Counterfactual explanations for Workforce Scheduling and Routing Problems

Mathieu Lerouge
mathieulerouge.github.io
MICS, CentraleSupélec Université Paris-Saclay

ICORES
19 February 2023

CentraleSupélec | université PARIS-SACLAY

DecisionBrain™
Smarter Decisions. Better Results.

Plan

1. Introduction
2. Related works
3. Mathematical model of the WSRP
4. Generating counterfactual explanations
5. Conclusion

Plan

1. Introduction
2. Related works
3. Mathematical model of the WSRP
4. Generating counterfactual explanations
5. Conclusion

Plan

1. Introduction
   - Motivations
   - Use case
   - Our proposition
2. Related works
3. Mathematical model of the WSRP
4. Generating counterfactual explanations
5. Conclusion

Motivations

3 observations:

(1) Real-world situations modeled as **Combinatorial Optimization (CO) problems** (*e.g.* workforce management);

(2) CO problems solved using **optimization systems** that are developed by experts (*e.g.* DecisionBrain);

(3) Optimization systems are used by **non-expert people**.

↪ End-users may experience a **black box feeling**.

Let see (1), (2) and (3) in our use case.

Motivations

3 observations:

(1) Real-world situations modeled as **Combinatorial Optimization (CO) problems** (*e.g.* workforce management);

(2) CO problems solved using **optimization systems** that are developed by experts (*e.g.* DecisionBrain);

(3) Optimization systems are used by **non-expert people**.

↪ End-users may experience a **black box feeling**.

Let see (1), (2) and (3) in our use case.

Motivations

3 observations:

(1) Real-world situations modeled as **Combinatorial Optimization (CO) problems** (*e.g.* workforce management);

(2) CO problems solved using **optimization systems** that are developed by experts (*e.g.* DecisionBrain);

(3) Optimization systems are used by **non-expert people**.

↪ End-users may experience a **black box feeling**.

Let see (1), (2) and (3) in our use case.

Motivations

3 observations:

(1) Real-world situations modeled as **Combinatorial Optimization (CO) problems** (*e.g.* workforce management);

(2) CO problems solved using **optimization systems** that are developed by experts (*e.g.* DecisionBrain);

(3) Optimization systems are used by **non-expert people**.

↪ End-users may experience a **black box feeling**.

Let see (1), (2) and (3) in our use case.

Motivations

3 observations:

(1) Real-world situations modeled as **Combinatorial Optimization (CO) problems** (*e.g.* workforce management);

(2) CO problems solved using **optimization systems** that are developed by experts (*e.g.* DecisionBrain);

(3) Optimization systems are used by **non-expert people**.

↪ End-users may experience a **black box feeling**.

Let see (1), (2) and (3) in our use case.

Motivations

3 observations:

(1) Real-world situations modeled as **Combinatorial Optimization (CO) problems** (*e.g.* workforce management);

(2) CO problems solved using **optimization systems** that are developed by experts (*e.g.* DecisionBrain);

(3) Optimization systems are used by **non-expert people**.

↪ End-users may experience a **black box feeling**.

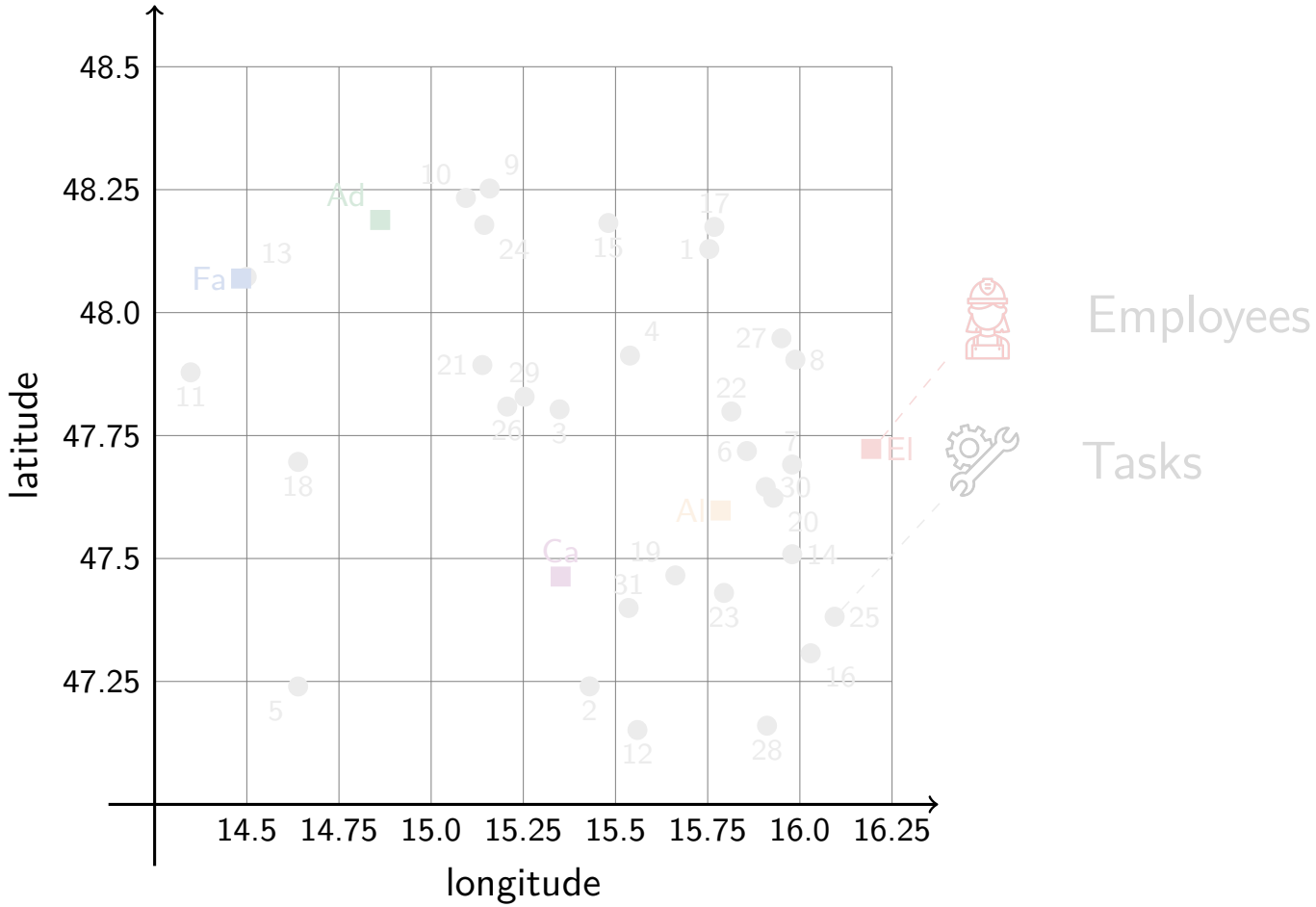
Let see (1), (2) and (3) in our use case.

