2008 -
2009**

2010 - 2016

2016 - 2023



2008 - 2009 : Research Scholar

Department of Industrial Engineering and Operations Research,
Columbia University in the City of New York
New York, USA



2006 - 2010

2010 - 2016

2016 - 2023

- ✚ 2010 – 2016: **Wissenschaftlicher Mitarbeiter = (Senior) Research Scientist**
 - 2010-2013: **Zuse Institute Berlin (ZIB)**, Berlin, Germany
 - 2013-2014: **Technische Universität Berlin**, Berlin, Germany
 - 2014-2016: **Zuse Institute Berlin (ZIB)**, Berlin, Germany
- ✚ 2015 – 2016: **Head of Research Group**, Zuse Institute Berlin (ZIB)



2006 - 2010

2010 - 2016

2016 - 2023

✚ 2016 - 2023: **Tenured Senior Research Scientist**

Centre National de la Recherche Scientifique (CNRS), France

**Laboratory HEUDIASYC (Computer Engineering)
of the University of Technology of Compiègne
(Grande École d'Ingénieurs)**





2006-2010

2010-2016

2016-2023

2023 –

✚ **2023: Professore Associato di Ingegneria Informatica**

Università degli Studi di Modena e Reggio Emilia

Dept. of Systems and Methods of Engineering

(currently on leave from CNRS)

UNIMORE
UNIVERSITÀ DEGLI STUDI DI
MODENA E REGGIO EMILIA



Non solo università e ricerca

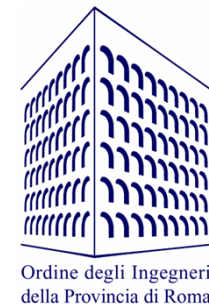
✚ Yes, I have been a consultant

- 2006 - 2007 Accenture Consulting



✚ Consulente (accademico) per aziende (studio e trasferimento di mathematical optimization models and algorithms)

✚ Più recentemente nominato **Coordinatore di Dipartimento**
Ordine degli Ingegneri - Provincia di Roma



Principali temi di ricerca

Teoria ed applicazioni di:

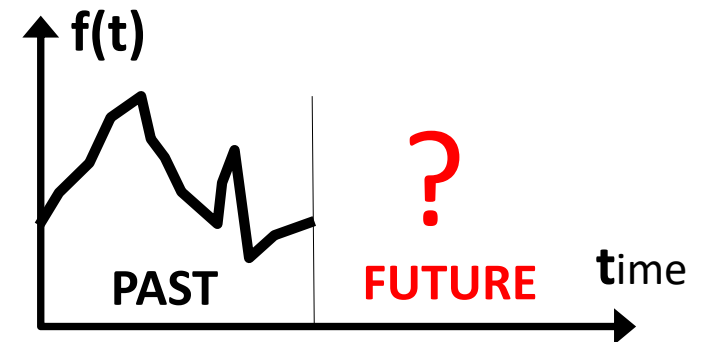
✚ Mixed Integer Linear Programming

- Polyhedral analysis (strong formulations)
- Exact Cutting-plane methods
- Fast hybrid exact-heuristics ([MATHeuristics](#))
- Hybridization with Machine Learning-based solution techniques ([Learning heuristics](#))

$$\begin{aligned} \max \quad & c' x \\ A x & \leq b \\ x & \in \mathbb{R}_+^n \times \mathbb{Z}_+^p \end{aligned}$$

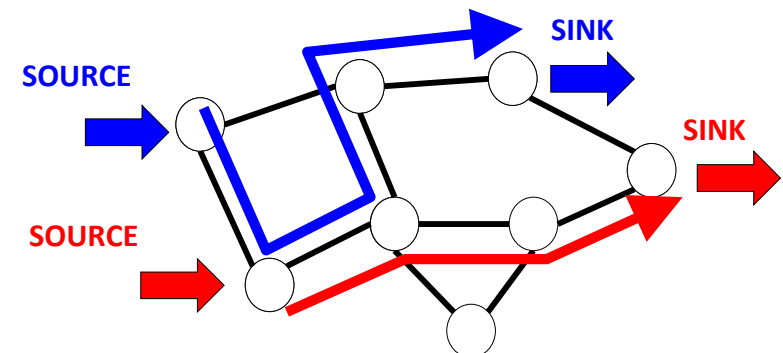
✚ Optimization under Data Uncertainty

- Robust Optimization
- Cardinality-constrained uncertainty sets



✚ Capacitated Network Design

- (Strong) valid inequalities characterization
- Efficient flow-routing algorithms



Temi di ricerca: real-world optimization

OBIETTIVO: colmare il "gap" tra teoria e pratica

+ Wireless Network Design

- User service coverage with quality-of-service guarantees
- Robustness against signal propagation uncertainty
- **Integration with drones** acting as 5G flying base stations



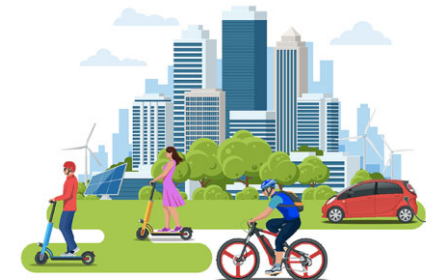
+ Design of computer networks

- Capacity and data routing design
- Robustness against traffic uncertainty and failures



+ Shared mobility in smart cities

- Gestione di flotte di veicoli condivisi (car- and e-scooter-sharing)
- Integrazione di misurazioni real-time di traffico fornite da reti di sensori



in numerosi progetti di ricerca industriale con e.g.

PART I

Robust Optimization

Un semplice problema di spostamento

Partiamo da un aneddoto:

- ✚ Lavorando come **scienziato CNRS in Francia**, mi sono spesso trovato a dover effettuare viaggi con breve preavviso tra Roma a Parigi
- ✚ Prenotare un **volo con compagnie di bandiera last-minute** era spesso **proibitivo in termini di costi e molto inefficiente in termini di tempo** (non esiste un volo diretto Lufthansa Roma FCO - Parigi CDG)
- ✚ L'unica alternativa efficiente in termini di tempo e costi era prendere un **volo Ryanair Roma Ciampino - Paris Beauvais**
- ✚ Con Ryanair dovevo prendere un volo la mattina presto e si presentava il problema di **scegliere il percorso in auto per raggiungere l'aeroporto**

Un semplice problema di spostamento (1)

Proviamo a formalizzare meglio il problema che affrontavo **per recarmi in aeroporto:**

✚ **OBIETTIVO DI OGNI VIAGGIO:**

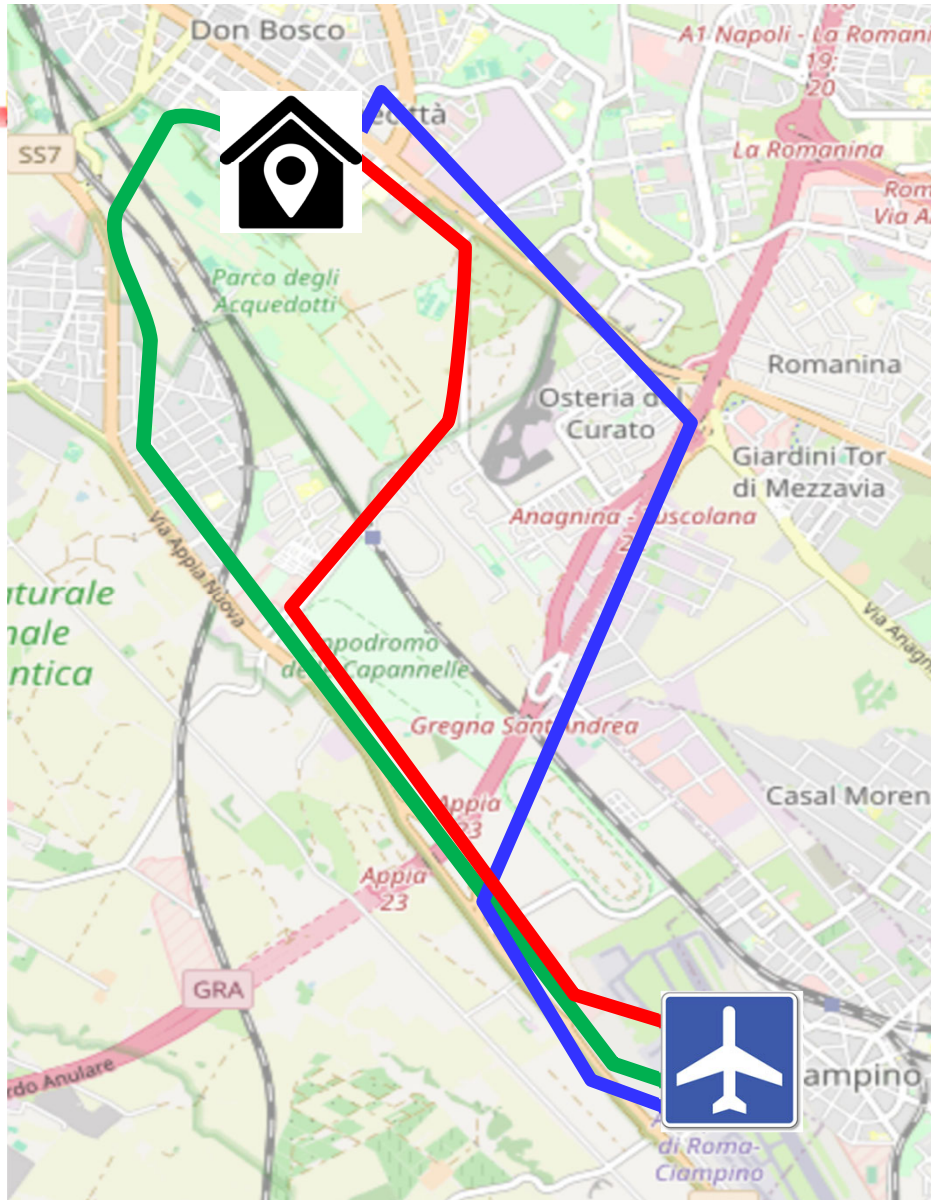
minimizzare i tempi di spostamento in auto da casa all'aeroporto

(Appio Claudio -> Aeroporto Ciampino)

✚ Effettivamente si tratta di un **PROBLEMA DI OTTIMIZZAZIONE:**

- **dati tutti i possibili percorsi** dal quartiere Appio Claudio all'Aeroporto di Ciampino
- si deve individuare il **percorso con il tempo di spostamento minimo**

Un semplice problema di spostamento (2)



Tre percorsi possibili:

- 1) tramite **Grande Raccordo Anulare (GRA)**
- 2) tramite **via Appia Nuova**
- 3) tramite **via delle Capannelle**

TEMPI DI PERCORRENZA

- | | |
|---------------------------------|----------|
| (1) via GRA | = 14 min |
| (2) via Appia Nuova | = 17 min |
| (3) via delle Capannelle | = 16 min |

**EVIDENTEMENTE
IL PERCORSO (1)
È IL PIÙ VELOCE**

L'impatto dell'incertezza di dati (1)

Il ragionamento che vi ho esposto ha però un difetto:



NON HO CONSIDERATO L'INCERTEZZA DEI DATI



- 1) i **tempi di percorrenza** non sono certi,
ma **variano** in base alle **condizioni di traffico**
- 2) ho fatto riferimento ai **tempi minimi di percorrenza**,
ottenuti in **condizioni di traffico trascurabile**
- 3) ho quindi definito una **soluzione troppo ottimista (BEST CASE)**



**Come possiamo migliorare
il processo decisionale?**



L'impatto dell'incertezza di dati (2)



✚ Ragioniamo in termini di **intervalli di tempo**, identificando il **tempo di percorrenza minimo e massimo**

✚ Tempi di percorrenza intervalli [min, max]

(1) via GRA [14,37] min

(2) via Appia Nuova [17,34] min

(3) via delle Capannelle [16,40] min

Se consideriamo i tempi massimi,

- il percorso (1) non è più il migliore
- il percorso (2) diventa il più veloce (richiede 3 minuti in meno)

L'impatto dell'incertezza di dati (3)

Possiamo ritenerci **soddisfatti considerando i tempi massimi?**

EFFETTIVAMENTE, NO

VANTAGGIO: siamo protetti da condizioni di traffico sfavorevoli
(FULL PROTECTION)

SVANTAGGIO: siamo **troppo pessimisti (WORST CASE)**
è (altamente) improbabile che tutte le vie percorse
siano contemporaneamente soggette alle peggiori
condizioni possibili di traffico

L'impatto dell'incertezza di dati (3)



Come possiamo migliorare
il processo decisionale?



Il nostro obiettivo è **superare i limiti dell'approccio WORST-CASE**
(troppo pessimista e conservativo)

Ora potete comprendere meglio il sottotitolo del seminario:
"Beyond pure-worst case optimization"

In questo caso ci vengono in aiuto:

- la **(ROBUST) MATHEMATICAL OPTIMIZATION**
- l'**ARTIFICIAL INTELLIGENCE**

Verso l'ottimizzazione matematica

Per migliorare il nostro processo decisionale possiamo ricorrere a:
modelli e algoritmi di ottimizzazione matematica

Essenzialmente, il nostro compito è:

- tradurre il problema di ottimizzazione in **termini matematici (modellazione)**
- risolvere il modello tramite un **algoritmo di ottimizzazione integrato con metodi di Artificial Intelligence (soluzione)**

Applicazioni diverse ma con simile struttura matematica



AIRCRAFT FLEET
ROUTE MANAGEMENT



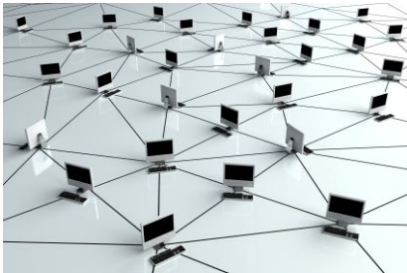
ENERGY
PRODUCTION



PRODUCTION LINE
MANAGEMENT



ROAD ROUTING
PLANNING



DESIGN OF TELECOM
NETWORKS



SUPPLY CHAIN
MANAGEMENT

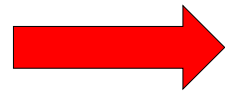


SURGERY
SCHEDULING

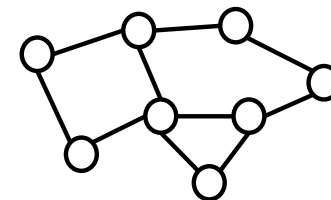


PUBLIC TRANSPORT
SCHEDULING

**COSA HANNO
IN COMUNE?**



PROBLEMI DI OTTIMIZZAZIONE SU RETI



Problema di ottimizzazione su reti - una descrizione informale

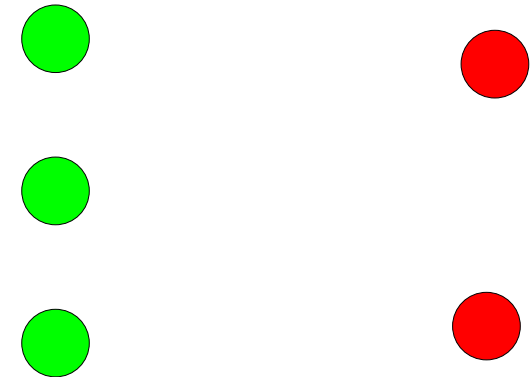
Tipicamente, ogni problema di ottimizzazione su reti può essere ricondotto al problema matematico di instradare un flusso su una rete
(problem of routing flows through a network)

Problema di ottimizzazione su reti - una descrizione informale

Tipicamente, ogni problema di ottimizzazione su reti può essere ricondotto al problema matematico di instradare un flusso su una rete (problem of routing flows through a network)

PIÙ PRECISAMENTE:

- esistono dei punti di SUPPLY e dei punti di DEMAND
- dobbiamo trasferire una quantità di beni da ogni punto di supply ad ogni punto di demand

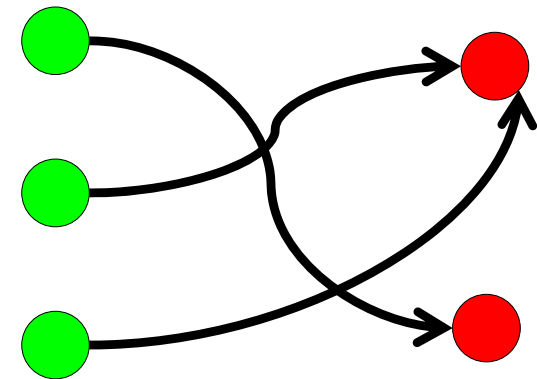


Problema di ottimizzazione su reti - una descrizione informale

Tipicamente, ogni problema di ottimizzazione su reti può essere ricondotto al problema matematico di instradare un flusso su una rete (problem of routing flows through a network)

PIÙ PRECISAMENTE:

- ✚ esistono dei **punti di SUPPLY** e dei **punti di DEMAND**
- ✚ dobbiamo **trasferire una quantità di beni** da ogni punto di supply ad ogni punto di demand
- ✚ delle **rotte (o cammini - routes)** connettono i punti

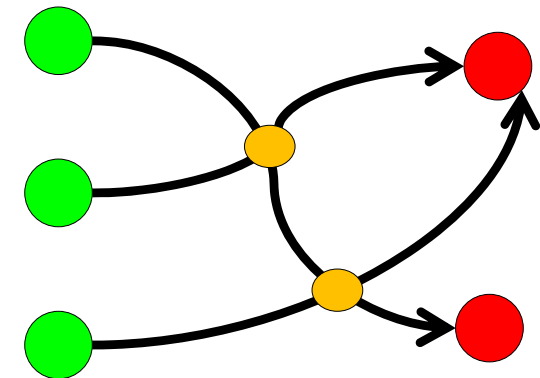


Problema di ottimizzazione su reti - una descrizione informale

Tipicamente, ogni problema di ottimizzazione su reti può essere ricondotto al problema matematico di instradare un flusso su una rete (problem of routing flows through a network)

PIÙ PRECISAMENTE:

- ✚ esistono dei **punti di SUPPLY** e dei **punti di DEMAND**
- ✚ dobbiamo **trasferire una quantità di beni** da ogni punto di supply ad ogni punto di demand
- ✚ delle **rotte (o cammini - routes)** connettono i punti
- ✚ i cammini possono contenere dei **punti di diramazione intermedi** (intermediate transshipment points)

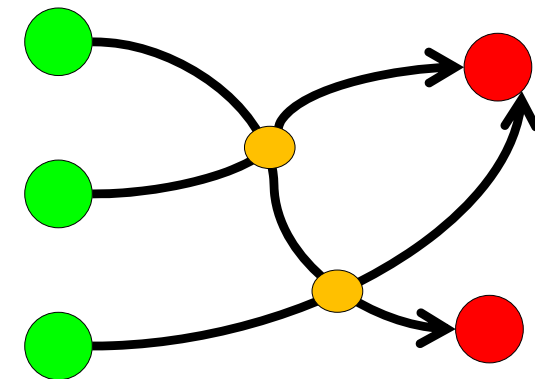


Problema di ottimizzazione su reti - una descrizione informale

Tipicamente, ogni problema di ottimizzazione su reti può essere ricondotto al problema matematico di instradare un flusso su una rete (problem of routing flows through a network)

PIÙ PRECISAMENTE:

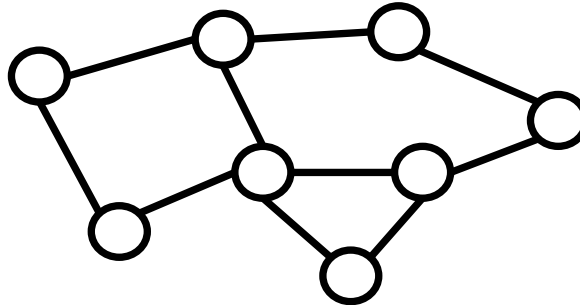
- ✚ esistono dei **punti di SUPPLY** e dei **punti di DEMAND**
- ✚ dobbiamo **trasferire una quantità di beni** da ogni punto di supply ad ogni punto di demand
- ✚ delle **rotte (o cammini - routes)** connettono i punti
- ✚ i cammini possono contenere dei **punti di diramazione intermedi** (intermediate transshipment points)
- ✚ dobbiamo **soddisfare dei vincoli associati** con i punti e le rotte (e.g., domanda di beni di un punto, massima capacità della rotta)
- ✚ se trasferiamo beni attraverso una rotta, dobbiamo **sostenere un costo**



Il **problema di ottimizzazione matematica** che affrontiamo è quello di **trasferire dei beni dai punti di supply ai punti di demand, rispettando i vincoli e minimizzando il costo totale**

Modellare reti tramite grafi (1)

Modellare una rete tramite un grafo è una scelta naturale



Un **grafo diretto** $G(N,A)$ è una **struttura matematica** caratterizzata da:

- un **insieme di nodi** N

- un **insieme di archi** A

(ognuno corrispondente ad una coppia di nodi distinti $(i,j) \neq (j,i)$)

- un **arco** (i,j) è una coppia di nodi ordinata, i.e.



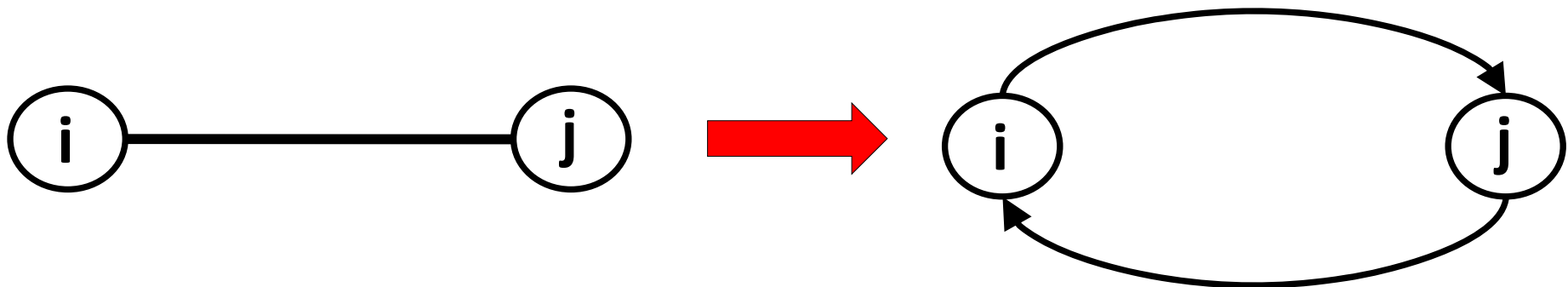
Modellare reti tramite grafi (2)

In moltissime applicazioni reali, **l'ordine dei nodi di un arco è importante**, quindi è molto comune usare grafi diretti

Nella teoria dei grafi, è però comune impiegare **grafi indiretti**, in cui **l'ordine dei nodi non è importante**



È però sempre possibile sostituire un ARCO INDIRETTO con due ARCHI DIRETTI



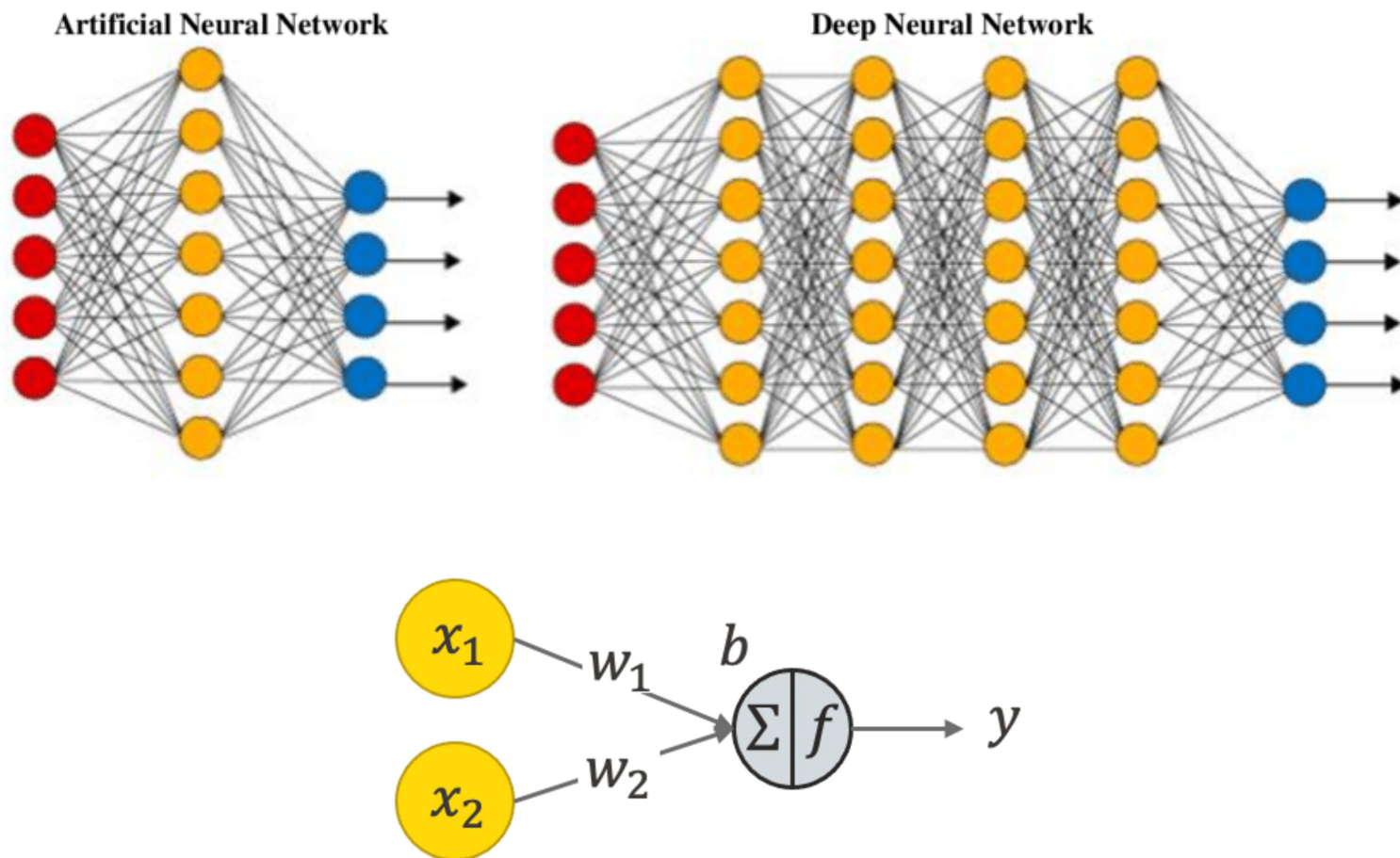
Modellare reti tramite grafi (3)

Nella seguente tabella, sono riportate le corrispondenze tra gli elementi di problemi reali di ottimizzazione su reti e gli elementi di un grafo

Nodes	Arcs	Flow
Intersections	Roads	Vehicles
Airports	Air lanes	Aircraft
Switching points	Wires, channels	Messages
Pumping stations	Pipes	Fluids
Work centers	Material-handling sources	Jobs

Grafi alla base di Artificial Neural Networks (3)

- ✚ È importante ricordare che i grafi sono anche alla base delle dei modelli **Neural Networks**



Mathematical Optimization (1)

OTTIMIZZAZIONE MATEMATICA

“The mathematical art of doing better”

Problemi di decisione complessi sono tradotti in termini matematici e vengono risolti da computer per trovare soluzioni di altissima qualità (al di là delle possibilità umane)

COMPONENTI FONDAMENTALI di un modello di ottimizzazione:

$$\max c'x$$



OBJECTIVE FUNCTION
(FUNZIONE OBIETTIVO)

$$Ax \leq b$$



FEASIBILITY CONSTRAINTS
(VINCOLI DI AMMISSIBILITÀ)

$$x \geq 0$$



DECISION VARIABLES
(VARIABILI DECISIONALI)

Mathematical Optimization 2)

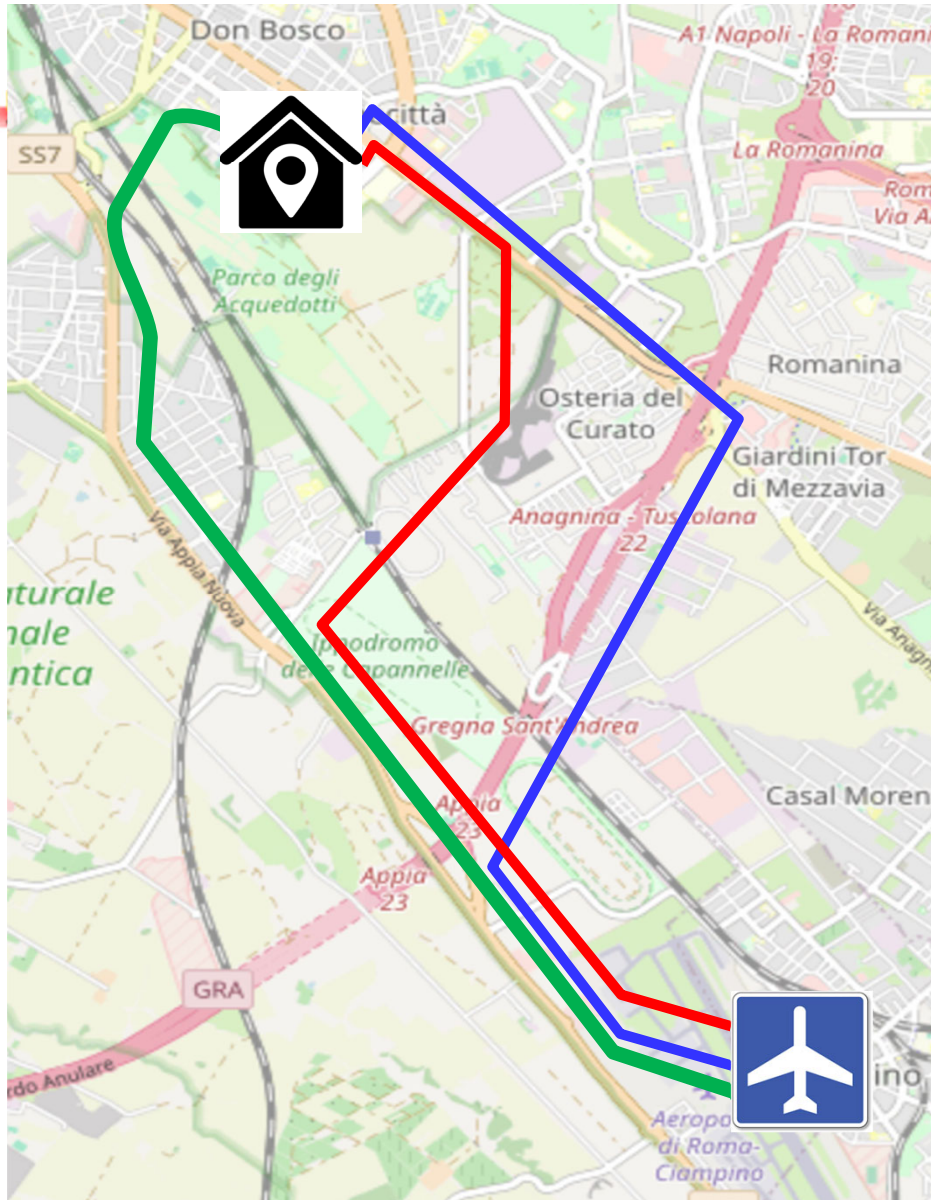
- 1) **Identificare chiaramente** il problema di ottimizzazione
- 2) **Formalizzare una descrizione essenziale** del problema ed **elencarne gli elementi** fondamentali e la loro **notazione matematica**
- 3) Definire le **variabili decisionali**
(rappresentanti le decisioni che possono essere prese)
- 4) Definire i **vincoli di ammissibilità**
(rappresentanti i limiti imposti alle decisioni)
- 5) Definire la **funzione obiettivo**
(rappresentante l'obiettivo da raggiungere)



può sembrare una procedura naturale da eseguire
ma spesso viene violata
(persino in lavori scientifici!!!)

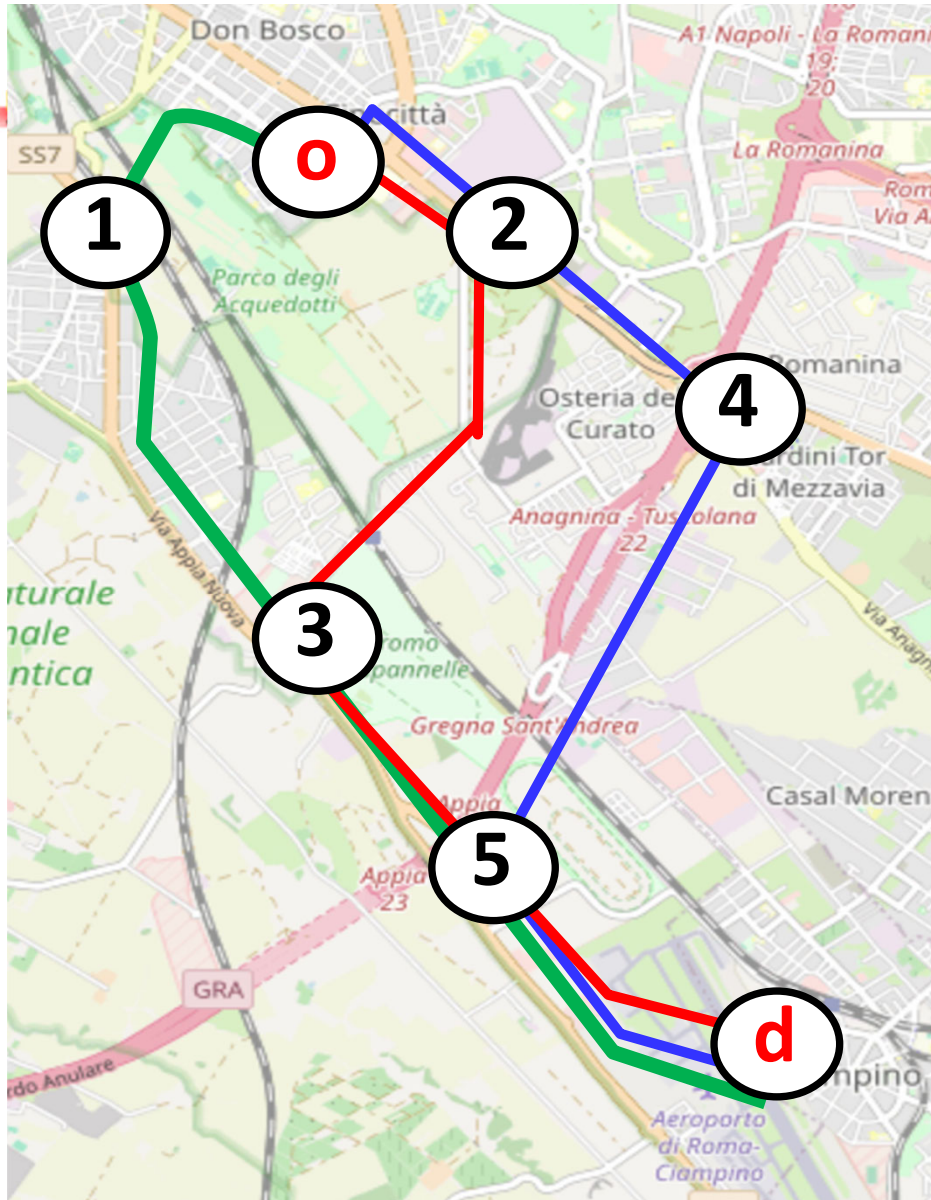


PASSO 1 - Identificare il problema



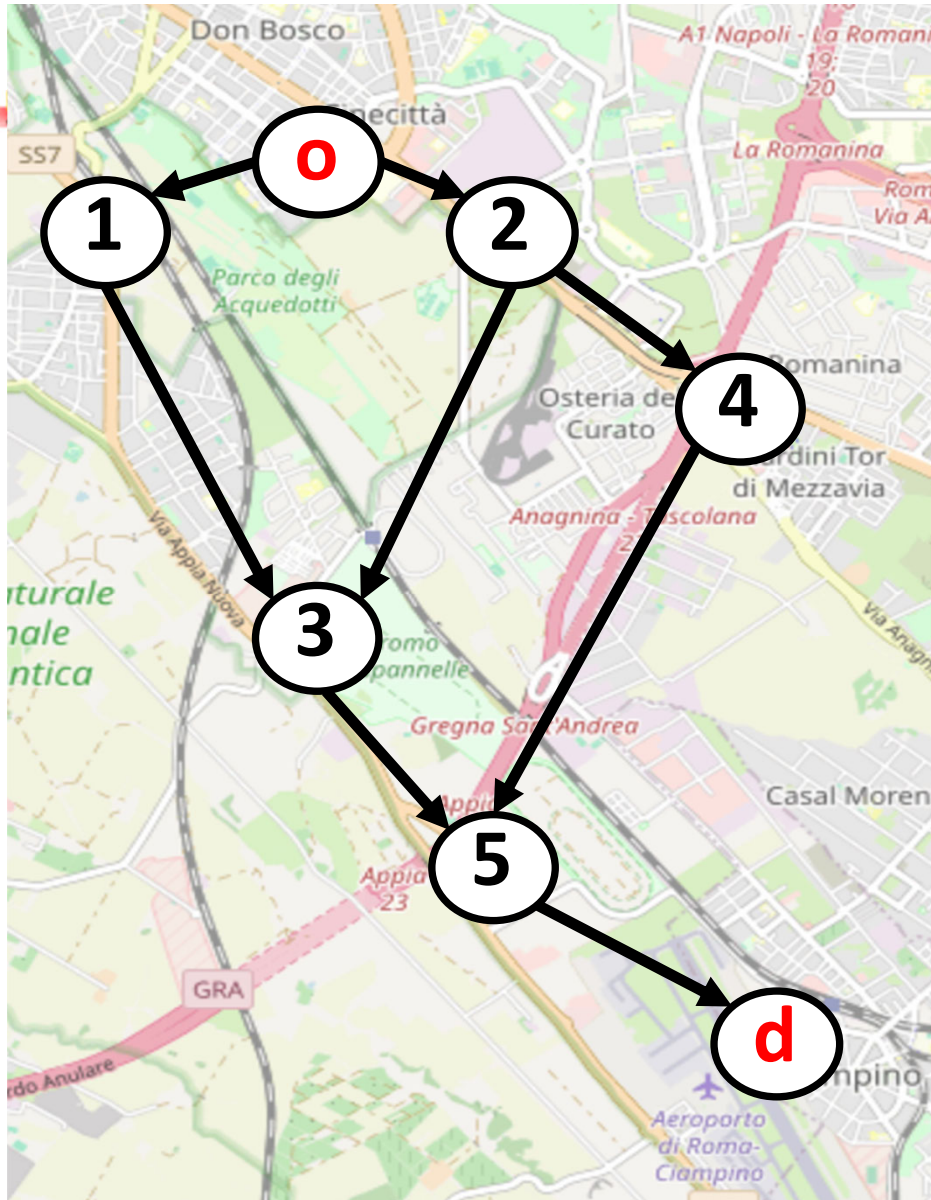
- Il problema consiste nel **trovare il percorso più veloce** per spostarsi da casa in aeroporto impiegando l'automobile
- È naturale **modellare la rete stradale**, impiegata per spostarsi, **tramite un grafo**
- NOTA: per semplicità, consideriamo solo i tratti ed intersezioni stradali principali

PASSO 1 - Identificare il problema



- Come primo passo, possiamo individuare i punti di intersezione stradale principali del problema ed associarli a nodi

PASSO 1 - Identificare il problema



- Il problema consiste nel trovare il percorso più veloce per spostarsi da casa in aeroporto impiegando l'automobile
- È naturale **modellare la rete stradale**, impiegata per spostarsi, tramite un **grafo diretto $G(V,A)$** con:
 $V = \{0, 1, 2, 3, 4, 5, d\}$
 $A = \{(0,1), (0,2), (1,3), (2,3), (2,4), (3,5), (4,5), (5,d)\}$
- NOTA: per semplicità, consideriamo solo i tratti ed intersezioni stradali principali

Passo 2: descrizione essenziale e notazione matematica

DATI:

- la rete stradale modellata tramite il **grafo** $G(V,A)$
- il **tempo di percorrenza** t_{ij} di ogni arco (i,j) del grafo
- i nodi dello spostamento
 - origine o
 - destinazione d

DESIDERIAMO:

individuare un cammino (i.e., una sequenza di archi) dal nodo origine " o " al nodo destinazione " d "

TALE CHE:

il tempo di percorrenza totale del cammino sia minimizzato

(tempo totale = SOMMA tempi degli archi inclusi nel cammino)

Passo 3: variabili decisionali

DECISIONE FONDAMENTALE:

Dato un arco stradale $(i,j) \in A$ dobbiamo decidere se percorrerlo o meno per spostarci dal nodo origine o al nodo destinazione d

Si tratta di una **decisione binaria** (i.e., acceso/spento, 0/1) e la modelliamo quindi con **VARIABILI BINARIE** (anche dette VARIABILI BOOLEANE).

Nello specifico:

- definiamo una variabile binaria $x(i,j)$ per ogni arco (i,j) del grafo, i.e.:

$$x(i,j) \in \{0,1\} \quad \forall (i,j) \in A$$

tale che:

$$x(i,j) = \begin{cases} 1 & \text{se l'arco } (i,j) \text{ è usato per spostarsi da o a d} \\ 0 & \text{se l'arco } (i,j) \text{ NON è usato per spostarsi da o a d} \end{cases}$$

Passo 4: vincoli di ammissibilità (1)

VINCOLO FONDAMENTALE:

se entriamo in un nodo stradale, dobbiamo uscirne
eccetto nel caso del:

- nodo origine (da cui usciamo solamente)
- nodo destinazione (in cui entriamo solamente)

Passo 4: vincoli di ammissibilità (1)

VINCOLO FONDAMENTALE:

se entriamo in un nodo stradale, dobbiamo uscirne

eccetto nel caso del:

- nodo origine (da cui usciamo solamente)
- nodo destinazione (in cui entriamo solamente)

Possiamo modellare questo aspetto con i c.d. **VINCOLI DI CONSERVAZIONE DEL FLUSSO** (riferiti alla **metafora della rete in cui scorre un flusso da una sorgente origine ad un pozzo destinazione**), i quali impongono che:

per ogni nodo, la somma dei flussi entranti deve essere uguale alla somma dei flussi uscenti

Passo 4: vincoli di ammissibilità (1)

VINCOLO FONDAMENTALE:

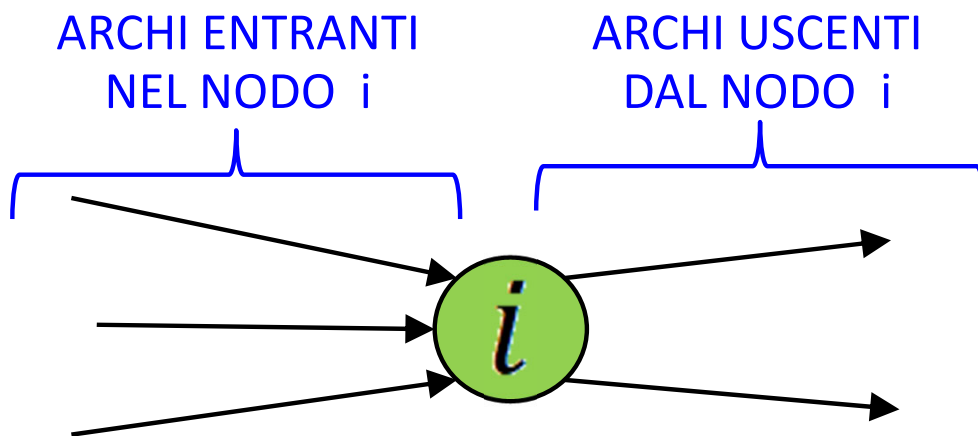
se entriamo in un nodo stradale, dobbiamo uscirne

eccetto nel caso del:

- nodo origine (da cui usciamo solamente)
- nodo destinazione (in cui entriamo solamente)

Possiamo modellare questo aspetto con i c.d. **VINCOLI DI CONSERVAZIONE DEL FLUSSO** (riferiti alla **metafora della rete in cui scorre un flusso da una sorgente origine ad un pozzo destinazione**), i quali impongono che:

per ogni nodo, la somma dei flussi entranti deve essere uguale alla somma dei flussi uscenti



$$\sum \text{FLUSSI ARCHI ENTRANTI} = \sum \text{FLUSSI ARCHI USCENTI}$$

equivalente a:

$$\sum \text{FLUSSI ARCHI ENTRANTI} - \sum \text{FLUSSI ARCHI USCENTI} = 0$$

Passo 4: vincoli di ammissibilità (2)