Plan

1. Introduction
   - Motivations
   - Use case
   - Our proposition
2. Related works
3. Mathematical model of the WSRP
4. Generating counterfactual explanations
5. Conclusion

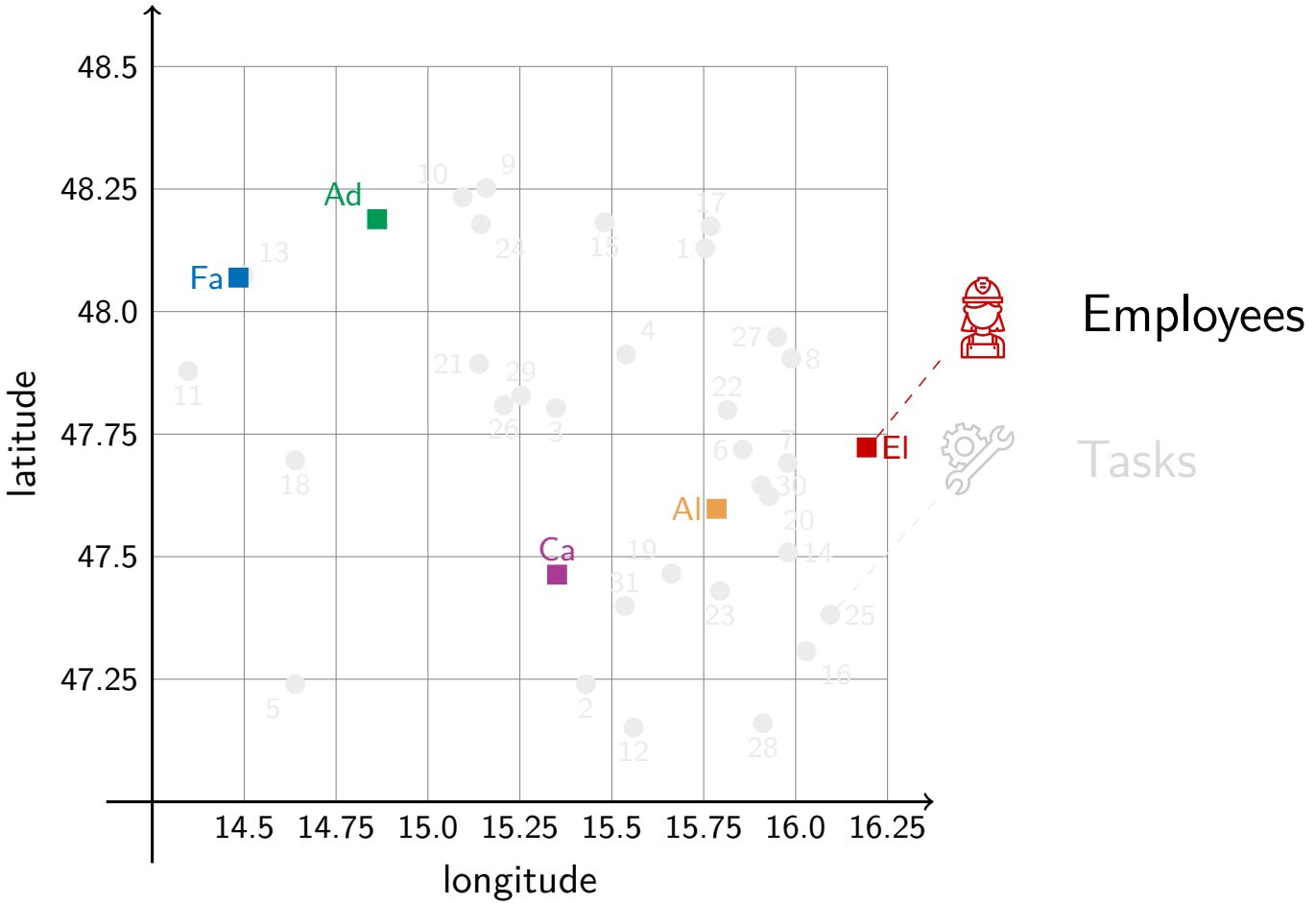
Use case - (1) CO problem

Workforce Scheduling and Routing Problem (WSRP):



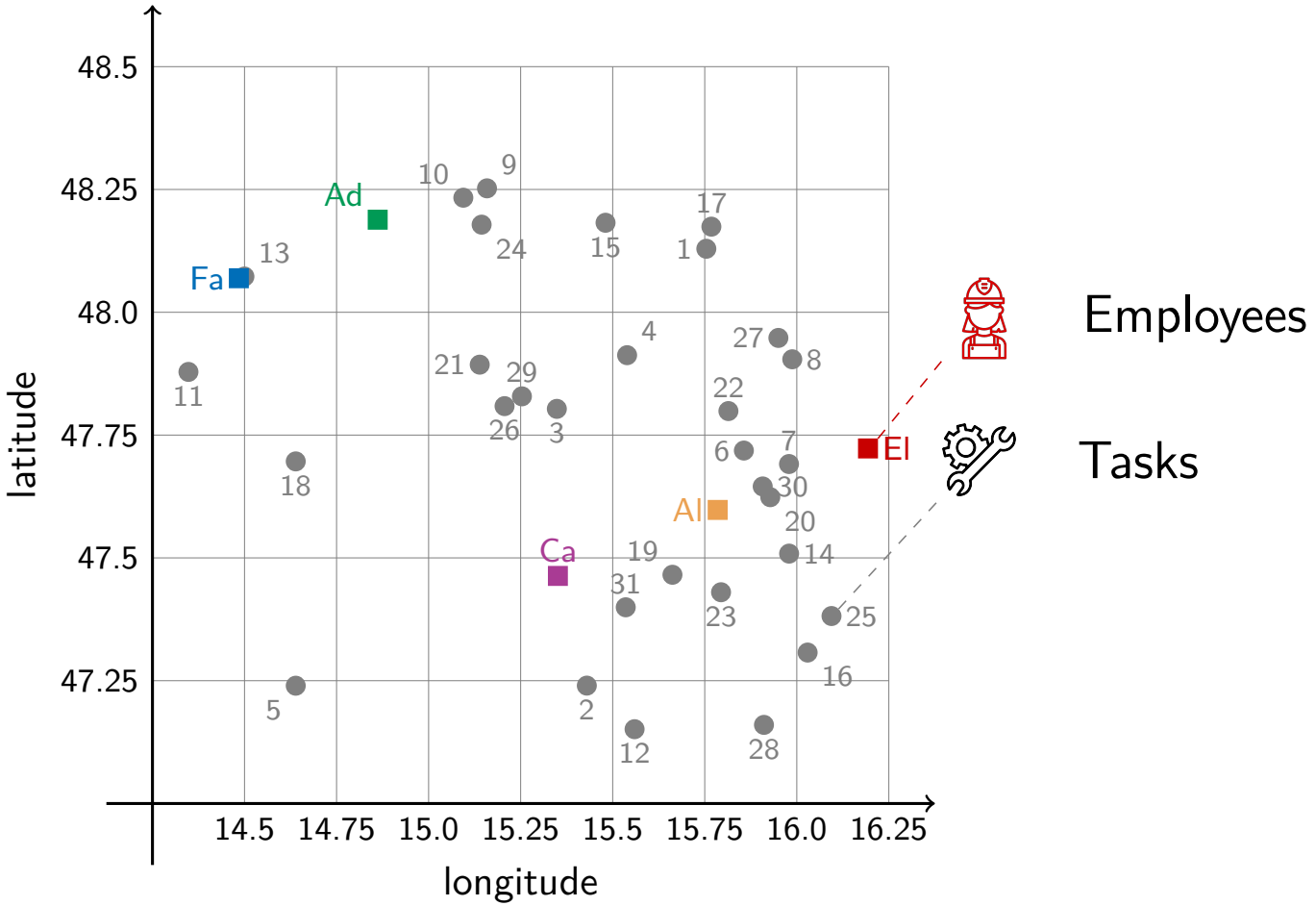
Use case - (1) CO problem

Workforce Scheduling and Routing Problem (WSRP):