Nel caso del nostro problema i **VINCOLI DI CONSERVAZIONE DEL FLUSSO** diventano:

$$\sum_{(j,i) \in A} x(j,i) - \sum_{(i,j) \in A} x(i,j) = 0$$

PER OGNI NODO
 $i \neq o, d$
SI ENTRA E SI ESCE

$$\sum_{(o,j) \in A} x(o,j) = 1$$

NODO ORIGINE o :
SI ESCE SOLAMENTE
(ATTIVANDO SOLO UNO TRA
GLI ARCHI USCENTI)

$$\sum_{(i,o) \in A} x(i,o) = 1$$

NODO DESTINAZIONE d :
SI ENTRA SOLAMENTE
(ATTIVANDO SOLO UNO TRA
GLI ARCHI ENTRANTI)

Passo 5: funzione obiettivo

OBIETTIVO:

minimizzare il tempo totale per spostarsi dall'origine alla destinazione

Questa condizione può essere espressa dalla **funzione obiettivo**:

$$\text{MIN} \sum_{(i,j) \in A} t(i,j) \cdot x(i,j)$$

che somma il tempo di percorrenza $t(i,j)$ di tutti gli archi stradali che vengono percorsi (i.e., gli archi (i,j) la cui corrispondente variabile binaria è attivata $x(i,j) = 1$)

Il modello di ottimizzazione completo

$$\text{MIN} \sum_{(i,j) \in A} t(i,j) \cdot x(i,j)$$

$$\sum_{(j,i) \in A} x(j,i) - \sum_{(i,j) \in A} x(i,j) = 0$$

PER OGNI NODO
 $i \neq o, d$
SI ENTRA E SI ESCE

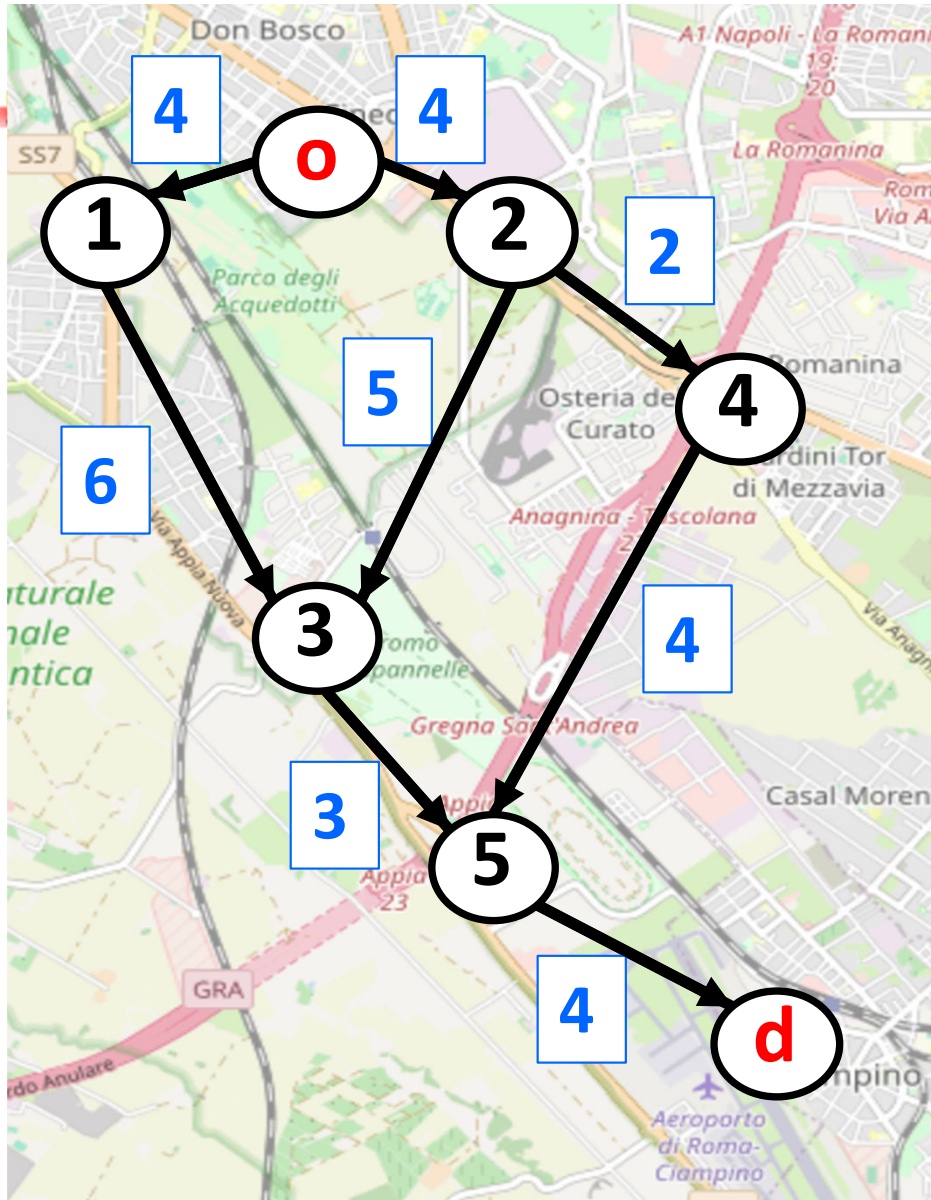
$$\sum_{(o,j) \in A} x(o,j) = 1$$

NODO ORIGINE o :
SI ESCE SOLAMENTE
(ATTIVANDO SOLO UNO TRA
GLI ARCHI USCENTI)

$$\sum_{(i,o) \in A} x(i,o) = 1$$

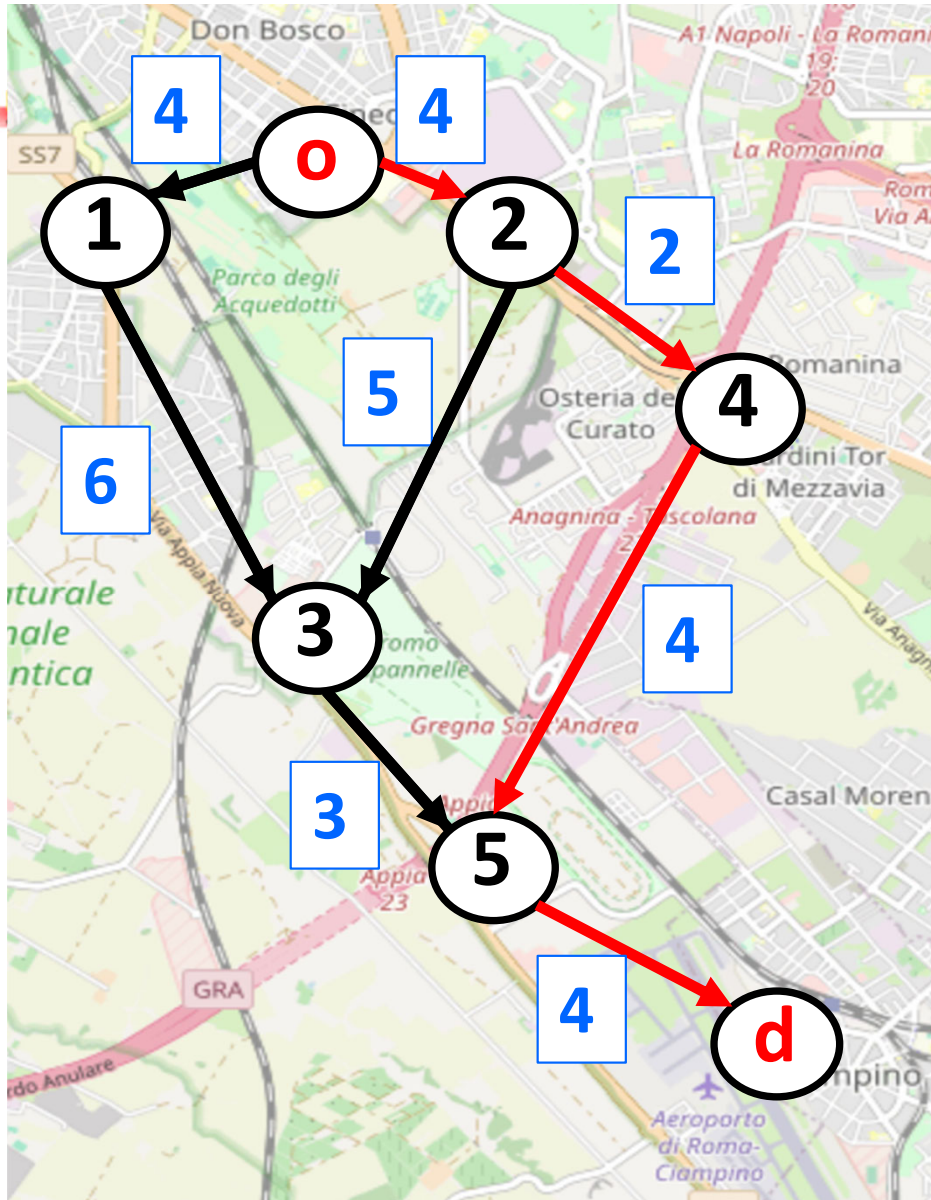
NODO DESTINAZIONE d :
SI ENTRA SOLAMENTE
(ATTIVANDO SOLO UNO TRA
GLI ARCHI ENTRANTI)

La soluzione ottima (BEST CASE)



✚ In questo grafo, ogni arco ha associato il suo **tempo minimo di percorrenza (best case)**

La soluzione ottima (BEST CASE)



✚ In questo grafo, ogni arco ha associato il suo **tempo minimo di percorrenza (best case)**

Il **percorso ottimo** è:

(o,2) (2,4) (4,5) (5,d)

in cui le variabili decisionali hanno valore:

$$x(o,1) = 0$$

$$x(o,2) = 1$$

$$x(1,3) = 0$$

$$x(2,3) = 0$$

$$x(2,4) = 1$$

$$x(3,5) = 0$$

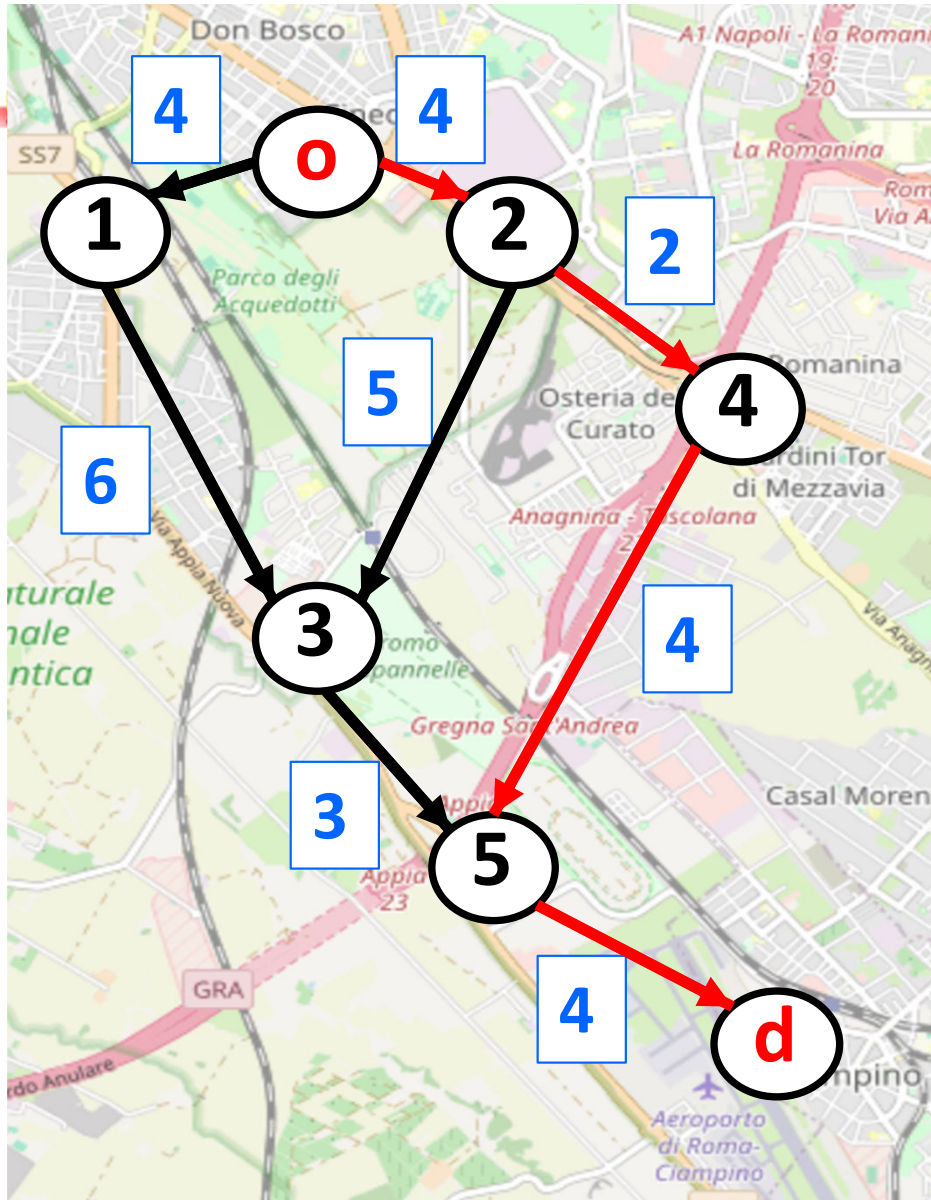
$$x(4,5) = 1$$

$$x(5,d) = 1$$

ed il tempo totale è:

$$4 + 2 + 4 + 4 = 14 \text{ minuti}$$

La soluzione ottima (BEST CASE)



✚ In questo grafo, ogni arco ha associato il suo **tempo minimo di percorrenza (best case)**

Il **percorso ottimo** è:

(o,2) (2,4) (4,5) (5,d)

in cui le variabili decisionali hanno valore:

$$x(o,1) = 0$$

$$x(o,2) = 1$$

$$x(1,3) = 0$$

$$x(2,3) = 0$$

$$x(2,4) = 1$$

$$x(3,5) = 0$$

$$x(4,5) = 1$$

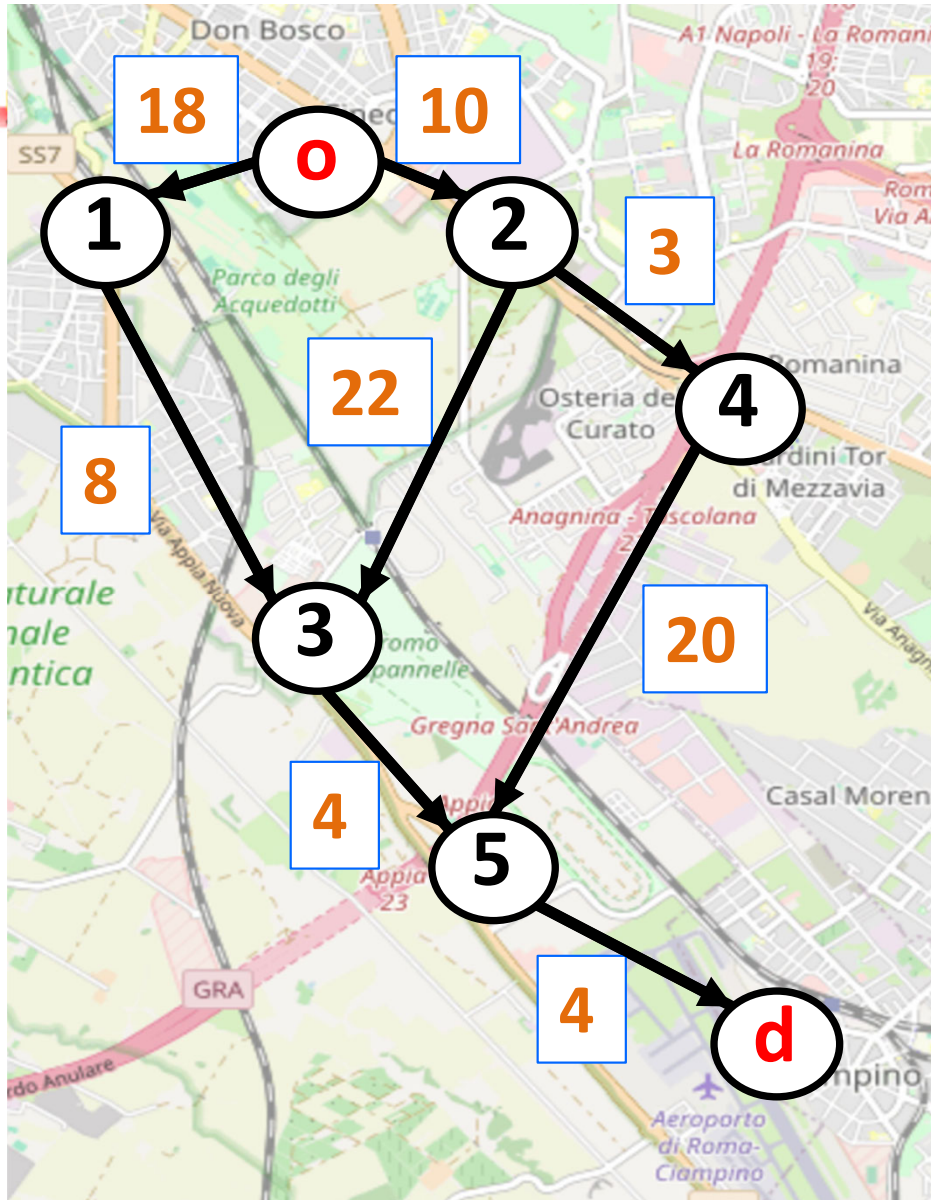
$$x(5,d) = 1$$

ed il tempo totale è:

$$4 + 2 + 4 + 4 = 14 \text{ minuti}$$

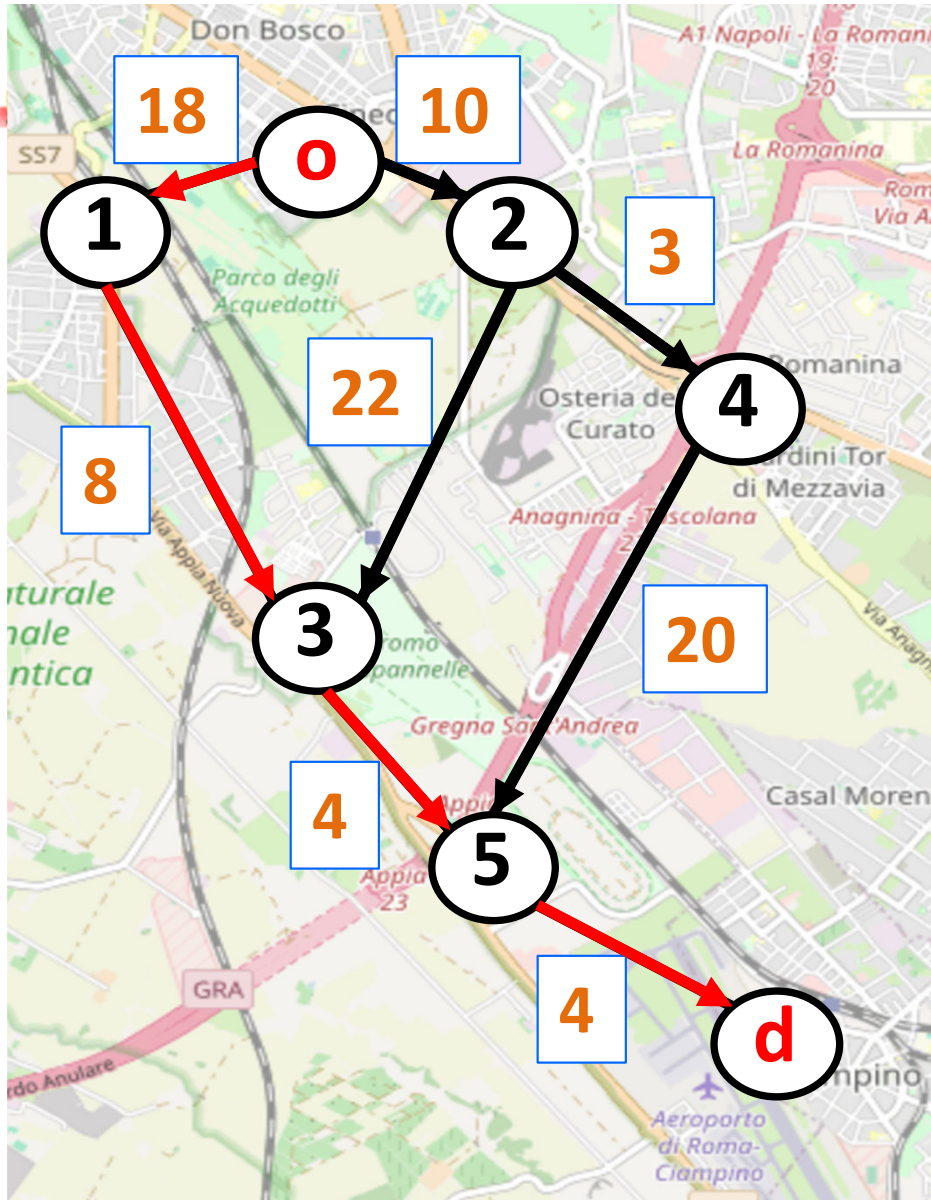
SI PUÒ RISOLVERE
EFFICIENTEMENTE
QUESTO PROBLEMA CON IL
FAMOSO ALGORITMO DI
DIJKSTRA

La soluzione ottima (WORST CASE)



✚ In questo grafo, ogni arco ha associato il suo **tempo massimo di percorrenza (worst case)**

La soluzione ottima (WORST CASE)



In questo grafo, ogni arco ha associato il suo **tempo massimo di percorrenza (worst case)**

Il **percorso ottimo** è CAMBIATO:

(o,1) (1,3) (3,5) (5,d)

in cui le variabili decisionali hanno valore:

$$x(o,1) = 0$$

$$x(o,2) = 1$$

$$x(1,3) = 0$$

$$x(2,3) = 0$$

$$x(2,4) = 1$$

$$x(3,5) = 0$$

$$x(4,5) = 1$$

$$x(5,d) = 1$$

ed il tempo totale è:

$$18 + 8 + 4 + 4 = 34 \text{ minuti}$$

NUOVAMENTE,
SI PUÒ RISOLVERE
EFFICIENTEMENTE
QUESTO PROBLEMA CON IL
FAMOSO ALGORITMO DI
DIJKSTRA

Raffinare la rappresentazione dell'incertezza

SITUAZIONE

Ora sappiamo che

- l'approccio **BEST-CASE** è troppo ottimista
- l'approccio **WORST-CASE** è troppo pessimista

Entrambi si basano su un'ipotesi **ALL-OR-NOTHING**:

tutti gli archi sono congestionati (tempo massimo -worst case)

OPPURE

tutti gli archi **NON** sono congestionati (tempo minimo - best case)

Raffinare la rappresentazione dell'incertezza

SITUAZIONE

Ora sappiamo che

- l'approccio **BEST-CASE** è troppo ottimista
- l'approccio **WORST-CASE** è troppo pessimista

Entrambi si basano su un'ipotesi **ALL-OR-NOTHING**:

tutti gli archi sono congestionati (tempo massimo -worst case)

OPPURE

tutti gli archi **NON** sono congestionati (tempo minimo - best case)



Come possiamo migliorare
il processo decisionale?



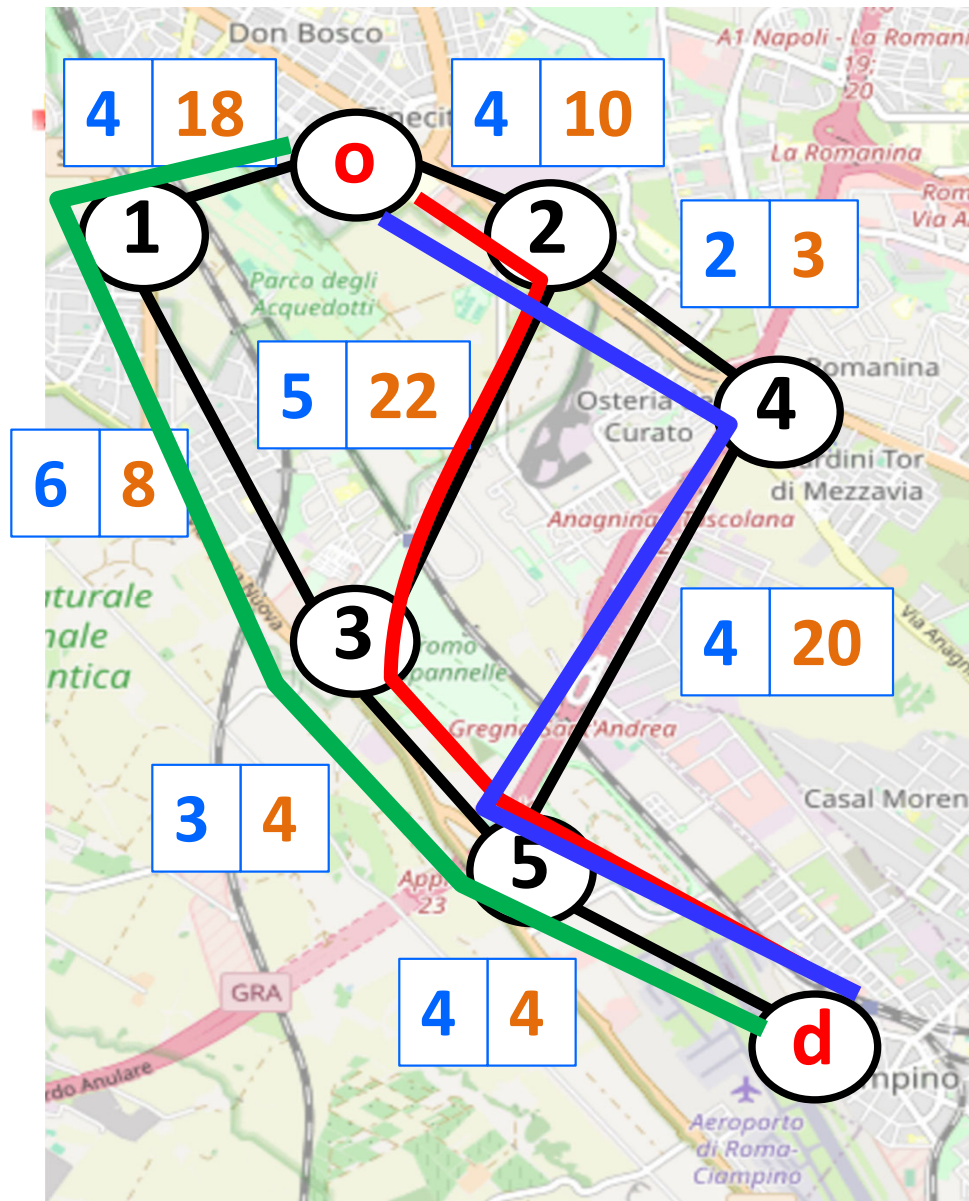
**INTRODUCIAMO IL CONCETTO DI
BUDGET DI INCERTEZZA (UNCERTAINTY BUDGET)**



Budget di incertezza

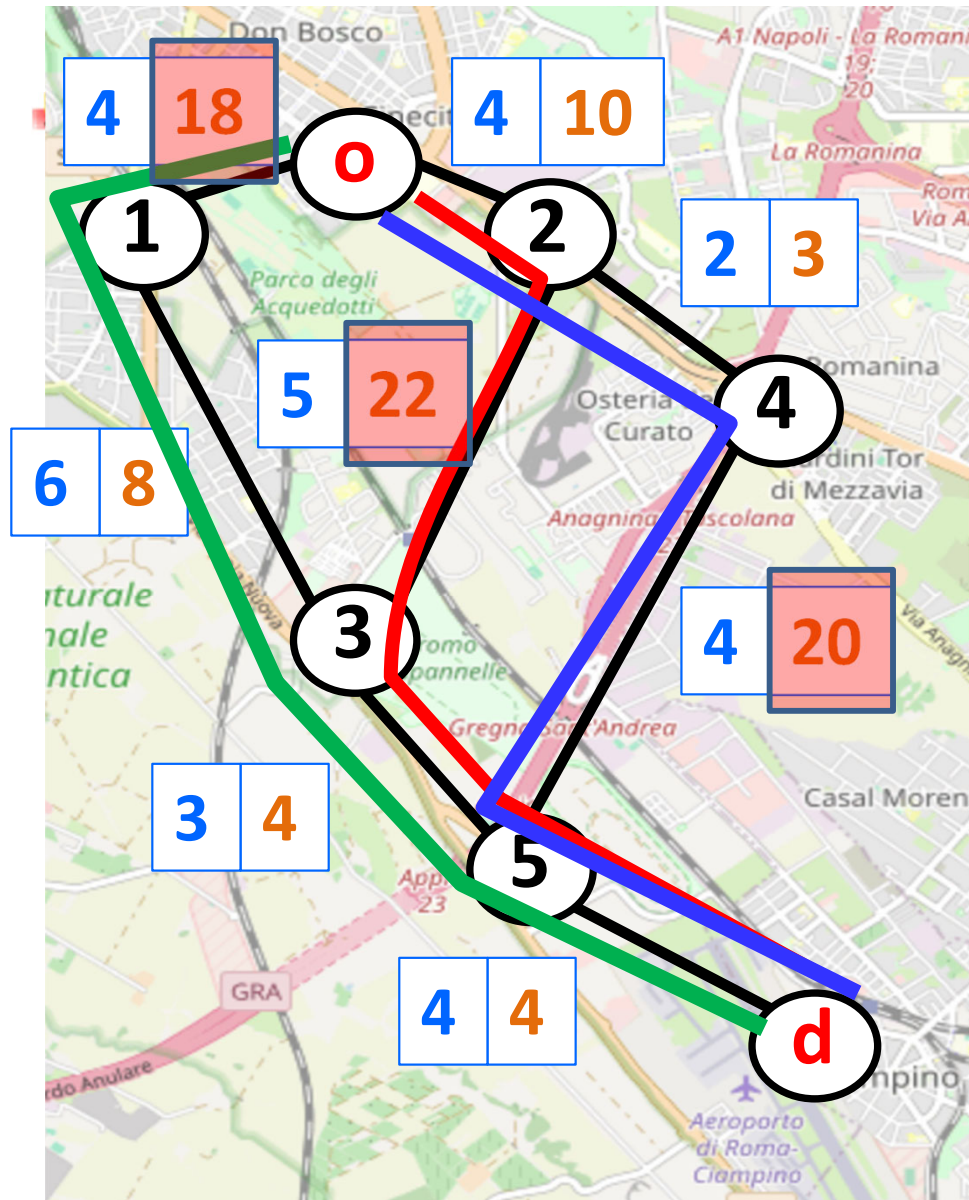
- Il **budget di incertezza** rappresenta il **numero di archi worst-case** che ammettiamo di incontrare nel problema
- Il budget di incertezza **ci permette interventi "più chirurgici"** introducendo l'incertezza ma senza arrivare agli eccessi di un puro worst case
- Nel nostro problema, **rappresenta il numero di archi stradali congestionati associati con il tempo massimo di percorrenza**
- **Questi archi "worst-case" non sono conosciuti a priori, ma dipendono dalla specifica soluzione**

Impatto del budget di incertezza ($\Gamma = 1$) (1)



- In questo grafo, ogni arco ha associato il suo **tempo minimo e massimo di percorrenza**

Impatto del budget di incertezza ($\Gamma = 1$) (1)



✚ In questo grafo, ogni arco ha associato il suo **tempo minimo e massimo di percorrenza**

Se fissiamo $\Gamma = 1$ permettiamo ad un solo arco di avere tempo massimo e l'arco colpito da questo incremento dipende dal percorso.

PERCORSO:

(1) via **GRA**

arco (4,5) passa da 4 a 20 minuti

(2) via **Appia Nuova**

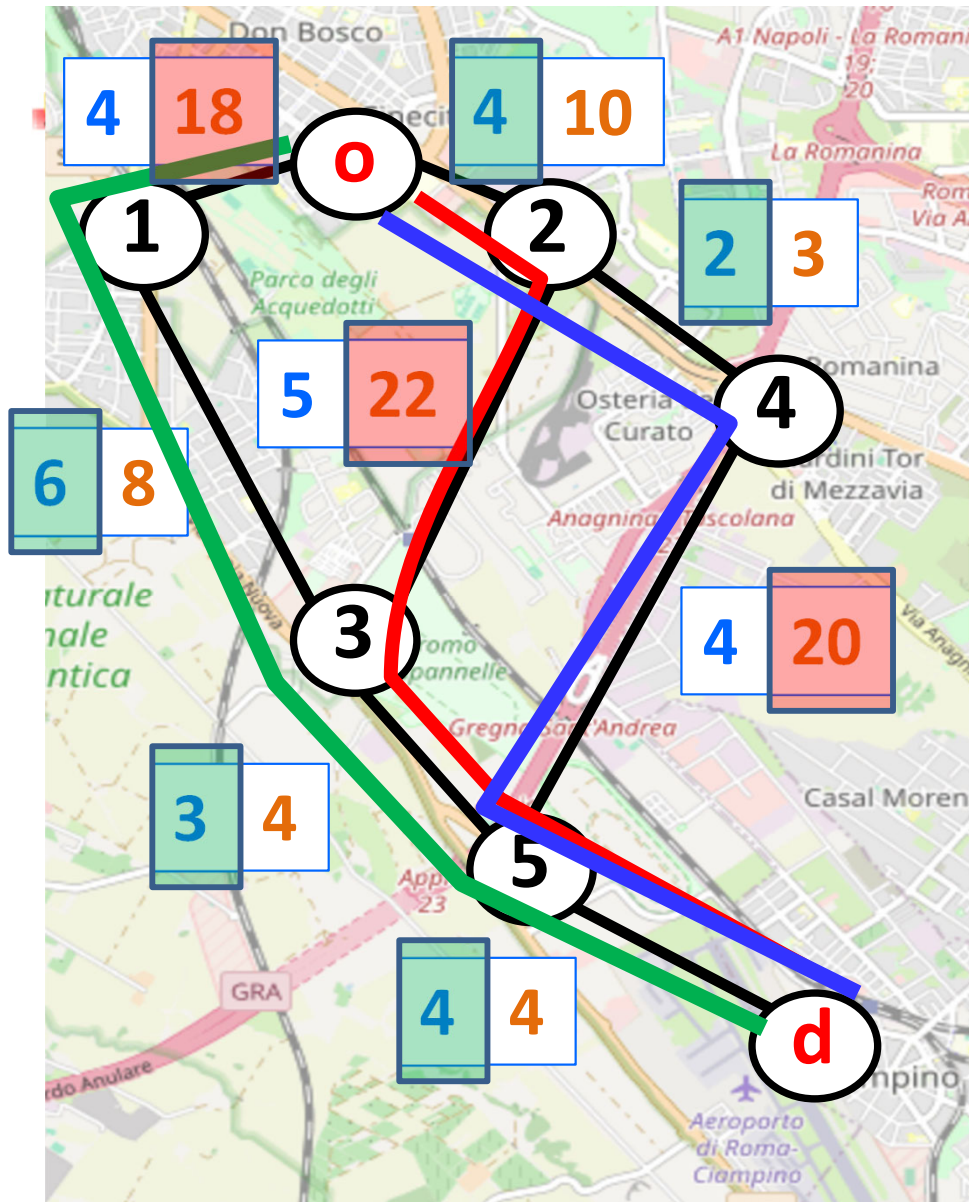
arco (0,1) passa da 4 a 18 minuti

(3) via delle **Capannelle**

arco (2,3) passa da 5 a 22 minuti

ARCHI "WORST" TUTTI DIVERSI!

Impatto del budget di incertezza ($\Gamma = 1$) (2)



✚ Se fissiamo $\Gamma = 1$ permettiamo ad un solo arco di avere tempo massimo e l'arco colpito da questo incremento dipende dal percorso.

✚ Tempi di percorrenza =

Tempo best + variazione $\Gamma = 1$

(1) via GRA $4+2+20+4 = 30$

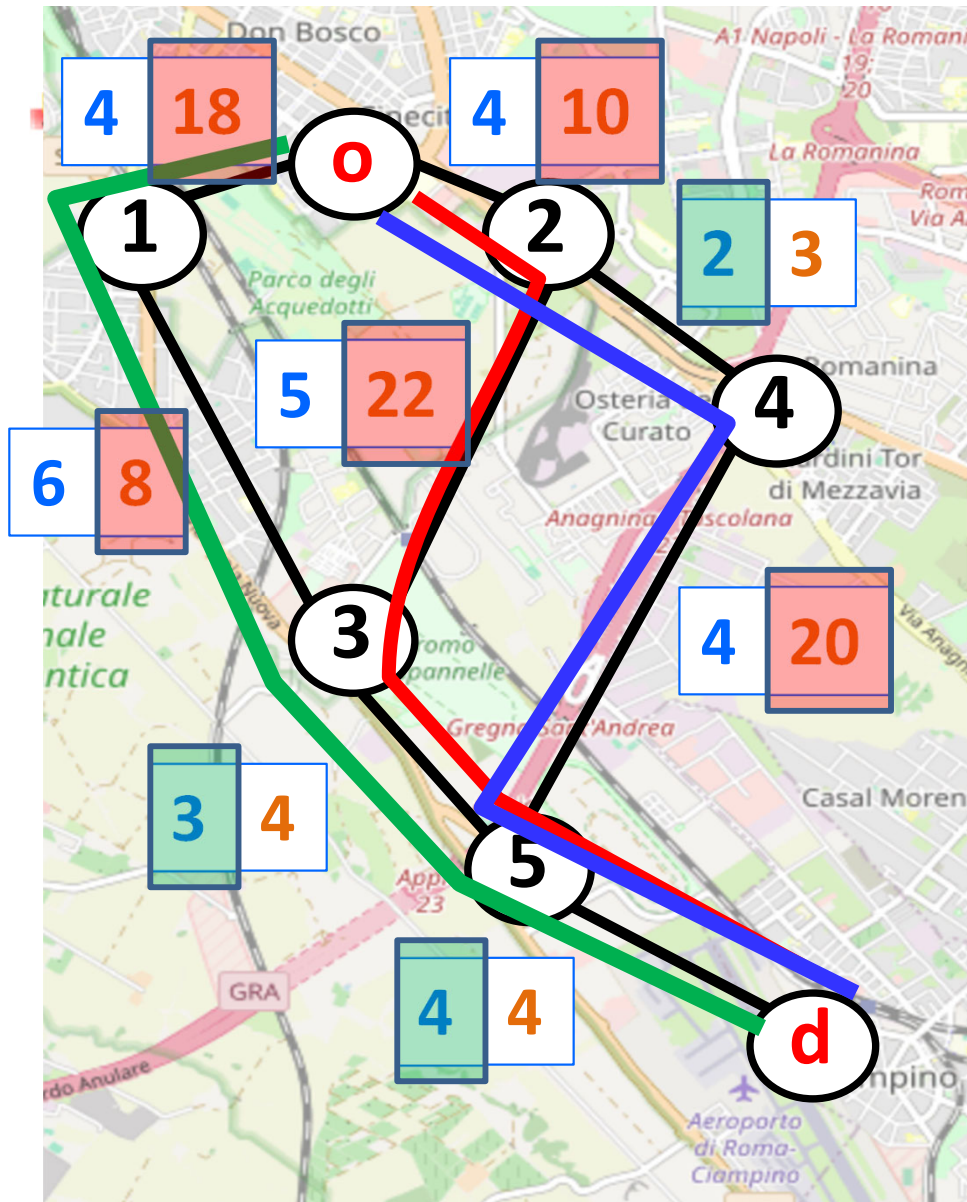
(2) via Appia Nuova $18+6+3+4 = 31$

(3) via delle Capannelle $4+22+3+4 = 33$

Per $\Gamma = 1$ (1 singolo arco a tempo max)

- il percorso (1) torna ad essere il migliore (rispetto al worst-case)

Impatto del budget di incertezza ($\Gamma = 2$)



✚ Se fissiamo $\Gamma = 2$ permettiamo a 2 archi di avere tempo massimo e gli archi colpiti da questo incremento dipendono dal percorso.

✚ Tempi di percorrenza =

Tempo best + variazione $\Gamma = 2$

(1) via GRA $10+2+20+4 = 36$

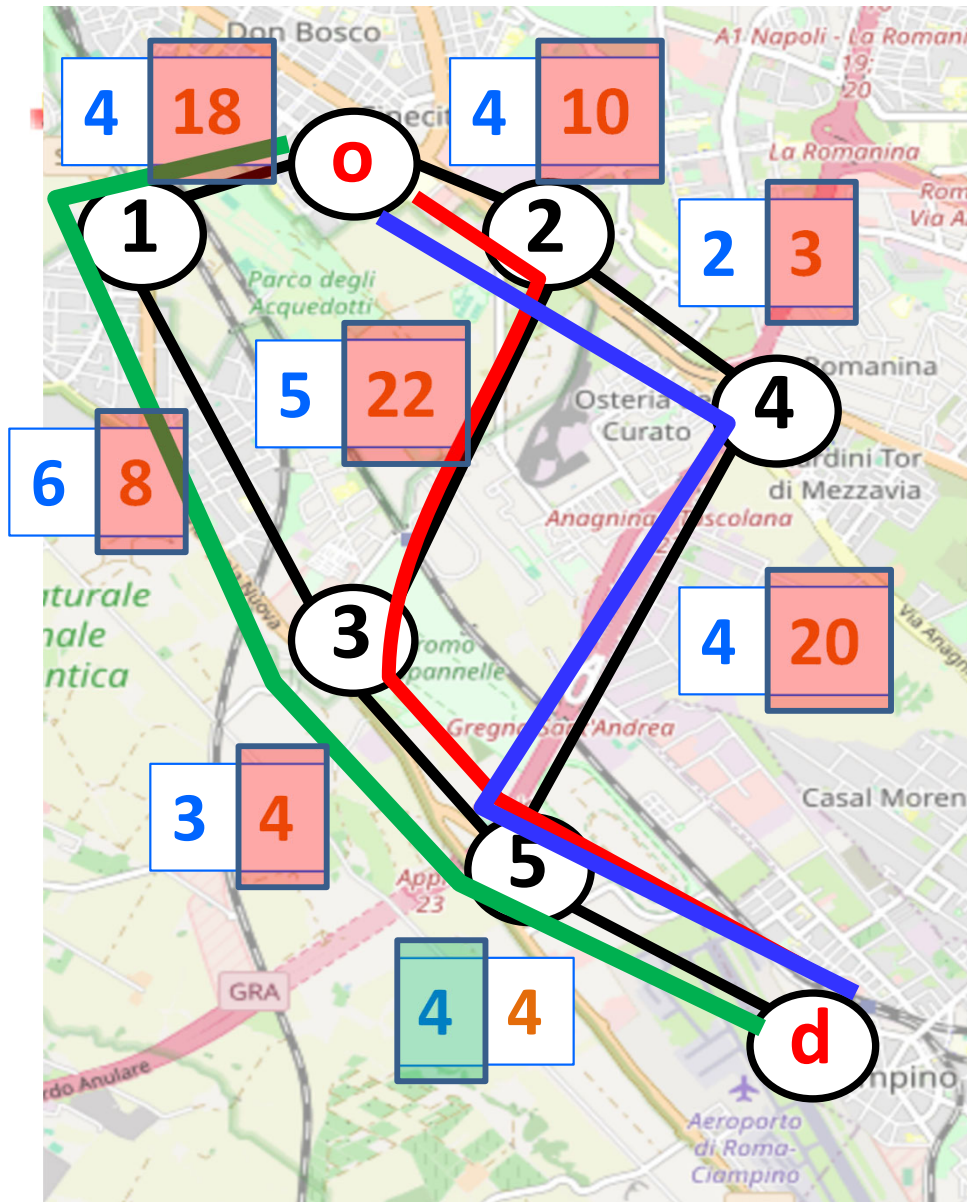
(2) via Appia Nuova $18+8+3+4 = 33$

(3) via delle Capannelle $10+22+3+4 = 39$

Per $\Gamma = 2$ (2 archi a tempo max)

- il percorso (2) diventa il migliore

Impatto del budget di incertezza ($\Gamma = 3$)



✚ Se fissiamo $\Gamma = 3$ permettiamo a 3 archi di avere tempo massimo e gli archi colpiti da questo incremento dipendono dal percorso.