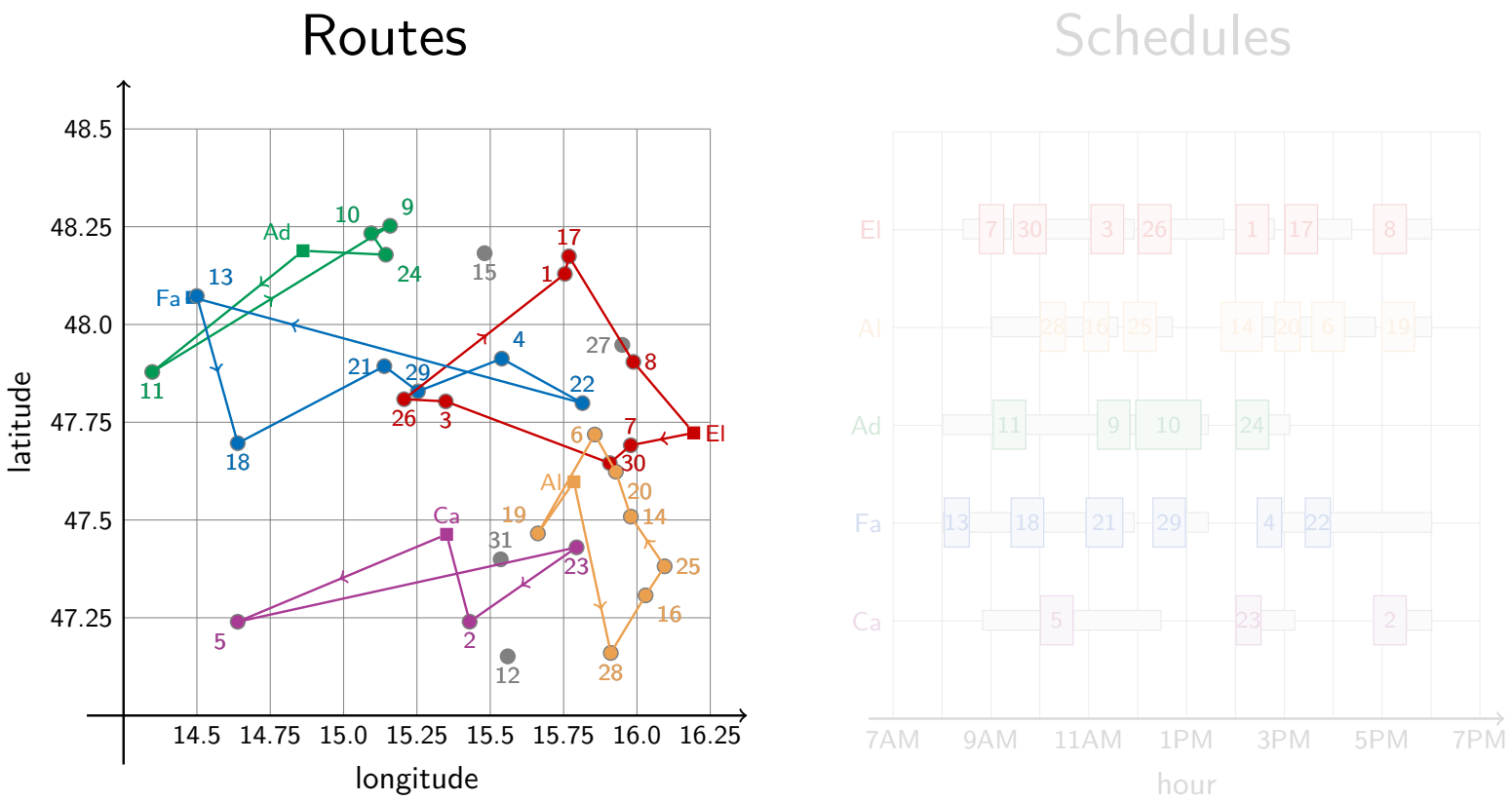
Use case - (1) CO problem

Workforce Scheduling and Routing Problem (WSRP):



Use case - (1) CO problem

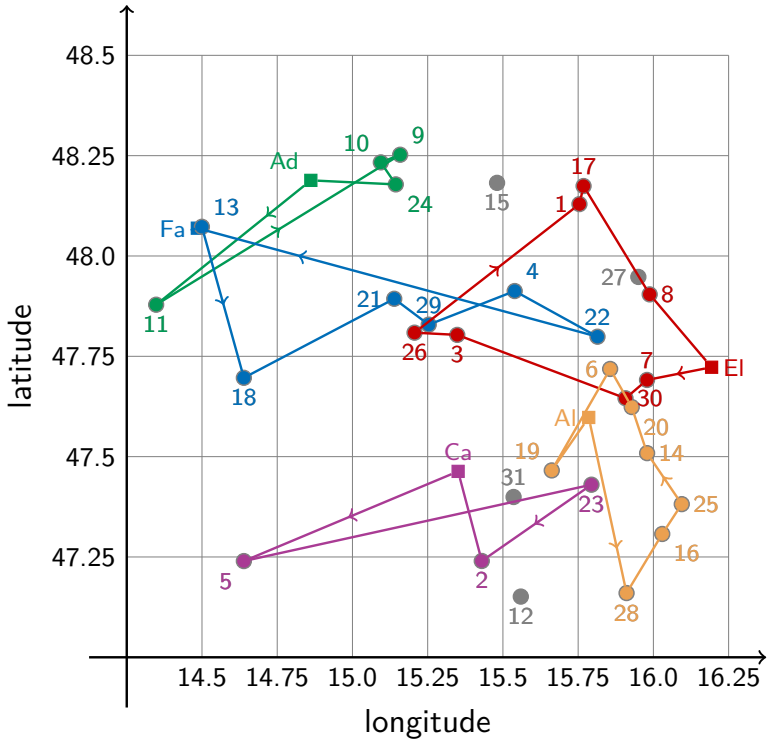
Workforce Scheduling and Routing Problem (WSRP):



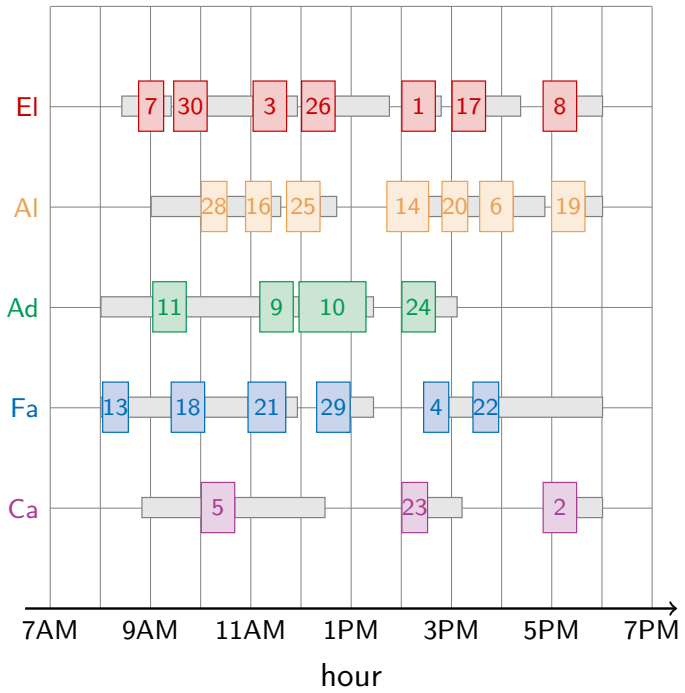
Use case - (1) CO problem

Workforce Scheduling and Routing Problem (WSRP):

Routes



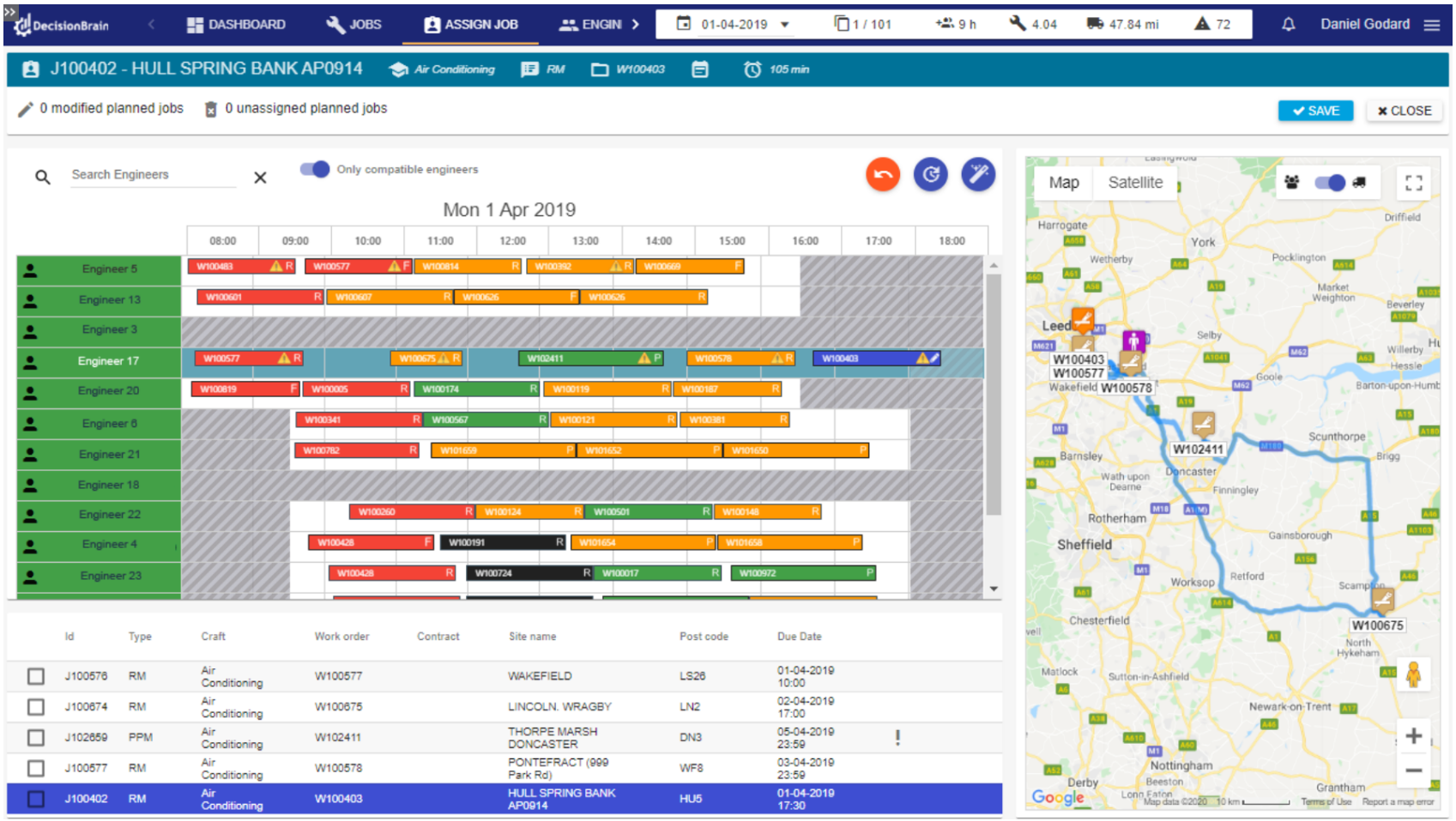
Schedules



Our use case - (2) Optimization system

WSRP-solving system:

e.g DecisionBrain's Dynamic Scheduler

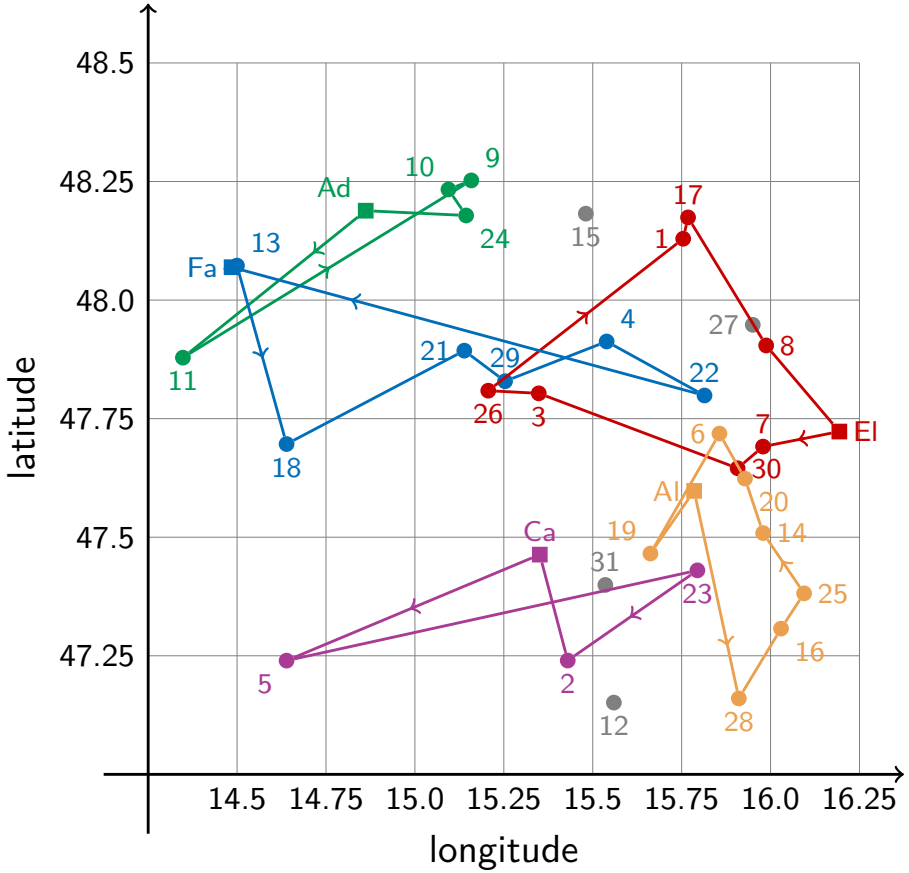


Use case - (3) Non-expert end-user

Planner:

Our use case - (3) Non-expert end-user

A problematic situation for a planner:

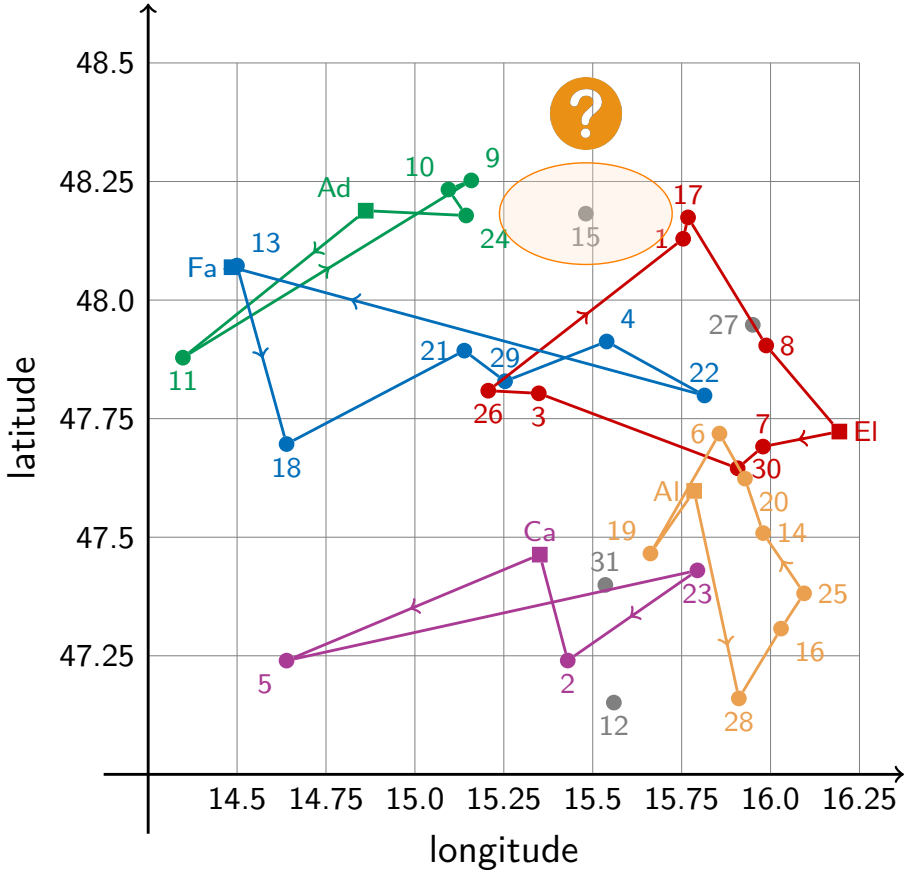


"How to make Ellen perform task 15 in addition to the tasks of her route?"

↪ If no explanations, then black box feeling...

Our use case - (3) Non-expert end-user

A problematic situation for a planner:

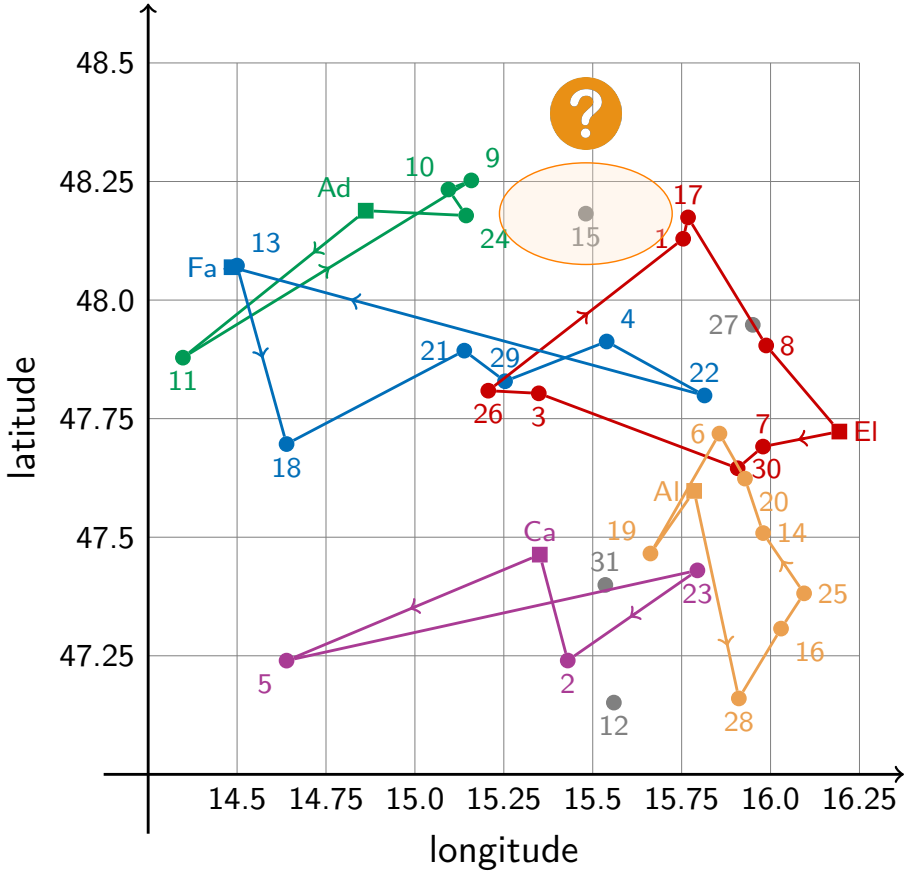


"How to make Ellen perform task 15 in addition to the tasks of her route?"

↪ If no explanations, then black box feeling...

Our use case - (3) Non-expert end-user

A problematic situation for a planner:

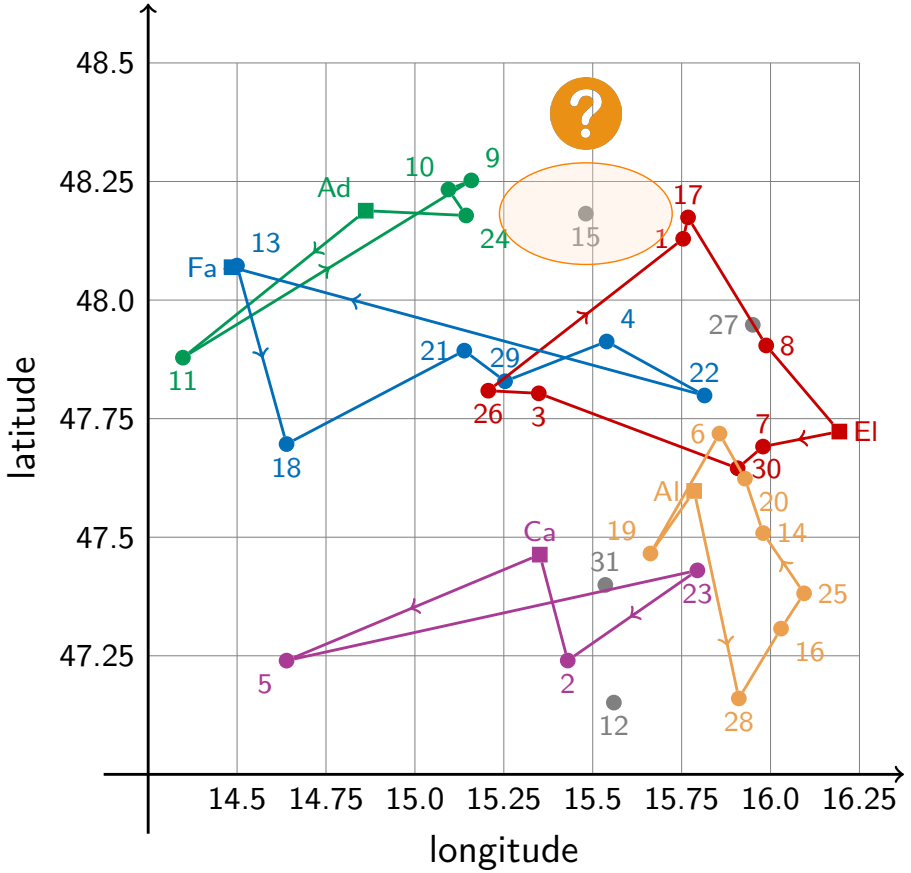


"How to make Ellen perform task 15 in addition to the tasks of her route?"