✚ Tempi di percorrenza =

Tempo best + variazione $\Gamma = 3$

(1) via GRA $10+3+20+4 = 37$

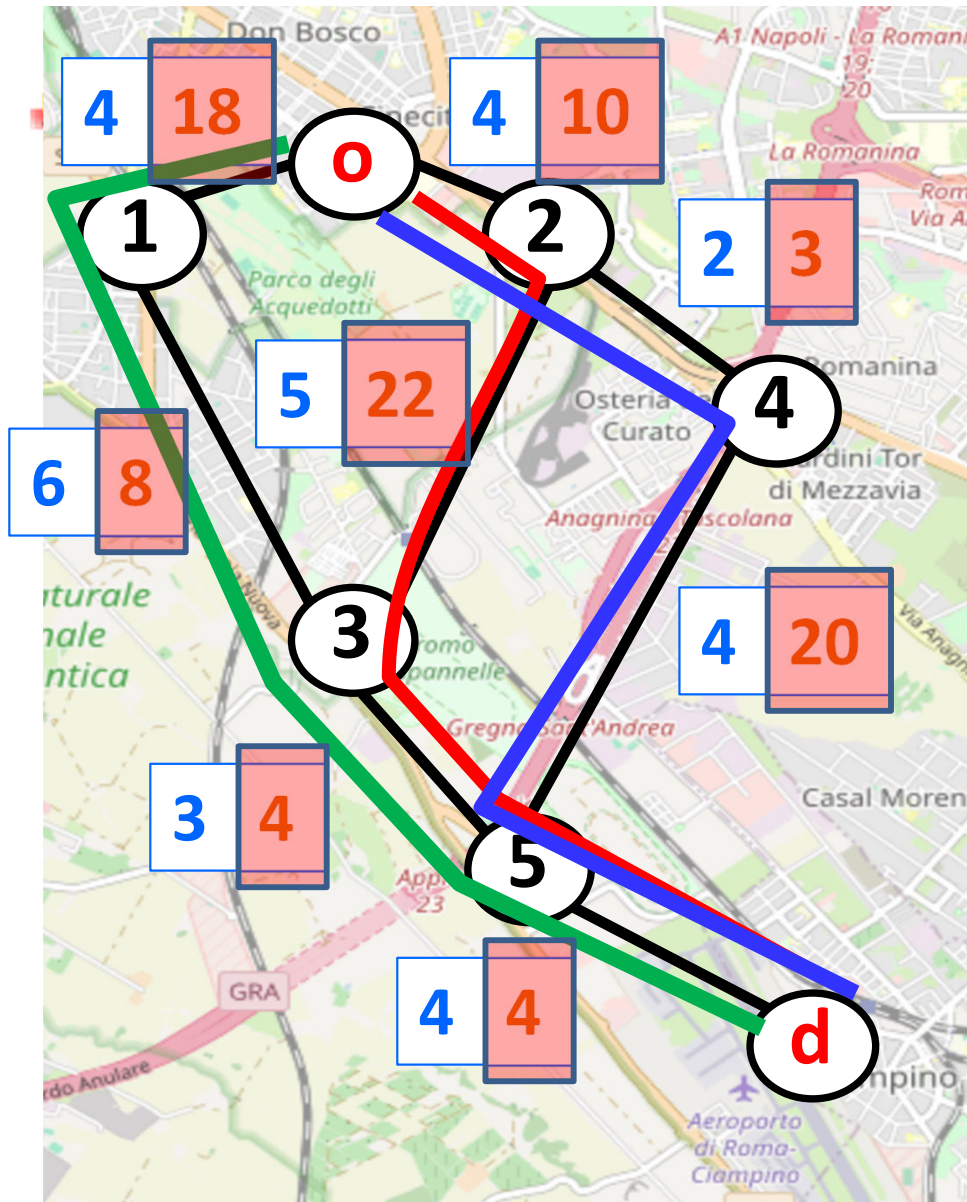
(2) via Appia Nuova $18+8+4+4 = 34$

(3) via delle Capannelle $10+22+4+4 = 40$

Per $\Gamma = 3$ (3 archi a tempo max)

- il percorso (2) resta il migliore

Impatto del budget di incertezza ($\Gamma = 4$)



✚ Se fissiamo $\Gamma = 4$ permettiamo a 4 archi di avere tempo massimo e gli archi colpiti da questo incremento dipendono dal percorso.

✚ DATO CHE L'ARCO (5,d) HA TEMPO COSTANTE = 4 minuti (no incertezza), la situazione non cambia ed i tempi sono:

(1) via GRA	$10+3+20+4 = 37$
(2) via Appia Nuova	$18+8+4+4 = 34$
(3) via delle Capannelle	$10+22+4+4 = 40$

Per $\Gamma = 4$ (4 archi a tempo max)

- il percorso (2) resta il migliore

Verso la Robust Optimization

Per migliorare la nostra rappresentazione dell'incertezza dei dati, stiamo muovendoci verso la:

ROBUST OPTIMIZATION (Ottimizzazione Matematica Robusta)

Discutiamone i principi in modo matematicamente più formale.

It's a stochastic world

La maggior parte dei problemi di ottimizzazione del mondo reale **coinvolgono data incerti**

FINANCE

Valore azioni



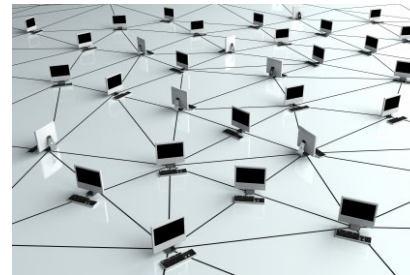
AIRCRAFT SCHEDULING

Ritardi di volo



TELECOM NETWORK DESIGN

Volumi di traffico



SURGERY SCHEDULING

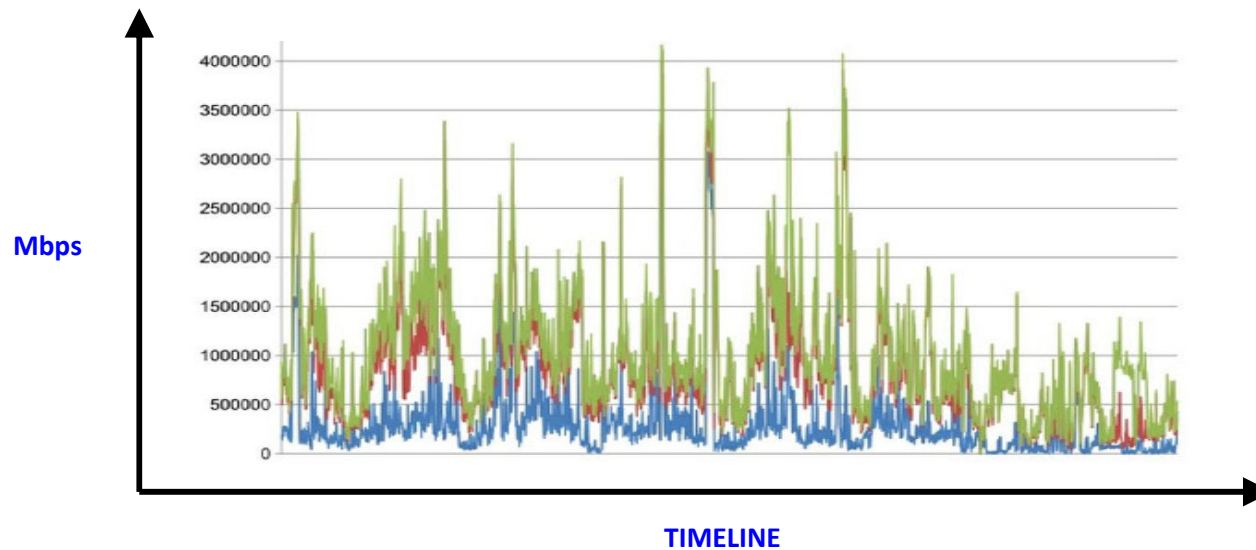
Numero di operazioni attese



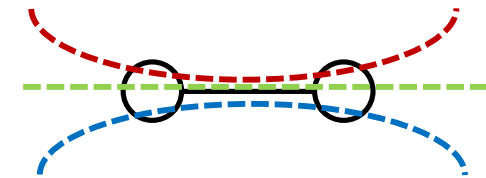
La **questione (spinosa) della presenza di dati incerti in problemi di ottimizzazione** era stata **già identificata da George Dantzig**, uno dei padri fondatori dell'ottimizzazione matematica moderna nel lavoro pionieristico

"Linear Programming under Uncertainty" (Management Science, 1955)

Un esempio di dati incerti: traffico su reti TLC



Fluttuazioni di traffico nella USA Abilene Network
3 connessioni su un link
(periodo di osservazione: una settimana)



In ognuna delle 3 connessioni, il traffico fluttua considerevolmente



La fluttuazione complessiva è molto alta



Tipica soluzione pratica per affrontare l'incertezza di traffico: sovra-dimensionare la capacità di rete sovrastimando (altamente) la domanda

?

**POSSIAMO DEFINIRE DELLE SOLUZIONI ROBUSTE MIGLIORI
TRAMITE L'OTTIMIZZAZIONE MATEMATICA**

?

Disambiguazione: cosa intendiamo per Robustezza?

- ✚ L'aggettivo **ROBUSTO**, è ampiamente utilizzato in Computer Science e Engineering con significati molto diversi.

ESEMPI:

- Robust Telecom network Design (= robustezza verso guasti di connessione)
- Robust Road Routing (= robustezza verso decisori non-razionali)

✚ IN QUESTA PRESENTAZIONE:

**SOLUZIONE
ROBUSTA** = soluzione protetta da
deviazioni nei valori di dati di input

Disambiguazione: cosa intendiamo per Robustezza?

- ✚ L'aggettivo **ROBUSTO**, è ampiamente utilizzato in Computer Science e Engineering con significati molto diversi.

ESEMPI:

- Robust Telecom network Design (= robustezza verso guasti di connessione)
- Robust Road Routing (= robustezza verso decisori non-razionali)

✚ IN QUESTA PRESENTAZIONE:

SOLUZIONE ROBUSTA = soluzione protetta da deviazioni nei valori di dati di input

ALCUNE OSSERVAZIONI CRUCIALI



1. La questione di come modellare l'incertezza di dati e l'incertezza è ancora aperta
2. Nel corso degli anni, **multi modelli sono stati proposti**
3. In generale, non esiste un modello che domina gli altri da un punto di vista teorico



Disambiguazione: cosa intendiamo per Robustezza?

- ✚ L'aggettivo **ROBUSTO**, è ampiamente utilizzato in Computer Science e Engineering con significati molto diversi.

ESEMPI:

- Robust Telecom network Design (= robustezza verso guasti di connessione)
- Robust Road Routing (= robustezza verso decisori non-razionali)

✚ IN QUESTA PRESENTAZIONE:

SOLUZIONE ROBUSTA = soluzione protetta da deviazioni nei valori di dati di input

ALCUNE OSSERVAZIONI CRUCIALI



1. La questione di come modellare l'incertezza di dati e l'incertezza è ancora aperta
2. Nel corso degli anni, **molti modelli sono stati proposti**
3. In generale, non esiste un modello che domina gli altri da un punto di vista teorico



D'ALTRA PARTE

Si concorda sul fatto che **i professionals preferiscano alcuni modelli e algoritmi di ottimizzazione rispetto ad altri**. In particolare, quelli che comprendono più facilmente e al cui tuning possono partecipare più facilmente

La soluzione non era robusta...



✚ Un esempio semplicissimo può aiutare a cogliere l'impatto di deviazioni inaspettate nei dati di input: Supponiamo di avere calcolato una soluzione ottima $x=1, y=1$ per un problema con il vincolo:

$$100x + 200y \leq 300$$

La soluzione non era robusta...



✚ Un esempio semplicissimo può aiutare a cogliere l'impatto di deviazioni inaspettate nei dati di input: Supponiamo di avere calcolato una soluzione ottima $x=1, y=1$ per un problema con il vincolo:

$$100x + 200y \leq 300$$

Il vincolo è soddisfatto,

MA abbiamo **trascurato il fatto che il coefficiente 100 che moltiplica la variabile x può essere soggetto ad una deviazione del 10%.**

Quindi potrebbe avvenire che:

$$(100 + 10) + 200 > 300$$

**SOLUZIONE OTTIMA
EFFETTIVAMENTE INAMMISSIBILE!
(VINCOLO VIOLATO)**

✚ Che sarebbe successo se si fosse trattato di un'applicazione critica?
(e.g., individuazione di inquinanti in acqua potabile)

Dati incerti nell'ottimizzazione matematica

**OTTIMIZZAZIONE MATEMATICA
CLASSICA**



**IL VALORE DI TUTTI I DATI DI INPUT
È **CONOSCIUTO PRECISAMENTE**
QUANDO SI RISOLVE IL PROBLEMA**

Dati incerti nell'ottimizzazione matematica

OTTIMIZZAZIONE MATEMATICA
CLASSICA



IL VALORE DI TUTTI I DATI DI INPUT
È **CONOSCIUTO PRECISAMENTE**
QUANDO SI RISOLVE IL PROBLEMA

IPOSTESI RAGIONEVOLE PER TUTTI I PROBLEMI DI OTTIMIZZAZIONE? **NO!**

In molti problemi (del mondo reale), **i dati sono incerti**, i.e. **il loro valore non è precisamente conosciuto quando il problema deve essere risolto**

Dati incerti nell'ottimizzazione matematica

**OTTIMIZZAZIONE MATEMATICA
CLASSICA**



**IL VALORE DI TUTTI I DATI DI INPUT
È CONOSCIUTO PRECISAMENTE
QUANDO SI RISOLVE IL PROBLEMA**

IPOSTESI RAGIONEVOLE PER TUTTI I PROBLEMI DI OTTIMIZZAZIONE? NO!

In molti problemi (del mondo reale), **i dati sono incerti**, i.e. il loro valore non è **precisamente conosciuto quando il problema deve essere risolto**

ALCUNE CAUSE:

- ✚ errori di misurazione
- ✚ stime di valore da serie di dati storiche
- ✚ rappresentazione numerica finita dei computer

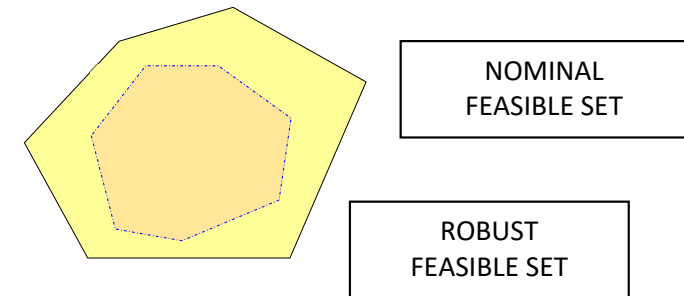
Tralasciare o ignorare la presenza di dati incerti può condurre a brutte sorprese:

- 1. soluzioni ottime, possono effettivamente rivelarsi di cattiva qualità**
- 2. soluzioni ammissibili, possono rivelarsi inammissibili (e quindi inutilizzabili)**

Robust mathematical optimization

La presenza di dati incerti viene gestita tramite l'**introduzione di vincoli aggiuntivi** che restringono l'insieme di soluzioni ammissibili

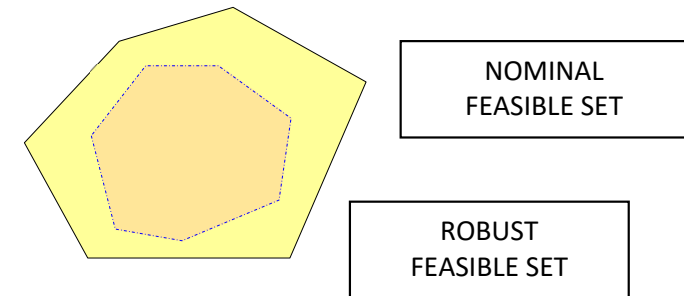
Sono ammissibili solo soluzioni robuste (i.e., protette contro deviazioni di valore nell'input)



Robust mathematical optimization

La presenza di dati incerti viene gestita tramite l'**introduzione di vincoli aggiuntivi** che restringono l'insieme di soluzioni ammissibili

Sono ammissibili solo soluzioni robuste (i.e., protette contro deviazioni di valore nell'input)



NOMINAL PROBLEM

$$\begin{aligned} \max \quad & c'x \\ \text{s.t.} \quad & Ax \leq b \\ & x \geq 0 \end{aligned}$$



COEFFICIENTI
INCERTI!!!



ROBUST COUNTERPART

$$a_{ij} = \bar{a}_{ij} + \delta_{ij}$$



VALORE
EFFETTIVO



VALORE DI
RIFERIMENTO
(CONOSCIUTO)



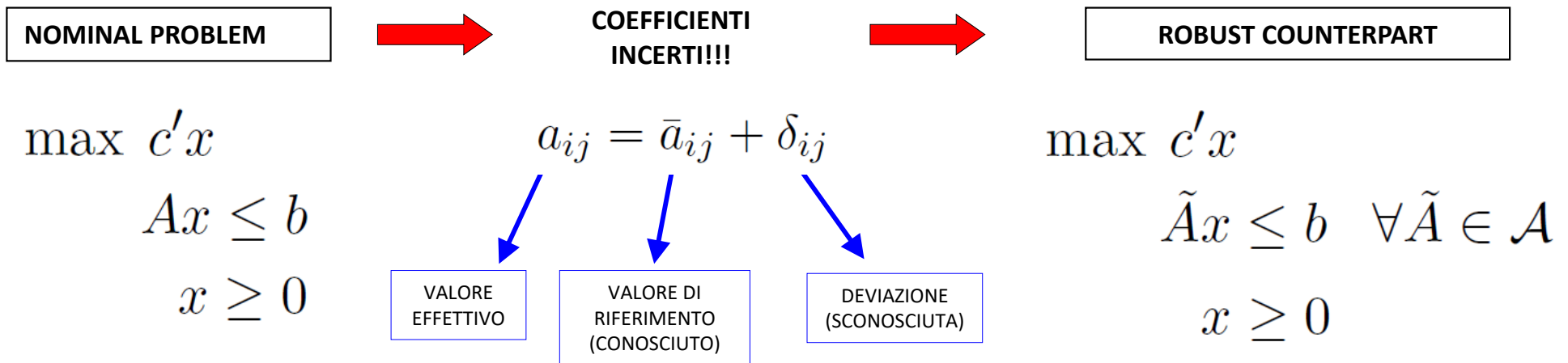
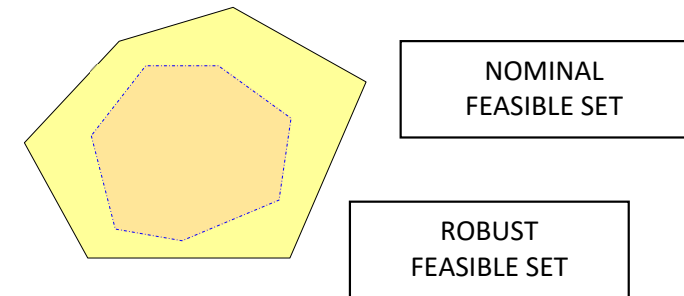
DEVIAZIONE
(SCONOSCIUTA)

$$\begin{aligned} \max \quad & c'x \\ \text{s.t.} \quad & \tilde{A}x \leq b \quad \forall \tilde{A} \in \mathcal{A} \\ & x \geq 0 \end{aligned}$$

Robust mathematical optimization

La presenza di dati incerti viene gestita tramite l'**introduzione di vincoli aggiuntivi** che restringono l'insieme di soluzioni ammissibili

Sono ammissibili solo soluzioni robuste (i.e., protette contro deviazioni di valore nell'input)



- ✚ $\mathcal{A}$ è un sottoinsieme di tutte le matrici che rappresentano tutte le possibili combinazioni di dati incerti per il problema
- ✚ considerare un sottoinsieme di matrici più grande, **aumenta la protezione, ma restringe anche l'insieme di soluzioni ammissibili robuste**
- ✚ la **protezione** nei confronti di dati incerti comporta un **deterioramento del valore della soluzione ottima (Price of Robustness)**

Alcune osservazioni (informali) importanti

Esistono deviazioni di valore "buone" e "cattive".

Per esempio, una diminuzione del tempo di percorrenza di una strada per diminuzione del traffico può essere considerata una deviazione buona, mentre un incremento del tempo è una deviazione cattiva.

Siamo principalmente interessati a caratterizzare le deviazioni cattive.

Le deviazioni cattive sono infatti quelle che ci colpiscono:

- **rendono soluzioni ottimali di scarsa qualità**
- **compromettono l'ammissibilità di soluzioni (violazione dei vincoli)**

Delle deviazioni cattive di piccola entità possono essere tollerate

(le soluzioni non diventano immediatamente mediocri o inammissibili per lievi deviazioni)

In generale, come abbiamo visto per l'esempio introduttivo di spostamento casa-aeroporto, la protezione worst case (i.e., protezione contro tutte le possibili deviazioni) è troppo conservativa

Robust Optimization: il modello Bertsimas-Sim (1)

È generalmente accettato nella comunità scientifica che un notevole balzo in avanti nella gestione dei dati incerti è stato fatto con l'introduzione del modello Gamma-Robustness di D. Bertsimas e M. Sim (MIT - Singapore Alliance)

Si tratta del modello che **abbiamo implicitamente impiegato** per raffinare la rappresentazione dell'incertezza nell'esempio di spostamento casa-aeroporto.

Procederò ora ad illustrarne le caratteristiche principali.