↪ If no explanations, then black box feeling...

Our use case - (3) Non-expert end-user

A problematic situation for a planner:



"How to make Ellen perform task 15 in addition to the tasks of her route?"

↪ If no explanations, then black box feeling...

Plan

1. Introduction
   - Motivations
   - Use case
   - Our proposition
2. Related works
3. Mathematical model of the WSRP
4. Generating counterfactual explanations
5. Conclusion

Our proposition

Tackling the black box feeling experienced by non-expert end-users solving WSRP instances, by generating **explanations** to help them better understand their WSRP solutions.

Our proposition

Tackling the black box feeling experienced by non-expert end-users solving WSRP instances, by generating **explanations** to help them better understand their WSRP solutions.

Our proposition

Tackling the black box feeling experienced by non-expert end-users solving WSRP instances, by generating **explanations** to help them better understand their WSRP solutions.

Plan

1. Introduction
2. Related works
3. Mathematical model of the WSRP
4. Generating counterfactual explanations
5. Conclusion

Plan

1. Introduction
2. Related works
   - Explanations in Artificial Intelligence
   - Our proposition (with more details)
3. Mathematical model of the WSRP
4. Generating counterfactual explanations
5. Conclusion

Literature on eXplainable Artificial Intelligence (XAI):

Works on explanations:

- **Many** in **Machine Learning** [Barredo Arrieta et al., 2020].
- **Some** in **other AI fields** including
  - Expert Systems, [Wick and Thompson, 1992],
  - Planning, *e.g.* [Chakraborti et al., 2020],
  - Constraint Programming, *e.g.* [Junker, 2004].
- **Few** ones in **Combinatorial Optimization (CO)**, *e.g.* [Korikov et al., 2021].

↪ Survey **concepts about explanations** in AI fields other than CO and transpose them to CO.

[Barredo Arrieta et al., 2020] XAI: Concepts, taxonomies, opportunities and challenges toward responsible AI
[Wick and Thompson, 1992] Reconstructive Expert System explanation
[Chakraborti et al., 2020] The emerging landscape of explainable AI planning and decision making
[Junker, 2004] QuickXplain: preferred explanations and relaxations for over-constrained problems
[Korikov et al., 2021] Counterfactual Explanations for Optimization-Based Decisions in the Context of the GDPR

Literature on eXplainable Artificial Intelligence (XAI):

Works on explanations:

- **Many** in **Machine Learning** [Barredo Arrieta et al., 2020].
- **Some** in **other AI fields** including
  - Expert Systems, [Wick and Thompson, 1992],
  - Planning, *e.g.* [Chakraborti et al., 2020],
  - Constraint Programming, *e.g.* [Junker, 2004].
- **Few** ones in **Combinatorial Optimization (CO)**, *e.g.* [Korikov et al., 2021].

↪ Survey **concepts about explanations** in AI fields other than CO and transpose them to CO.

[Barredo Arrieta et al., 2020] XAI: Concepts, taxonomies, opportunities and challenges toward responsible AI
[Wick and Thompson, 1992] Reconstructive Expert System explanation
[Chakraborti et al., 2020] The emerging landscape of explainable AI planning and decision making
[Junker, 2004] QuickXplain: preferred explanations and relaxations for over-constrained problems
[Korikov et al., 2021] Counterfactual Explanations for Optimization-Based Decisions in the Context of the GDPR

Literature on eXplainable Artificial Intelligence (XAI):

Works on explanations:

- **Many** in **Machine Learning** [Barredo Arrieta et al., 2020].
- **Some** in **other AI fields** including
  - Expert Systems, [Wick and Thompson, 1992],
  - Planning, *e.g.* [Chakraborti et al., 2020],
  - Constraint Programming, *e.g.* [Junker, 2004].
- **Few** ones in **Combinatorial Optimization (CO)**, *e.g.* [Korikov et al., 2021].

↪ Survey **concepts about explanations** in AI fields other than CO and transpose them to CO.

[Barredo Arrieta et al., 2020] XAI: Concepts, taxonomies, opportunities and challenges toward responsible AI
[Wick and Thompson, 1992] Reconstructive Expert System explanation
[Chakraborti et al., 2020] The emerging landscape of explainable AI planning and decision making
[Junker, 2004] QuickXplain: preferred explanations and relaxations for over-constrained problems
[Korikov et al., 2021] Counterfactual Explanations for Optimization-Based Decisions in the Context of the GDPR

Literature on eXplainable Artificial Intelligence (XAI):

Works on explanations:

- **Many** in **Machine Learning** [Barredo Arrieta et al., 2020].
- **Some** in **other AI fields** including
  - Expert Systems, [Wick and Thompson, 1992],
  - Planning, *e.g.* [Chakraborti et al., 2020],
  - Constraint Programming, *e.g.* [Junker, 2004].
- **Few** ones in **Combinatorial Optimization (CO)**, *e.g.* [Korikov et al., 2021].

↪ Survey **concepts about explanations** in AI fields other than CO and transpose them to CO.

[Barredo Arrieta et al., 2020] XAI: Concepts, taxonomies, opportunities and challenges toward responsible AI
[Wick and Thompson, 1992] Reconstructive Expert System explanation
[Chakraborti et al., 2020] The emerging landscape of explainable AI planning and decision making
[Junker, 2004] QuickXplain: preferred explanations and relaxations for over-constrained problems
[Korikov et al., 2021] Counterfactual Explanations for Optimization-Based Decisions in the Context of the GDPR

Literature on eXplainable Artificial Intelligence (XAI):

Works on explanations:

- **Many** in **Machine Learning** [Barredo Arrieta et al., 2020].
- **Some** in **other AI fields** including
  - Expert Systems, [Wick and Thompson, 1992],
  - Planning, *e.g.* [Chakraborti et al., 2020],
  - Constraint Programming, *e.g.* [Junker, 2004].
- **Few** ones in **Combinatorial Optimization (CO)**, *e.g.* [Korikov et al., 2021].

↪ Survey **concepts about explanations** in AI fields other than CO and transpose them to CO.

[Barredo Arrieta et al., 2020] XAI: Concepts, taxonomies, opportunities and challenges toward responsible AI
[Wick and Thompson, 1992] Reconstructive Expert System explanation
[Chakraborti et al., 2020] The emerging landscape of explainable AI planning and decision making
[Junker, 2004] QuickXplain: preferred explanations and relaxations for over-constrained problems
[Korikov et al., 2021] Counterfactual Explanations for Optimization-Based Decisions in the Context of the GDPR

Some recurrent concepts in XAI methods:

Explanations are often:

- **local** [Wick and Thompson, 1992], *i.e.* focusing on a specific output generated by the system.
- **counterfactual** [Wachter et al., 2018], *i.e.* explanations which present alterations in the inputs that would have resulted in different outputs (*e.g.* end-user-specified one).
- formulated as texts using **templates**, *e.g.* [Krarup et al., 2021].

[Wick and Thompson, 1992] Reconstructive expert system explanation
[Wachter et al., 2018] Counterfactual explanations w/o opening the black box: automated decisions and the GDPR
[Krarup et al., 2021] Contrastive explanations of plans through model restrictions

Some recurrent concepts in XAI methods:

Explanations are often:

- **local** [Wick and Thompson, 1992], *i.e.* focusing on a specific output generated by the system.
- **counterfactual** [Wachter et al., 2018], *i.e.* explanations which present alterations in the inputs that would have resulted in different outputs (*e.g.* end-user-specified one).
- formulated as texts using **templates**, *e.g.* [Krarup et al., 2021].