Robust Optimization: il modello Bertsimas-Sim (2)

Partiamo da un semplice modello di ottimizzazione lineare:

$$\begin{aligned} \max \quad & \sum_{j \in J} c_j x_j \\ \text{s.t.} \quad & \sum_{j \in J} a_{ij} x_j \leq b_i \quad i \in I = \{1, \dots, m\} \\ & x_j \geq 0 \quad j \in J = \{1, \dots, n\} \end{aligned}$$

e, senza perdita di generalità, ipotizziamo che **gli unici coefficienti incerti sono i coefficienti di input a_{ij}**

Quindi ipotizziamo che:

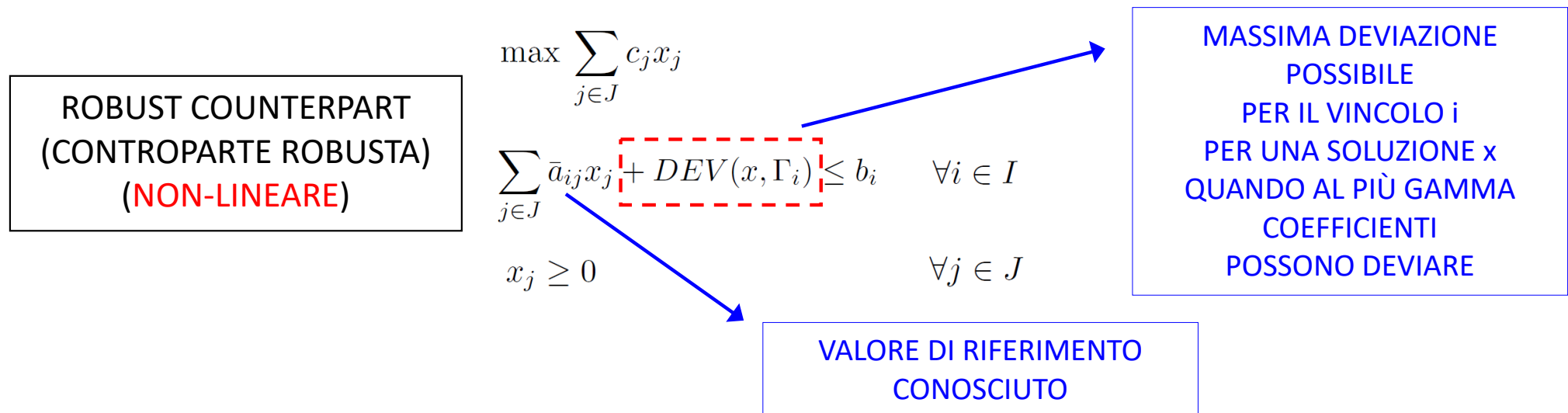
- per ogni coefficiente incerto a_{ij} conosciamo
 - un **valore di riferimento** $\bar{a}_{ij}$
 - un **valore di deviazione massima** $d_{ij}^{\max}$
- il valore (incerto) di un coefficiente a_{ij} cade all'interno **dell'intervallo di incertezza**: $[a_{ij} - d_{ij}^{\max}, a_{ij} + d_{ij}^{\max}]$
- per ogni vincolo i -esimo del problema, **desideriamo proteggerci da Γ_i coefficienti che deviano dal proprio valore di riferimento $\bar{a}_{ij}$**

Robust Optimization; il modello Bertsimas-Sim (3)

In base alle precedenti ipotesi, il **modello NON-ROBUSTO lineare** iniziale:

$$\begin{aligned} \max \quad & \sum_{j \in J} c_j x_j \\ \text{s.t.} \quad & \sum_{j \in J} a_{ij} x_j \leq b_i \quad i \in I = \{1, \dots, m\} \\ & x_j \geq 0 \quad j \in J = \{1, \dots, n\} \end{aligned}$$

può essere riscritto come il seguente **modello ROBUSTO non-lineare**



che però ha il "difetto" di essere diventato **non-lineare**

La magia duale del modello Bertsimas-Sim (1)

$$DEV(x, \Gamma_i) =$$

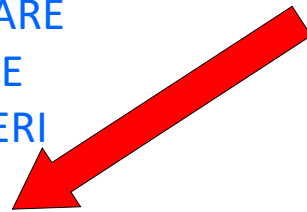
$$\max \sum_{j \in J} d_{ij}^{\max} x_j y_{ij}$$

$$\sum_{j \in J} y_{ij} \leq \Gamma_i$$

$$y_{ij} \in \{0, 1\} \quad j \in J$$

PROBLEMA DI
OTTIMIZZAZIONE
CHE ESPRIME LA RICERCA
DELLA MASSIMA DEVIATIONE
DEI COEFFICIENTI
CONSENTITA DA AL PIÙ
GAMMA_i DEVIATIONI

IL RILASSAMENTO LINEARE
AMMETTE SOLUZIONE
OTTIMA A VALORI INTERI



$$\max \sum_{j \in J} d_{ij}^{\max} x_j y_{ij}$$

$$\sum_{j \in J} y_{ij} \leq \Gamma_i$$

$$0 \leq y_{ij} \leq 1 \quad j \in J$$

DEFINIZIONE
DEL PROBLEMA
DUALE



QUESTO **PROBLEMA DUALE** PUÒ ESSERE
REINSERITO NEL PROBLEMA ORIGINALE
ELIMINANDO LA NON-LINEARITÀ

$$\min \Gamma_i w_i + \sum_{j \in J} z_{ij}$$

$$w_i + z_{ij} \geq d_{ij}^{\max} x_j \quad j \in J$$

$$w_i \geq 0$$

$$z_{ij} \geq 0 \quad j \in J$$

La magia duale del modello Bertsimas-Sim (2)

$$\begin{aligned} \max \quad & \sum_{j \in J} c_j x_j \\ & \sum_{j \in J} \bar{a}_{ij} x_j + \Gamma_i w_i + \sum_{j \in J} z_{ij} \leq b_i & i \in I \\ & w_i + z_{ij} \geq d_{ij}^{\max} x_j & i \in I, j \in J \\ & w_i \geq 0 & i \in I \\ & z_{ij} \geq 0 & i \in I, j \in J \\ & x_j \geq 0 & j \in J \end{aligned}$$

Una **controparte robusta LINEARE** che richiede :

- $|I| + |I| \times |J|$ variabili aggiuntive
- $|I| \times |J|$ vincoli aggiuntivi

Notiamo che le **difficoltà** associate all'aumento di dimensione e complessità del problema sono **più che compensate dall'eliminazione dei termini non-lineari**

La magia duale del modello Bertsimas-Sim (3)

- ✚ Il **parametro Gamma Γ** ci consente di controllare con facilità la robustezza del modello di ottimizzazione robusta
- ✚ Per **$\Gamma = 0$** non imponiamo alcuna protezione
(ci proteggiamo rispetto a 0 deviazioni worst di valore, effettivamente siamo nel caso dell'ottimizzazione classica)
PRICE OF ROBUSTNESS = 0
- ✚ Per **$\Gamma = \text{num. coefficienti incerti}$** imponiamo la protezione completa
(ci proteggiamo rispetto a tutte le deviazioni worst posizioni)
PRICE OF ROBUSTNESS = MASSIMO VALORE
- ✚ Per **$0 < \Gamma < \text{num. coefficienti incerti}$** abbiamo casi di protezione intermedi
Più aumentiamo Γ , più ci proteggiamo, ma incrementa anche il **Price of Robustness**

KISS Bertsimas and Sim

- ✚ Teoria accessibile e matematicamente elegante per gestire l'incertezza dei dati
- ✚ Punto di partenza per innumerevoli altri studi
- ✚ Nonostante altri sviluppi teorici, a distanza di 20 anni resta ancora un riferimento cruciale per teoria ed applicazioni di Ottimizzazione Robusta



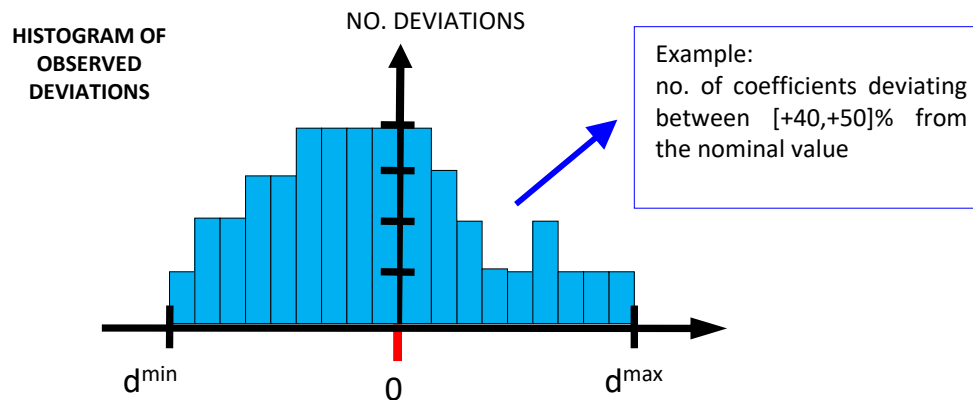
Keep It Simple
and
Straightforward!

modello di incertezza comprensibile ed efficace

controllo semplice della robustezza
(parametro Gamma)

Un passo ulteriore: Multiband Robust Optimization (1)

- In problemi reali, dati storici riguardanti i dati incerti sono tipicamente disponibili
- Questi dati storici possono essere impiegati per definire **ISTROGRAMMI** rappresentanti la distribuzione delle deviazioni

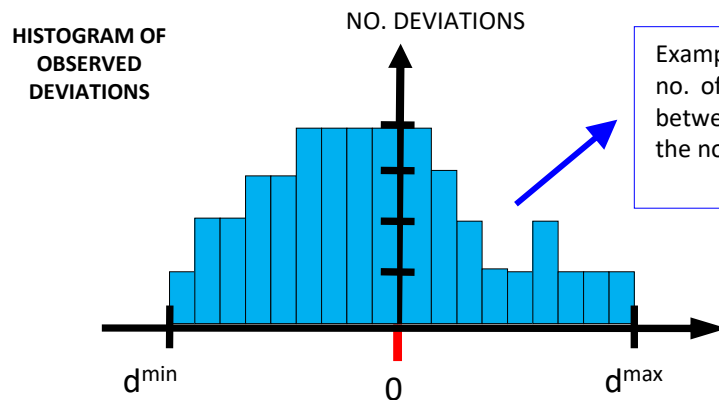


Un passo ulteriore: Multiband Robust Optimization (1)

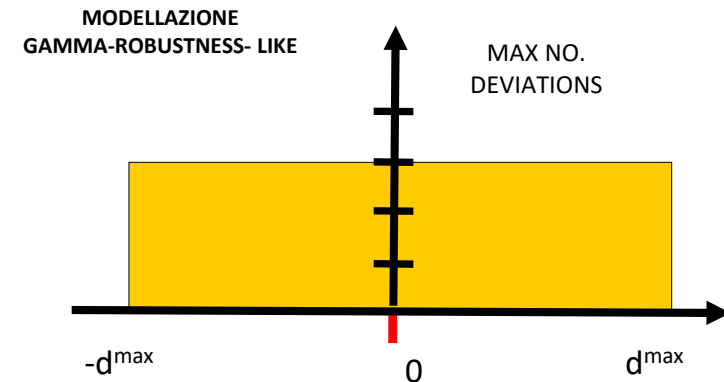
- In problemi reali, dati storici riguardanti i dati incerti sono tipicamente disponibili
- Questi dati storici possono essere impiegati per definire **ISTROGRAMMI** rappresentanti la distribuzione delle deviazioni

**SIAMO IN GRADO DI SFRUTTARE EFFICACEMENTE QUESTE
INFORMAZIONI CON IL MODELLO BERTSIMAS-SIM?**

(basato sull'ipotesi di avere distribuzioni di probabilità simmetriche su intervalli di deviazione simmetrici)



Example:
no. of coefficients deviating
between $[+40, +50]\%$ from
the nominal value

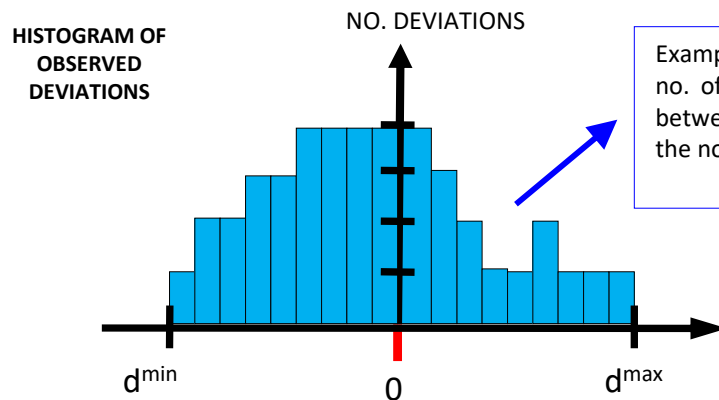


Un passo ulteriore: Multiband Robust Optimization (1)

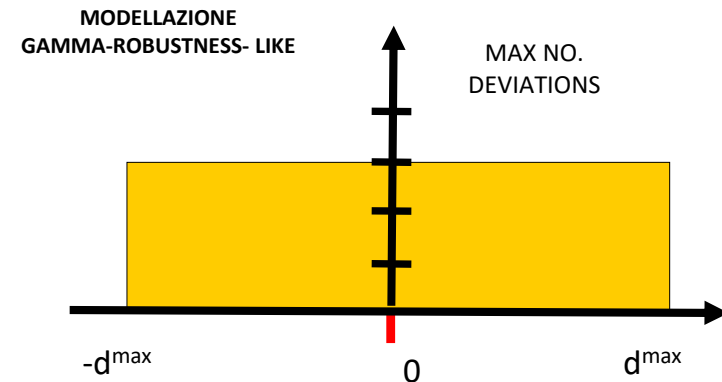
- In problemi reali, **dati storici riguardanti i dati incerti** sono tipicamente disponibili
- Questi dati storici possono essere impiegati per definire **ISTROGRAMMI** rappresentanti la distribuzione delle deviazioni

SIAMO IN GRADO DI SFRUTTARE EFFICACEMENTE QUESTE INFORMAZIONI CON IL MODELLO BERTSIMAS-SIM?

(basato sull'ipotesi di avere distribuzioni di probabilità simmetriche su intervalli di deviazione simmetrici)



Example:
no. of coefficients deviating
between $[+40, +50]\%$ from
the nominal value



- Il comportamento delle deviazioni all'**interno** dell'intervallo di deviazioni è **completamente trascurato** e si **focalizza l'attenzione sulle deviazioni estreme**
- In base alla mia esperienza diretta, i **professionisti desidererebbero una rappresentazione più raffinata di questo tipo di incertezza**

Un passo ulteriore: Multiband Robust Optimization (1)

?

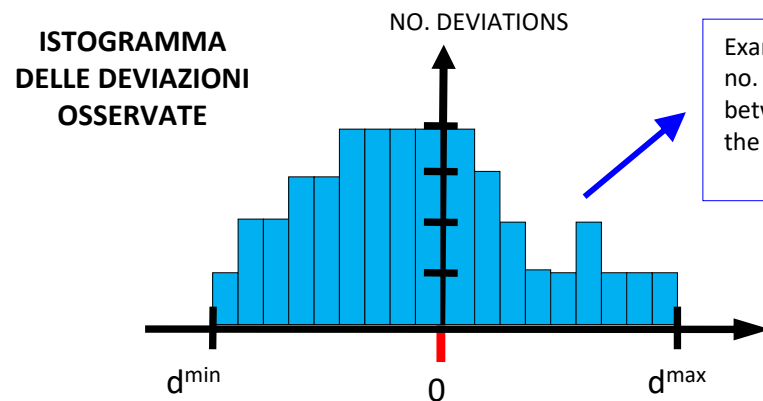
COSA POSSIAMO FARE PER AUMENTARE LA NOSTRA
CAPACITÀ DI MODELLAZIONE

?

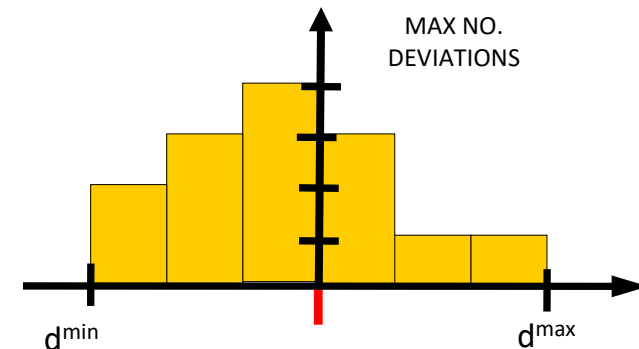
💡

ADOTTARE UN INSIEME MULTI-BAND

💡



Example:
no. of coefficients deviating
between $[+40, +50]\%$ from
the nominal value



insieme ad una mia collega tedesca (Christina Büsing, TU Berlin) **abbiamo studiato le conseguenze teoriche di questa semplice ma efficace idea**, ideando quella che abbiamo chiamato **Multiband Robust Optimization**, una strongly-data-driven generalizzazione del modello Bertsimas-Sim

Multiband Robustness - risultato teorico centrale

PROPOSITION 1 (Büsing & D'Andreagiovanni 2012)

The polytope associated with (DEV01) is integral.

✚ Proof based on showing that the coefficient matrix of (DEV01) is totally unimodular

THEOREM 1 (Büsing & D'Andreagiovanni)

The Robust Counterpart of (MILP) under multi-band uncertainty is equivalent to:

$$\begin{aligned}
 & \max \sum_{j \in J} c_j x_j && (RLP) \\
 & \sum_{j \in J} \bar{a}_{ij} x_j - \sum_{k \in K} l_k v_i^k + \sum_{k \in K} u_k w_i^k + \sum_{j \in J} z_i^j \leq b_i && i \in I \\
 & -v_i^k + w_i^k + z_i^j \geq d_{ij}^k x_j && i \in I, j \in J, k \in K \\
 & v_i^k, w_i^k \geq 0 && i \in I, k \in K \\
 & z_i^j \geq 0 && i \in I, j \in J \\
 & x_j \geq 0 && j \in J \\
 & x_j \in \mathbb{Z}_+ && j \in J_{\mathbb{Z}} \subseteq J
 \end{aligned}$$

✚ Proof based on exploiting the integrality of (DEV01) and strong duality

✚ If the original problem is purely linear, then also the counterpart is purely linear

Multiband Robustness - achievements

- ✚ **ispirato da reali necessità del partner industriale Nokia Siemens Networks** per la modellazione di dati incerti in reti fisse di telecomunicazione nazionale nell'ambito del **progetto di ricerca federale tedesco ROBUKOM**
- ✚ ricerche ulteriori finanziate nell'ambito di iniziative federali di eccellenza nella ricerca (**progetto ROUAN, Einstein Center for Mathematics, Berlin**)
- ✚ articolo scientifico basato su Multiband Robustness premiato in conferenza internazionale maggiore
- ✚ ripreso e adottato da molti altri ricercatori internazionali
- ✚ **adattato e proposto nell'ambito di un progetto di consulenza** per gestione di offerte di acquisto e vendita di energia nel **mercato elettrico** per una delle maggiori utility company a livello mondiale

PART II

Robust

Artificial Intelligence

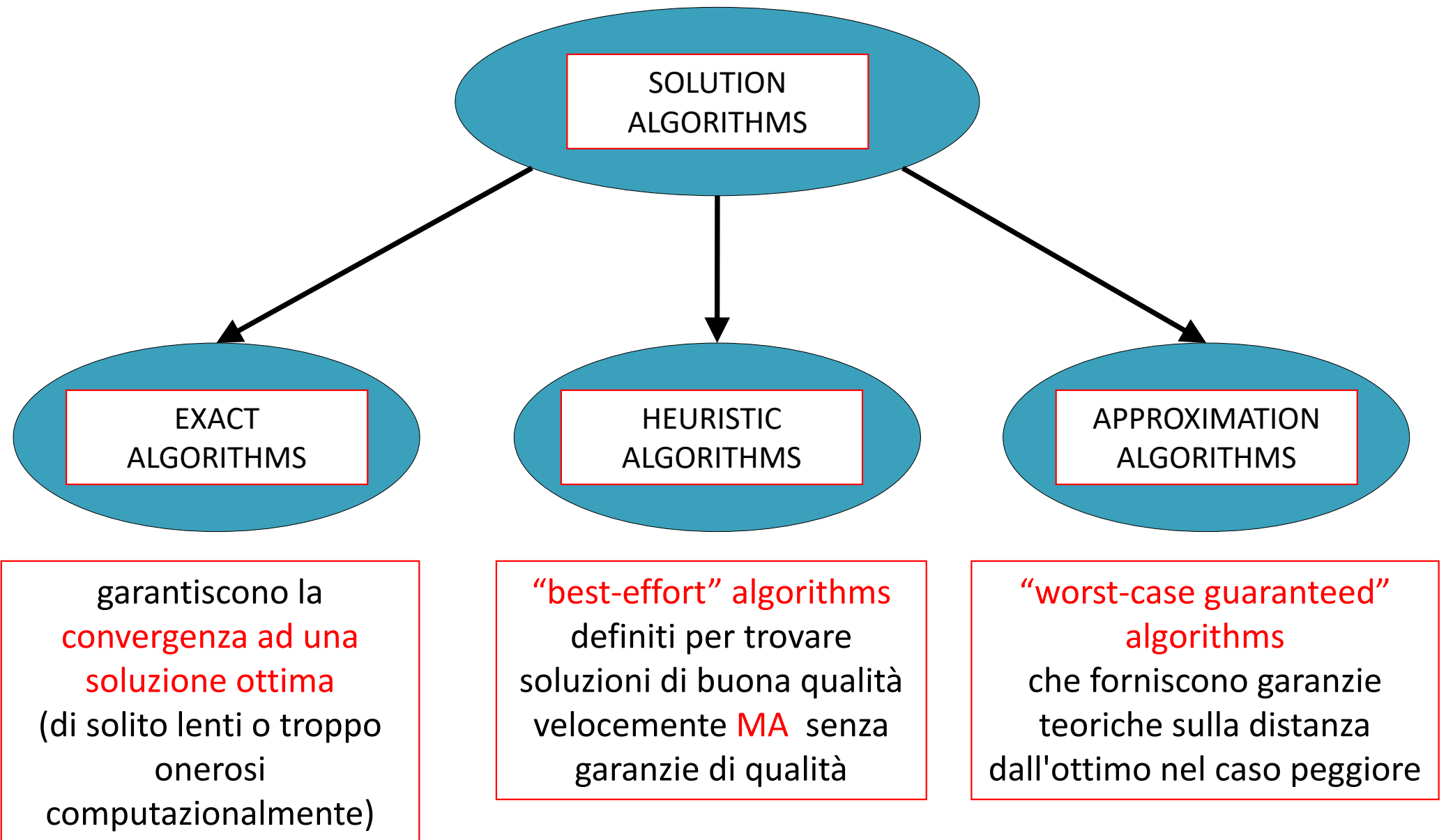
Alcune definizioni (in inglese)

- ✚ **Algorithm:** a well-defined computational procedure that requires a set of values **(the input)** and produces a set of values **(the output)**
- ✚ The algorithm can be seen as **the sequence of computational steps that transform the input into the output**



- ✚ **Another point of view:**
See the algorithm as a **tool for solving a specific computational problem**
(e.g., il problema di ordinare un insieme di elementi dal più piccolo al più grande - SORTING PROBLEM)

Optimization Algorithms



Meta- o Math- heuristics

- ✚ Un **algoritmo euristico** è tipicamente definito come un algoritmo per la risoluzione efficiente di problemi di ottimizzazione NP-Hard (traducibile come "difficili dal punto di vista teorico") che **sacrificano garanzie teoriche di qualità per velocità di esecuzione**.
- ✚ Lo scopo ultimo di un'euristica è quello di trovare una soluzione di alta qualità velocemente. **MA, per definizione le euristiche non sono in grado di valutare la qualità di una soluzione** (derivare un c.d. optimality gap)
- ✚ **Alcuni esempi notevoli di euristiche:**
 - greedy algorithms
 - local search