[Wick and Thompson, 1992] Reconstructive expert system explanation
[Wachter et al., 2018] Counterfactual explanations w/o opening the black box: automated decisions and the GDPR
[Krarup et al., 2021] Contrastive explanations of plans through model restrictions

Some recurrent concepts in XAI methods:

Explanations are often:

- **local** [Wick and Thompson, 1992], *i.e.* focusing on a specific output generated by the system.
- **counterfactual** [Wachter et al., 2018], *i.e.* explanations which present alterations in the inputs that would have resulted in different outputs (*e.g.* end-user-specified one).
- formulated as texts using **templates**, *e.g.* [Krarup et al., 2021].

[Wick and Thompson, 1992] Reconstructive expert system explanation
[Wachter et al., 2018] Counterfactual explanations w/o opening the black box: automated decisions and the GDPR
[Krarup et al., 2021] Contrastive explanations of plans through model restrictions

Some recurrent concepts in XAI methods:

Explanations are often:

- **local** [Wick and Thompson, 1992], *i.e.* focusing on a specific output generated by the system.
- **counterfactual** [Wachter et al., 2018], *i.e.* explanations which present alterations in the inputs that would have resulted in different outputs (*e.g.* end-user-specified one).
- formulated as texts using **templates**, *e.g.* [Krarup et al., 2021].

[Wick and Thompson, 1992] Reconstructive expert system explanation
[Wachter et al., 2018] Counterfactual explanations w/o opening the black box: automated decisions and the GDPR
[Krarup et al., 2021] Contrastive explanations of plans through model restrictions

Plan

1. Introduction
2. Related works
   - Explanations in Artificial Intelligence
   - Our proposition (with more details)
3. Mathematical model of the WSRP
4. Generating counterfactual explanations
5. Conclusion

Our proposition (with more details)

Tackling the black box feeling experienced by non-expert end-users solving WSRP instances, by generating **explanations** to help them better understand their WSRP solutions.

Generating **explanations** which are:

- focusing **on a given solution**,
- **counterfactual**,
- expressed as **texts** using **templates**.

Our proposition (with more details)

Tackling the black box feeling experienced by non-expert end-users solving WSRP instances, by generating **explanations** to help them better understand their WSRP solutions.

Generating **explanations** which are:

- focusing **on a given solution**,
- **counterfactual**,
- expressed as **texts** using **templates**.

Our proposition (with more details)

Tackling the black box feeling experienced by non-expert end-users solving WSRP instances, by generating **explanations** to help them better understand their WSRP solutions.

Generating **explanations** which are:

- focusing **on a given solution**,
- **counterfactual**,
- expressed as **texts** using **templates**.

Our proposition (with more details)

Tackling the black box feeling experienced by non-expert end-users solving WSRP instances, by generating **explanations** to help them better understand their WSRP solutions.

Generating **explanations** which are:

- focusing **on a given solution**,
- **counterfactual**,
- expressed as **texts** using **templates**.

Our proposition (with more details)

Tackling the black box feeling experienced by non-expert end-users solving WSRP instances, by generating **explanations** to help them better understand their WSRP solutions.

Generating **explanations** which are:

- focusing **on a given solution**,
- **counterfactual**,
- expressed as **texts** using **templates**.

Plan

1. Introduction
2. Related works
3. Mathematical model of the WSRP
4. Generating counterfactual explanations
5. Conclusion

Mathematical model of the WSRP

Instance data:

- $\mathcal{E} = \{1, \ldots, n\}$ set of **mobile employees**,
  each employee i characterized by:
  - a **skill level** $ske_i \in \mathbb{N}$;
  - a working **time-window** $[lbe_i, ube_i] \subset [\![0, 1440]\!] \subset \mathbb{N}$;
  - a **location**.
- $\mathcal{T} = \{1, \ldots, m\}$ set of **tasks**,
  each task j characterized by:
  - a minimum required **skill level** $skt_j \in \mathbb{N}$;
  - an availability **time-window** $[lbt_j, ubt_j] \subset [\![0, 1440]\!] \subset \mathbb{N}$;
  - a performing **duration** $dt_j \in \mathbb{N}$;
  - a **location**.
- **Travel duration** between two locations tr_{jk} with $(j, k) \in \mathcal{T}^2$.

Mathematical model of the WSRP

Instance data:

- $\mathcal{E} = \{1, \ldots, n\}$ set of **mobile employees**,
  each employee i characterized by:
  - a **skill level** $ske_i \in \mathbb{N}$;
  - a working **time-window** $[lbe_i, ube_i] \subset [\![0, 1440]\!] \subset \mathbb{N}$;
  - a **location**.
- $\mathcal{T} = \{1, \ldots, m\}$ set of **tasks**,
  each task j characterized by:
  - a minimum required **skill level** $skt_j \in \mathbb{N}$;
  - an availability **time-window** $[lbt_j, ubt_j] \subset [\![0, 1440]\!] \subset \mathbb{N}$;
  - a performing **duration** $dt_j \in \mathbb{N}$;
  - a **location**.
- **Travel duration** between two locations tr_{jk} with $(j, k) \in \mathcal{T}^2$.

Mathematical model of the WSRP

Instance data:

- $\mathcal{E} = \{1, \ldots, n\}$ set of **mobile employees**,
  each employee i characterized by:
  - a **skill level** $ske_i \in \mathbb{N}$;
  - a working **time-window** $[lbe_i, ube_i] \subset [\![0, 1440]\!] \subset \mathbb{N}$;
  - a **location**.
- $\mathcal{T} = \{1, \ldots, m\}$ set of **tasks**,
  each task j characterized by:
  - a minimum required **skill level** $skt_j \in \mathbb{N}$;
  - an availability **time-window** $[lbt_j, ubt_j] \subset [\![0, 1440]\!] \subset \mathbb{N}$;
  - a performing **duration** $dt_j \in \mathbb{N}$;
  - a **location**.
- **Travel duration** between two locations tr_{jk} with $(j, k) \in \mathcal{T}^2$.

Mathematical model of the WSRP

Instance data:

- $\mathcal{E} = \{1, \ldots, n\}$ set of **mobile employees**,
  each employee i characterized by:
  - a **skill level** $ske_i \in \mathbb{N}$;
  - a working **time-window** $[lbe_i, ube_i] \subset [\![0, 1440]\!] \subset \mathbb{N}$;
  - a **location**.
- $\mathcal{T} = \{1, \ldots, m\}$ set of **tasks**,
  each task j characterized by:
  - a minimum required **skill level** $skt_j \in \mathbb{N}$;
  - an availability **time-window** $[lbt_j, ubt_j] \subset [\![0, 1440]\!] \subset \mathbb{N}$;
  - a performing **duration** $dt_j \in \mathbb{N}$;
  - a **location**.
- **Travel duration** between two locations tr_{jk} with $(j, k) \in \mathcal{T}^2$.