Meta- o Math- heuristics

- ✚ Un **algoritmo euristico** è tipicamente definito come un algoritmo per la risoluzione efficiente di problemi di ottimizzazione NP-Hard (traducibile come "difficili dal punto di vista teorico") che **sacrificano garanzie teoriche di qualità per velocità di esecuzione**.
 - ✚ Lo scopo ultimo di un'euristica è quello di trovare una soluzione di alta qualità velocemente. **MA, per definizione le euristiche non sono in grado di valutare la qualità di una soluzione** (derivare un c.d. optimality gap)
 - ✚ **Alcuni esempi notevoli di euristiche:**
 - greedy algorithms
 - local search
 - ✚ **Metaheuristics** rappresentano delle euristiche “più raffinate”, di solito ispirate da processi naturali, come:
 - ant colony optimization (ispirato dal processo di ricerca del cibo delle formiche)
 - simulated annealing (ispirato dal processo di raffreddamento di materiali in metallurgia)
 - genetic algorithms (ispirati da processi di evoluzione di popolazioni)
- MATHEURISTICS** sono una forma recente di metaheuristic che **integrano metodi di ottimizzazione esatti**
- LEARNHEURISTICS** sono una forma ancora più recente di metaeuristiche che **integrano tecniche di Intelligenza Artificiale (in primis, Machine Learning)**

Artificial Intelligence - alcune definizioni

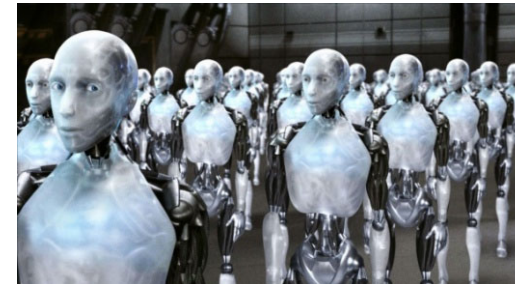
- **Artificial Intelligence (AI)**

"involves machines that can perform tasks that are characteristic of human intelligence"

John McCarthy ,1956

ESEMPLI:

- comprendere il linguaggio parlato
- identificare oggetti visivamente
- imparare a risolvere problemi di ottimizzazione

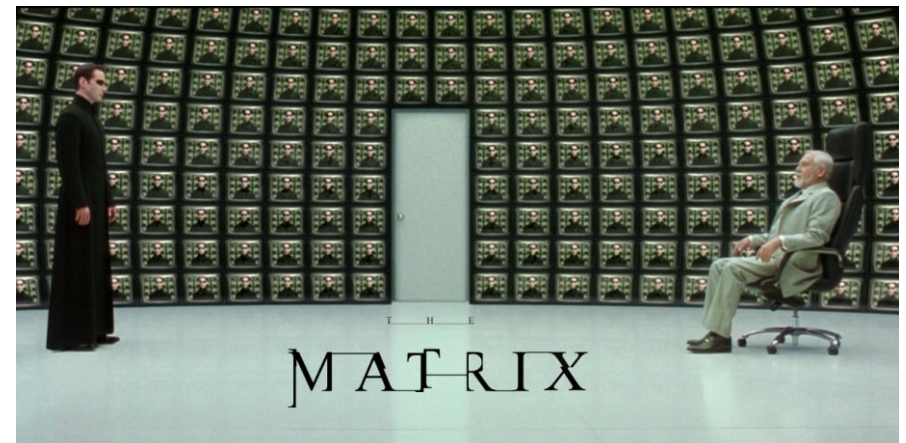


- **AI is different from AGI**

AGI = Artificial General Intelligence

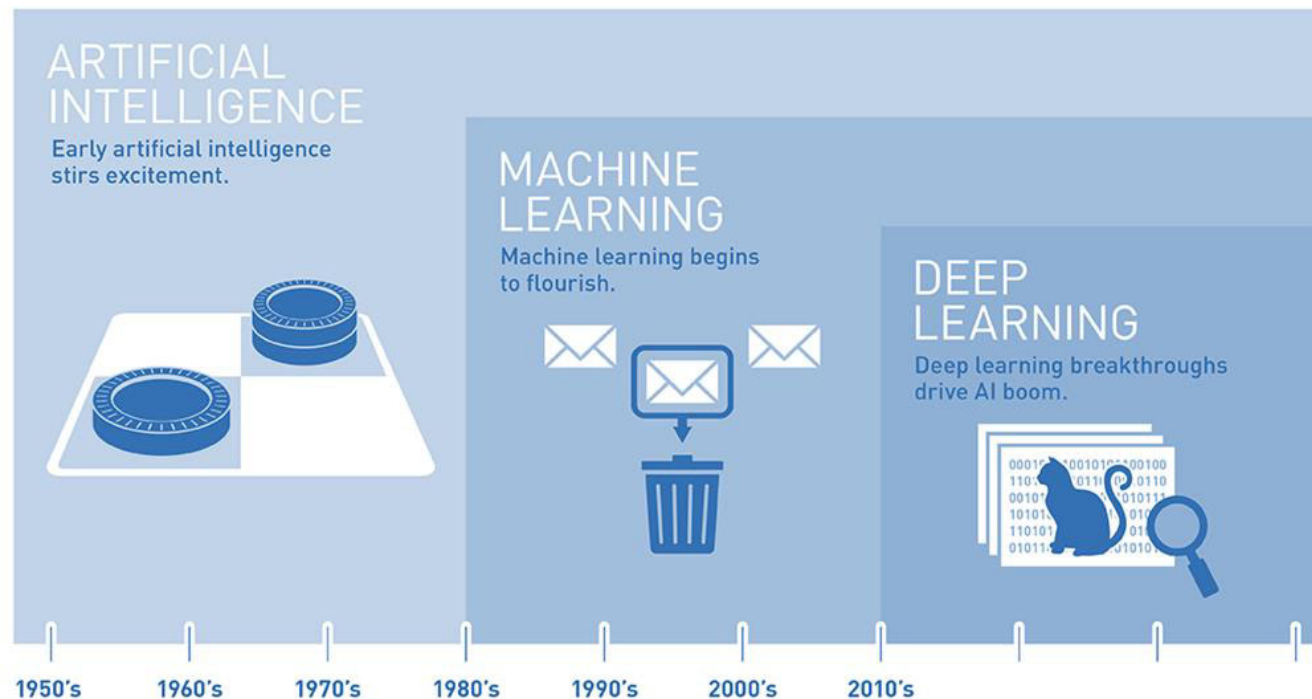
"a machine that can successfully perform any intellectual task that a human being is able to conduct"

Shall we see AGI?



Artificial Intelligence - una panoramica

- Tre buzzwords che probabilmente udite spesso:
 - Artificial Intelligence
 - Machine Learning
 - Deep Learning
- Machine Learning = **subfield dell'AI** (esiste cospicua AI di tipo non-learning)
- Deep Learning = **subfield del Machine Learning** - esistono molti non-deep learning algorithms



Artificial Intelligence - esempi - recommendation systems

- Tra tutti i risultati di ricerca, predire quelli che sono **più rilevanti per una persona** (e.g., Google)
- Raccomandare **prodotti simili** (e.g., Netflix, Amazon)

The screenshot shows an Amazon product page for a Logitech Desktop MK120 Durable, Comfortable, USB Mouse and keyboard Combo. The main product image is a black keyboard and mouse. To the right, there's a section titled 'Available from these sellers.' with details: Connectivity Technology: USB, Brand: Logitech, Keyboard Description: QWERTY. Below this is 'About this item' with a list of features: Comfortable low-profile keys, spill-resistant design, high-definition optical mouse, easy-to-read keys, and simple wired USB connection. A 'New & Used (9) from \$26.88' section is also present. A 'This product is available as Renewed.' section shows a renewed version for \$23.99 with a 4.5-star rating (10 reviews) and the Amazon Renewed Guarantee. On the right, a sponsored section shows a 'Wireless Keyboard and Mouse Combo, RATEL 2.4GHz Full...' for \$24.99 with a 15% off coupon. At the bottom, a 'Products related to this item' section displays a carousel of similar products, including Havit Wired Gaming, Wireless Keyboard and Mouse, Vaydeer USB3.0 Wireless, Ri-ck and Mo-RTY Mouse, Wireless Keyboard and Mouse, Manhattan Black Wired, and Wireless Keyboard and Mouse. The page is marked 'Page 1 of 108'.

Available from these sellers.

Connectivity Technology USB
Brand Logitech
Keyboard Description QWERTY

About this item

- Comfortable low-profile keys and standard keyboard layout with full-size F-keys and number pad
- Sleek and sturdy spill-resistant design* (*Tested under limited conditions—maximum of 60 ml liquid spillage. Do not immerse keyboard in liquid.)
- High-definition optical mouse for smooth, accurate cursor control
- Easy-to-read keys with bold, bright white characters
- Simple wired USB connection—works with Windows-based PCs and Linux

New & Used (9) from \$26.88

This product is available as Renewed.

Logitech Desktop MK120 Durable, Comfortable, USB Mouse and keyboard Combo (Renew...)
\$23.99
★★★★☆ (10)
Works and looks like new and backed by the Amazon Renewed Guarantee

Share 15% off coupon
Wireless Keyboard and Mouse Combo, RATEL 2.4GHz Full...
\$24.99 prime
Sponsored

Products related to this item

Sponsored

Havit Wired Gaming
Wireless Keyboard and
Vaydeer USB3.0 Wireless
Ri-ck and Mo-RTY Mouse
Wireless Keyboard and
Manhattan Black Wired
Wireless Keyboard and

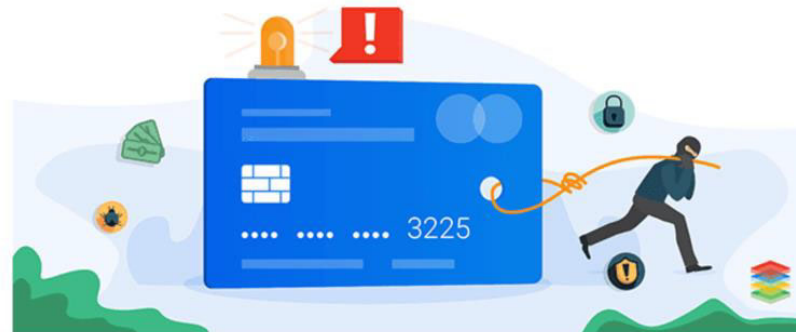
Page 1 of 108

Artificial Intelligence - esempi - finanza

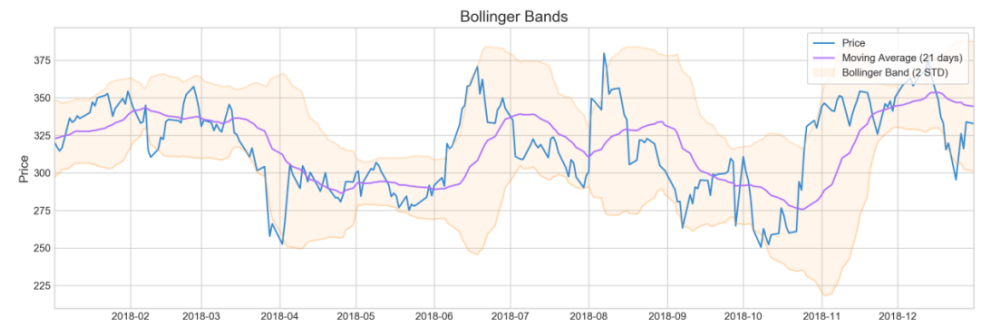
- ✚ Predire l'affidabilità finanziaria di un cliente



- ✚ Prevenire frodi

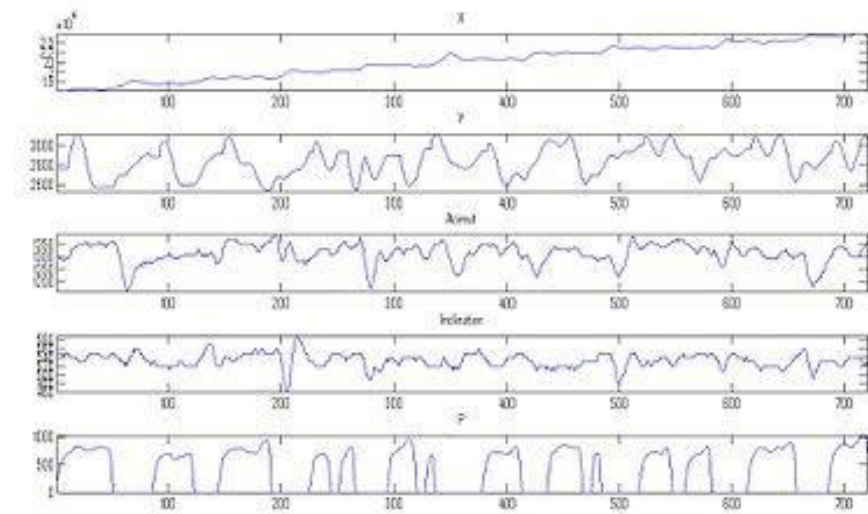
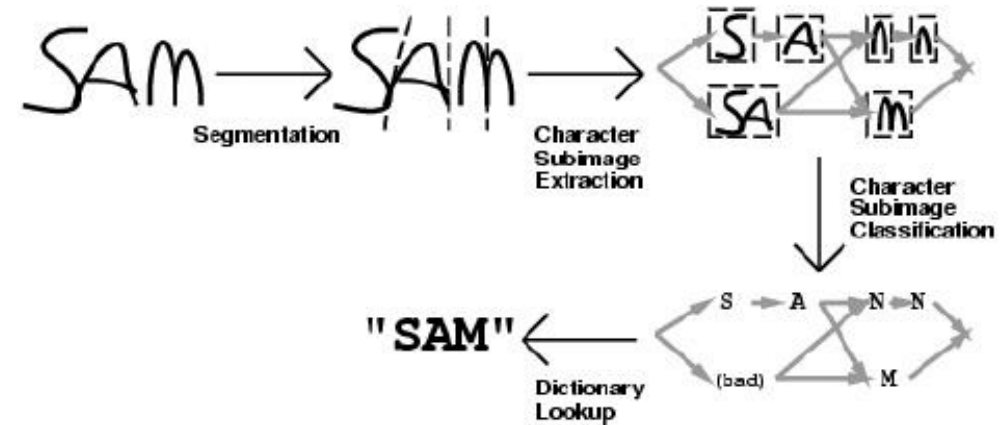


- ✚ Identificare trend promettenti nel mercato



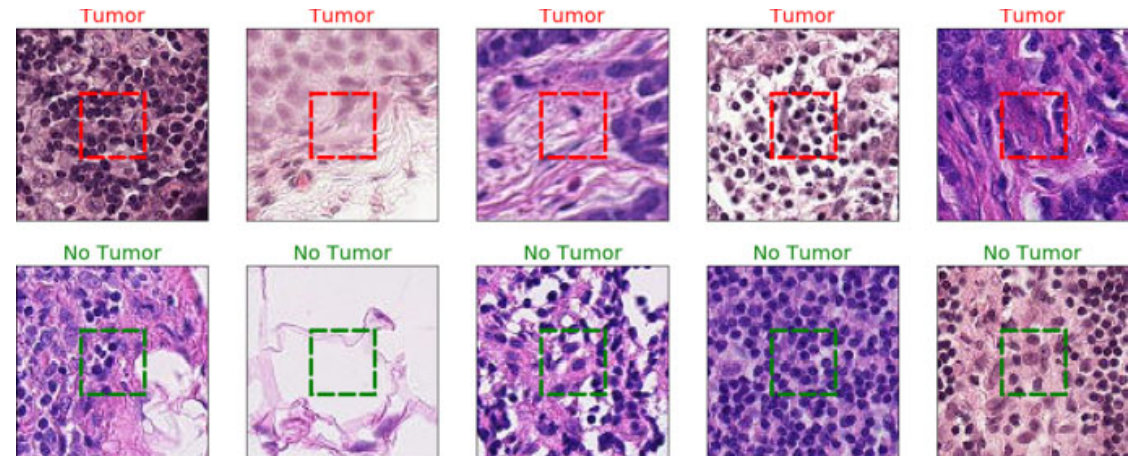
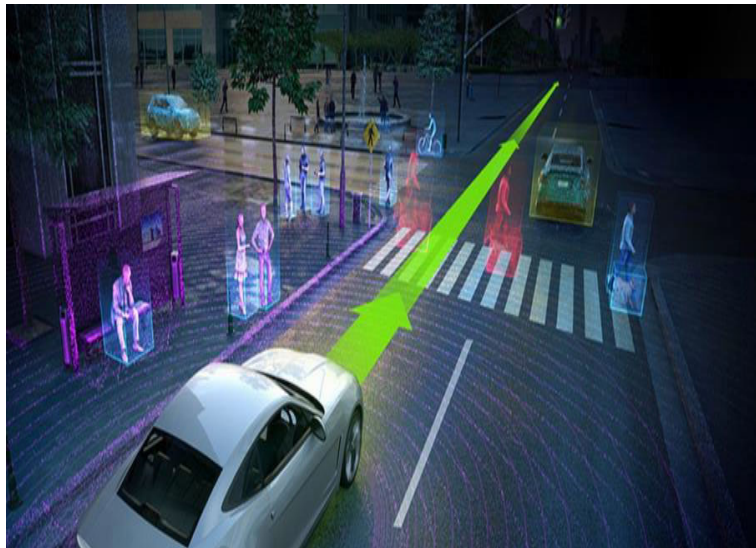
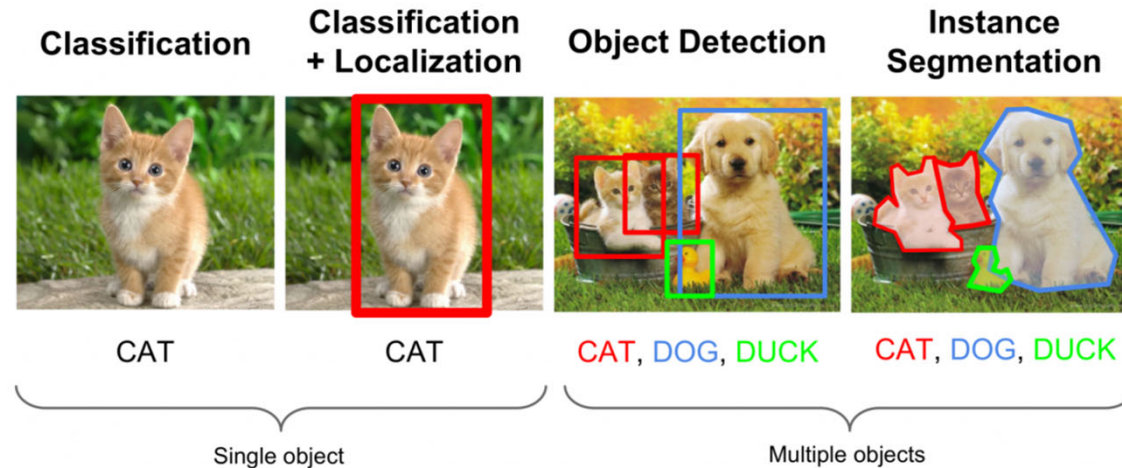
Artificial Intelligence - esempi - testi e audio

- ✚ Handwriting recognition
- ✚ Assistenti vocali(e.g., ALEXA)
- ✚ Sistemi di traduzione di lingue
- ✚ Natural Language Processing



Artificial Intelligence - esempi - immagini

✚ Riconoscimento di oggetti in immagini in applicazioni statiche e dinamiche





Focalizzeremo ora l'attenzione su:

modelli ed algoritmi di Artificial Intelligence

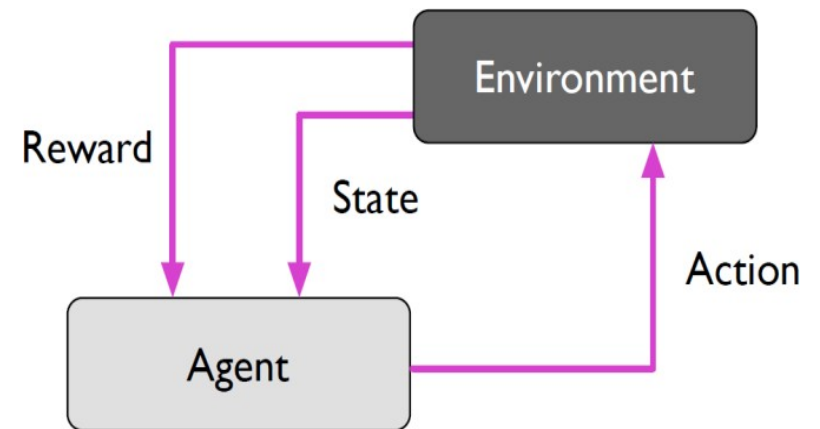
che possiamo impiegare per risolvere i complessi problemi di ottimizzazione robusta che abbiamo introdotto nella prima parte di questo seminario

Artificial Intelligence - Reinforcement learning

REINFORCEMENT LEARNING

Basato su **agenti** che:

- ad ogni iterazione, **si trovano in uno stato**
- possono **scegliere un'azione da un set di azioni che dipende dallo stato**
- **eseguita l'azione si spostano in un nuovo stato**



L'esecuzione di un'azione comporta

- **una ricompensa** (se l'azione ha portato un vantaggio)
- **una penalità** (se l'azione ha portato uno svantaggio)



ESEMPI DI APPLICAZIONE:

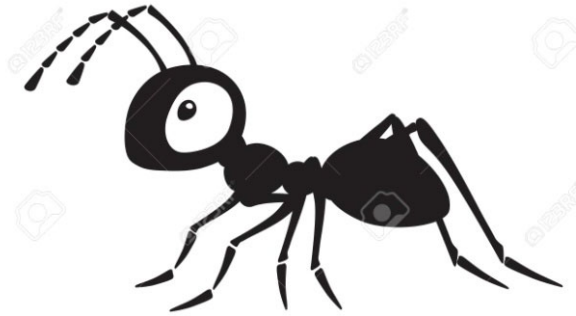
- 1) policy di movimento di robot
- 2) costruzione sequenziale di soluzioni di un problema
- 3) giochi (e.g., DeepMind),

Swarm Intelligence - Ant Colony Optimization

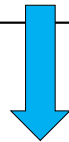


“Ant Colony Optimization (ACO) studies artificial systems that take inspiration from the *behavior of real ant colonies* and which are used to solve discrete optimization problems.” (Prof. Dr. **Marco Dorigo, father of ACO**)

Una formica nel mondo reale (1)

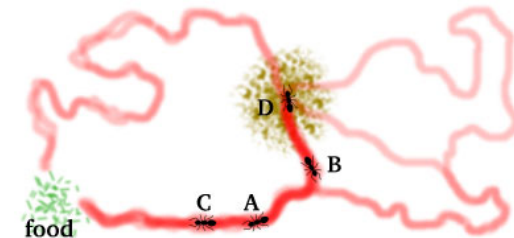
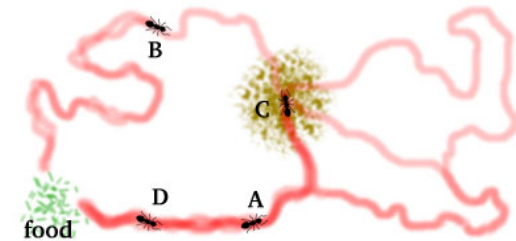
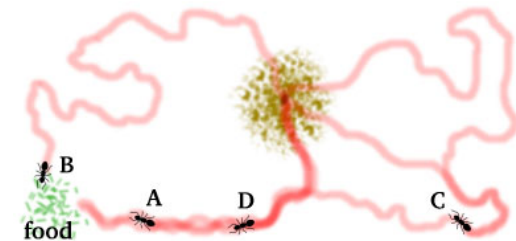
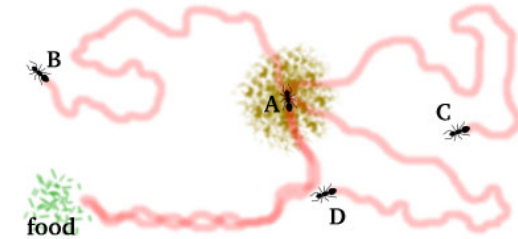
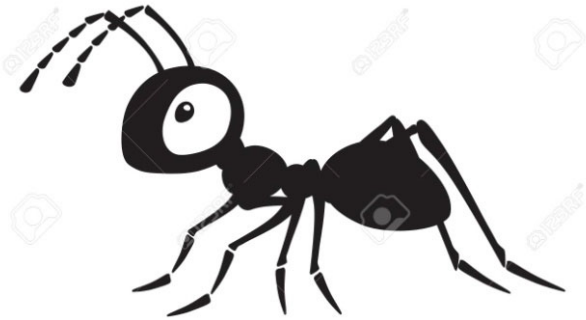


- Praticamente cieca
- Incapace di eseguire compiti complessi da sola
- Si affida a dinamiche di "swarm intelligence" per sopravvivere
- In grado di stabilire "cammini di lunghezza minima" dalla colonia a fonti di cibo
- Utilizza forme di comunicazione stigmergiche (scie di feromone - pheromone trails).



forma di comunicazione indiretta impiegata dagli insetti sociali per coordinare le loro attività mediante la modifica dell'ambiente in cui operano (rilascio di feromone)

Una formica nel mondo reale (1)



- Segue **pheromone trails** con alta probabilità
- Più formiche seguono una scia di pheromone, più la scia diventa attraente da seguire per altre formiche.
- Il processo di ricerca del cibo delle formiche costituisce un **positive feedback loop**, in cui la **probabilità di scelta di un cammino (path)** incrementa con il numero di volte che il cammino è stato scelto precedentemente

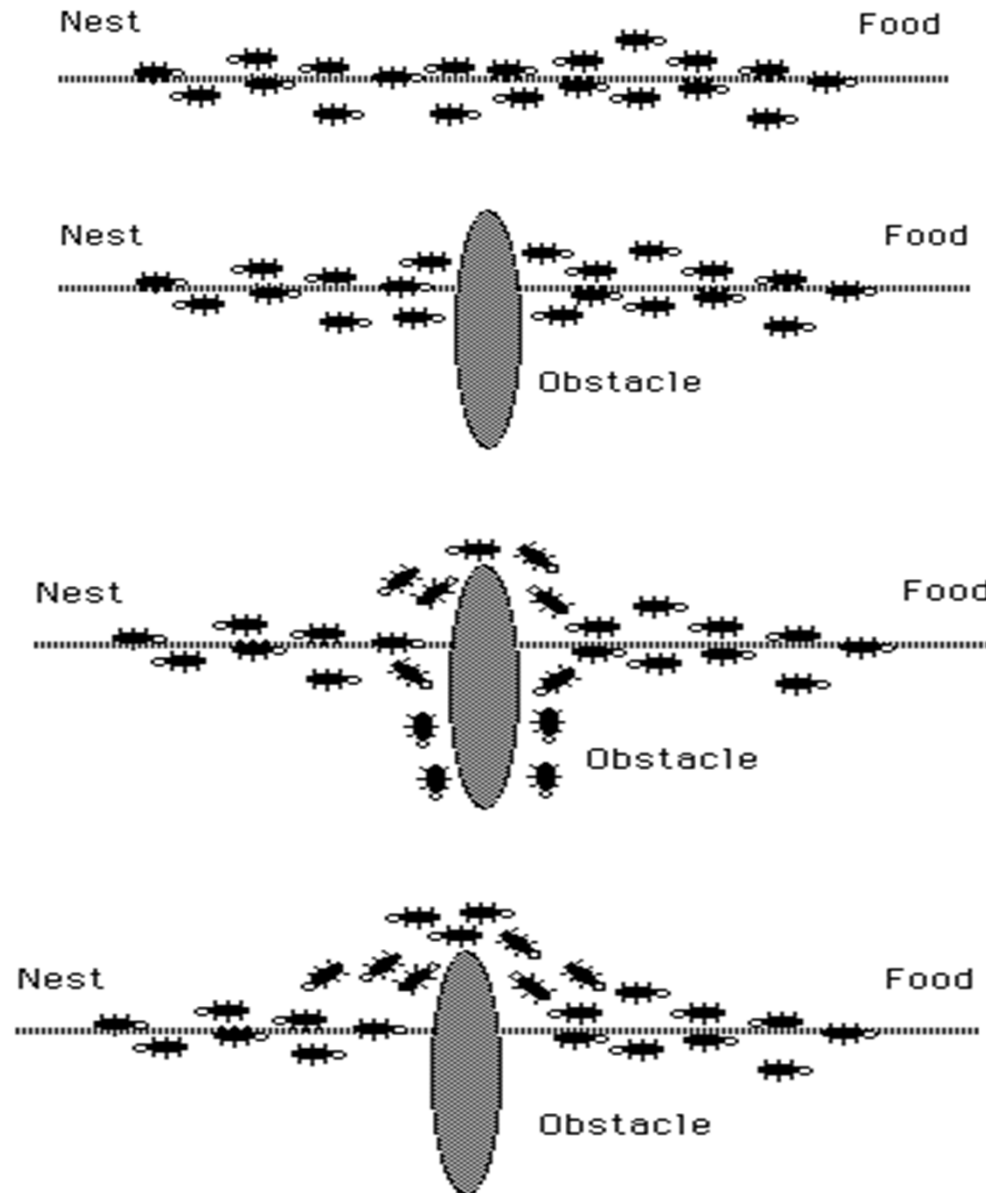
Compartamento di una colonia di formiche

STABLE
ROUTE

PERTURBATION

REORGANIZATION

NEW OPTIMAL
STABLE
ROUTE



Ant Colony Optimization (1)

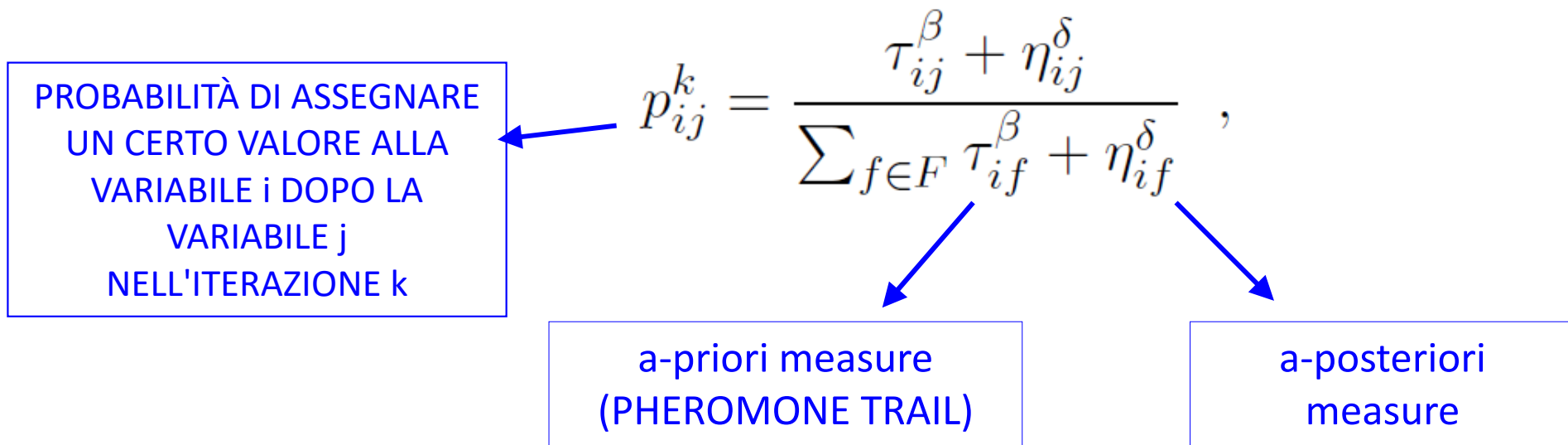
- Un famosissimo **algoritmo di swarm intelligence** ideato da Marco Dorigo (1992)
- Risultato di **studi di computational intelligence applicata al campo dell'Ottimizzazione Combinatoria**
- Inizialmente impiegato per risolvere il **Traveling Salesman Problem**
- Impiego poi esteso a molti altri problemi di ottimizzazione NP-Hard
- Oggetto di numerosissimi studi teorici e pratici sviluppati negli anni

Algorithm 1 General ACO Algorithm (ACOA)

```
1: while an arrest condition is not satisfied do  
2:   - ant-based solution construction  
3:   - pheromone trail update  
4: end while  
5: - daemon actions
```

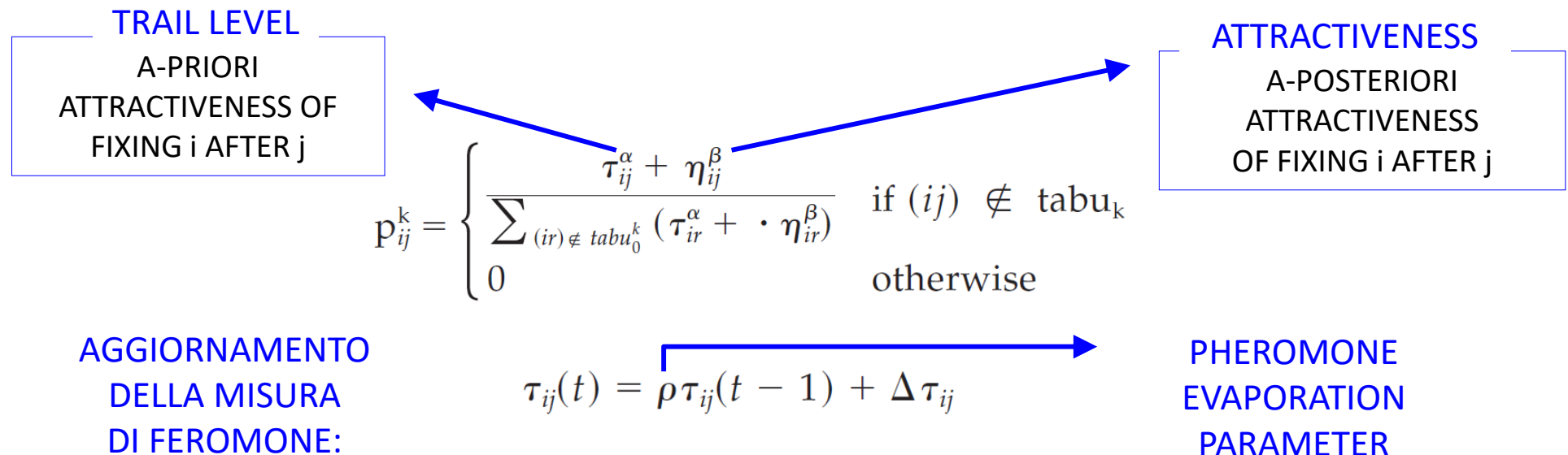
Ant Colony Optimization (2)

- ✚ Ad ogni esecuzione del ciclo di costruzione di soluzioni basato su formiche, un insieme di m formiche costruisce m soluzioni per il problema di ottimizzazione
- ✚ Ad ogni iterazione, una formica si trova in uno **stato** (= **soluzione parziale del problema**)
- ✚ La formica può completare ulteriormente la soluzione parziale **effettuando una mossa (azione o action)** (corrispondente a **determinare il valore di una variabile decisionale**)
- ✚ La **mossa è scelta in modo probabilistico** combinando:
 - una misura a-priori (pheromone)
 - una misura a-posteriori**di attrattività** di assegnare ad una variabile un certo valore and an a-posteriori measure



Ant Colony Optimization (3)

- ✚ Si tratta effettivamente di **una forma di algoritmo di Reinforcement Learning** che fissa probabilisticamente il valore di variabili decisionali tramite **agenti ispirati da formiche (ants)**
- ✚ Ogni ant prova a definire una soluzione ammissibile per il problema di ottimizzazione
- ✚ **Probabilità di fissare il valore di una variabile** i dopo una variabile j espresso dalla formula generale:



Ant Colony Optimization (3) - challenges

Difficoltà di applicazione:

- ✚ non esiste regola generale su come definire le funzioni di attrattività (devono essere adattate allo specifico problema)
- ✚ Scelta discutibile delle funzioni di reward/penalty
 - numerosi parametri difficili da settare
 - operazioni aritmetiche inefficienti
- ✚ **WEAK CONVERGENCE**: le soluzioni prodotte sono tipicamente di qualità random e devono essere raffinate

Una nota preliminare



NON INTENDIAMO USARE L'ANT COLONY CLASSICA

DRAWBACKS:

- too random and unpredictable behaviour
- the local search always does the real job

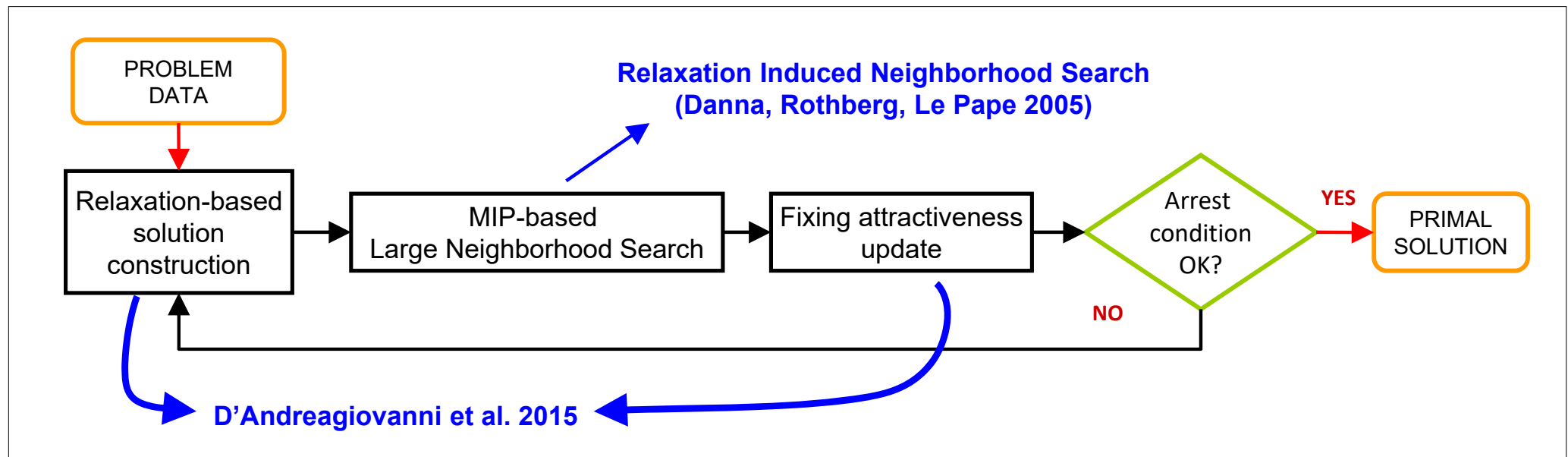


INVECE, l'obiettivo è impiegare l'ANT COLONY OPTIMIZATION come punto di partenza, interpretandola come una strategia probabilistica di valorizzazione delle variabili decisionali, RAFFORZATA TRAMITE TECNICHE DI "EXACT OPTIMIZATION"

Una hybrid robust learning heuristic (1)

APPROCCIO ORIGINALE CHE ABBIAMO PROPOSTO (D'Andreagiovanni, Krolikowski, Pulaj, 2015)

- Integrare più efficacemente tecniche di Ant Colony Optimization (ACO) con metodi di ottimizzazione esatta state-of-the-art
- RAFFORZARE** la fase di costruzione di soluzioni ACO con **modelli rilassati e rafforzati del modello di ottimizzazione**
- Sfruttare la potenza di software state-of-the-art di exact optimization per abbandonare l'uso di semplici euristiche di ricerca locale ed impiegare **Mixed Integer Programming-based exact large neighbourhood search**



Una hybrid robust learning heuristic (2)

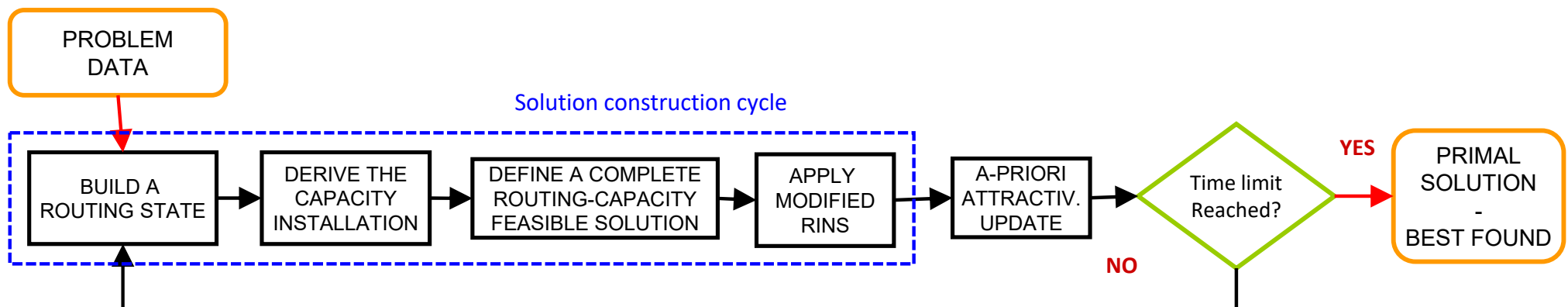
MIP HEURISTIC IMPROVEMENT

Una volta costruita una soluzione, cerchiamo di migliorarla tramite una large neighborhood search condotta in modo esatto

Modified Relaxation Induced Neighborhood Search (Mod-RINS)

- ✚ **Fix** those **variables** whose value is (almost) the same in a strong relaxation of the problem and in the current feasible solution
- ✚ **Solve exactly** the resulting (Mixed) Integer Problem **of reduced size** by using a powerful commercial solver

➡ SEMPLICE MA ESTREMAMENTE EFFICACE



Una hybrid robust learning heuristic (3)

PRINCIPALI CARATTERISTICHE DELL'ALGORITMO:

- semplificare le operazioni aritmetiche dell'Ant Colony Optimization
- basare le funzioni probabilistiche su un modello di ottimizzazione robusta e su un modello "matematicamente forte"

Una hybrid robust learning heuristic (3)

PRINCIPALI CARATTERISTICHE DELL'ALGORITMO:

- semplificare le operazioni aritmetiche dell'Ant Colony Optimization
- basare le funzioni probabilistiche su un modello di ottimizzazione robusta e su un modello "matematicamente forte"

+ Fast randomized fixing of path-routing variables

A-PRIORI FIXING ATTRACTIVENESS

FIXING PROBABILITY
(path variable j
at iteration t)

$$p_j(t) = \frac{\alpha \cdot \tau_j(t) + (1 - \alpha) \cdot \eta_j(t)}{\sum_{k \in J} \alpha \cdot \tau_k(t) + (1 - \alpha) \cdot \eta_k(t)} \in [0, 1]$$

A-POSTERIORI
FIXING ATTRACTIVENESS

COMBINATION
PARAMETER

- A-priori fixing attractiveness $\longrightarrow$ Linear relaxation (**robust formulation**)
- A-posteriori fixing attractiveness $\longrightarrow$ Linear relaxation (**deterministic formulation**)

+ Update of a-priori attractiveness after each complete fixing $\tau_j(t) = \tau_j(t-1) + \sum_{s=1}^m \Delta \tau_j^s$

+ Fast solution of the deterministic linear relaxation
(CPLEX proved too slow when solving the relaxation)

$$\Delta \tau_j^s = \tau_j(0) \cdot \left(1 - \frac{z_{curr} - LB}{\bar{z} - LB} \right)$$

Risultati computazionali

✚ Questa nuova LearnHeuristic da noi proposta ci ha permesso di ottenere **risultati computazionali davvero impressionanti** per un **problema di ottimizzazione robusta relativo al design pluriennale di reti di telecomunicazioni** :

- optimality gap molto bassi (individuazione di soluzioni probabilmente ottime)
- **miglioramenti del 3700%** rispetto al software di ottimizzazione IBM CPLEX in termini di qualità delle soluzioni individuate

✚ Per maggiori dettagli sull'algoritmo, contattatemi pure offline all'indirizzo:
f.dandreagiovanni@gmail.com

CONCLUSIONI



Abbiamo discusso:

- la **questione dell'impatto di dati incerti in problemi di ottimizzazione**
- **l'importanza di modellare l'incertezza di dati, al fine di evitare errori**
- introdotto la **Robust Optimization** come un metodo efficace ed efficiente per affrontare l'incertezza di dati
- la **Robust Artificial Intelligence** come una famiglia di metodi per la risoluzione di problemi di (Robust) Mathematical Optimization, focalizzando l'attenzione su **Learning Heuristics** (algoritmi ibridi euristici - AI Machine Learning)



Per ulteriori chiarimenti e discussioni, contattatemi pure offline all'indirizzo email: f.dandreagiovanni@gmail.com



**SARÒ ORA FELICE DI RISPONDERE
ALLE VOSTRE DOMANDE**