Mathematical model of the WSRP

Bi-objective Integer Linear Program (ILP) model:

lex max (total working duration, −total traveling duration)

s.t.
- flow constraints
- skill constraints
- occurence constraints
- tasks availability constraints
- employees working hours constraints
- sequencing constraints

$U_{ijk} \in \{0,1\}$ whether or not i goes from j to k, $\forall i \in \mathcal{E},\ \forall (j,k) \in \mathcal{T}^2$

$T_j \in \mathbb{N}$ start time of j, $\forall j \in \mathcal{T}$

Mathematical model of the WSRP

Bi-objective Integer Linear Program (ILP) model:

lex max (total working duration, –total traveling duration)

s.t.
- flow constraints
- skill constraints
- occurence constraints
- tasks availability constraints
- employees working hours constraints
- sequencing constraints

$U_{ijk} \in \{0,1\}$ whether or not i goes from j to k, $\forall i \in \mathcal{E},\ \forall (j,k) \in \mathcal{T}^2$

$T_j \in \mathbb{N}$ start time of j, $\forall j \in \mathcal{T}$

Mathematical model of the WSRP

Bi-objective Integer Linear Program (ILP) model:

lex max $\big($total working duration, $-$total traveling duration$\big)$

s.t.
- flow constraints
- skill constraints
- occurence constraints
- tasks availability constraints
- employees working hours constraints
- sequencing constraints

$U_{ijk} \in \{0,1\}$ whether or not i goes from j to k, $\quad \forall i \in \mathcal{E},\ \forall (j,k) \in \mathcal{T}^2$

$T_j \in \mathbb{N}$ start time of j, $\quad \forall j \in \mathcal{T}$

Mathematical model of the WSRP

Bi-objective Integer Linear Program (ILP) model:

lex max $\big($total working duration, $-$total traveling duration$\big)$

s.t.
- flow constraints
- skill constraints
- occurence constraints
- tasks availability constraints
- employees working hours constraints
- sequencing constraints

$U_{ijk} \in \{0,1\}$ whether or not i goes from j to k, $\forall i \in \mathcal{E},\ \forall (j,k) \in \mathcal{T}^2$

$T_j \in \mathbb{N}$ start time of j, $\forall j \in \mathcal{T}$

Mathematical model of the WSRP

Bi-objective Integer Linear Program (ILP) model:

lex max $\big($total working duration, $-$total traveling duration$\big)$

s.t.
- flow constraints
- skill constraints
- occurence constraints
- tasks availability constraints
- employees working hours constraints
- sequencing constraints

$U_{ijk} \in \{0,1\}$ whether or not i goes from j to k, $\forall i \in \mathcal{E},\ \forall (j,k) \in \mathcal{T}^2$

$T_j \in \mathbb{N}$ start time of j, $\forall j \in \mathcal{T}$

Sequencing constraints:

$$\forall\ (j,k) \in \mathcal{T}^2,\ j \neq k,$$

$$T_j + \underbrace{dt_j}_{\substack{\text{duration}\\ \text{of } j}} + \sum_{i \in \mathcal{E}} U_{ijk}\ \underbrace{tr_{jk}}_{\substack{\text{travel}\\ \text{duration}\\ j \to k}} \leq T_k + \left(1 - \sum_{i \in \mathcal{E}} U_{ijk}\right) \underbrace{ubt_j}_{\substack{\text{upper}\\ \text{bound}\\ \text{of } j}}$$

↪ If there is an employee performing j then k,

$$T_j + dt_j + 1 \times tr_{jk} \leq T_k + 0 \times ubt_j$$

↪ Otherwise,

$$T_j + dt_j + 0 \times tr_{jk} \leq T_k + 1 \times ubt_j$$

Sequencing constraints:

$$\forall\ (j,k) \in \mathcal{T}^2,\ j \neq k,$$

$$T_j + \underbrace{dt_j}_{\substack{\text{duration} \\ \text{of } j}} + \sum_{i \in \mathcal{E}} U_{ijk} \underbrace{tr_{jk}}_{\substack{\text{travel} \\ \text{duration} \\ j \to k}} \leq T_k + \left(1 - \sum_{i \in \mathcal{E}} U_{ijk}\right) \underbrace{ubt_j}_{\substack{\text{upper} \\ \text{bound} \\ \text{of } j}}$$

↪ If there is an employee performing j then k,

$$T_j + dt_j + 1 \times tr_{jk} \leq T_k + 0 \times ubt_j$$

↪ Otherwise,

$$T_j + dt_j + 0 \times tr_{jk} \leq T_k + 1 \times ubt_j$$

Mathematical model of the WSRP

Sequencing constraints:

$$\forall\ (j,k) \in \mathcal{T}^2,\ j \neq k,$$

$$T_j + \underbrace{dt_j}_{\substack{\text{duration} \\ \text{of } j}} + \sum_{i \in \mathcal{E}} U_{ijk}\ \underbrace{tr_{jk}}_{\substack{\text{travel} \\ \text{duration} \\ j \to k}} \leq T_k + \left(1 - \sum_{i \in \mathcal{E}} U_{ijk}\right) \underbrace{ubt_j}_{\substack{\text{upper} \\ \text{bound} \\ \text{of } j}}$$

↪ If there is an employee performing j then k,

$$T_j + dt_j + 1 \times tr_{jk} \leq T_k + 0 \times ubt_j$$

↪ Otherwise,

$$T_j + dt_j + 0 \times tr_{jk} \leq T_k + 1 \times ubt_j$$

Mathematical model of the WSRP

Sequencing constraints:

$$\forall\ (j,k) \in \mathcal{T}^2,\ j \neq k,$$

$$T_j + \underbrace{dt_j}_{\substack{\text{duration} \\ \text{of } j}} + \sum_{i \in \mathcal{E}} U_{ijk} \underbrace{tr_{jk}}_{\substack{\text{travel} \\ \text{duration} \\ j \to k}} \leq T_k + \left(1 - \sum_{i \in \mathcal{E}} U_{ijk}\right) \underbrace{ubt_j}_{\substack{\text{upper} \\ \text{bound} \\ \text{of } j}}$$

↪ If there is an employee performing j then k,

$$T_j + dt_j + 1 \times tr_{jk} \leq T_k + 0 \times ubt_j$$

↪ Otherwise,

$$T_j + dt_j + 0 \times tr_{jk} \leq T_k + 1 \times ubt_j$$

Mathematical model of the WSRP

Sequencing constraints:

$$\forall\ (j,k) \in \mathcal{T}^2,\ j \neq k,$$

$$T_j + \underbrace{dt_j}_{\substack{\text{duration} \\ \text{of } j}} + \sum_{i \in \mathcal{E}} U_{ijk} \underbrace{tr_{jk}}_{\substack{\text{travel} \\ \text{duration} \\ j \to k}} \leq T_k + \left(1 - \sum_{i \in \mathcal{E}} U_{ijk}\right) \underbrace{ubt_j}_{\substack{\text{upper} \\ \text{bound} \\ \text{of } j}}$$

↪ If there is an employee performing j then k,

$$T_j + dt_j + 1 \times tr_{jk} \leq T_k + 0 \times ubt_j$$

↪ Otherwise,

$$T_j + dt_j + 0 \times tr_{jk} \leq T_k + 1 \times ubt_j$$

Mathematical model of the WSRP

Bi-objective Integer Linear Program (ILP) model:

lex max $\big($total working duration, $-$total traveling duration$\big)$

s.t.
- flow constraints
- skill constraints
- occurence constraints
- tasks availability constraints
- employees working hours constraints
- sequencing constraints

$U_{ijk} \in \{0,1\}$ whether or not i goes from j to k, $\forall i \in \mathcal{E},\ \forall (j,k) \in \mathcal{T}^2$

$T_j \in \mathbb{N}$ start time of j, $\forall j \in \mathcal{T}$

Plan

1. Introduction
2. Related works
3. Mathematical model of the WSRP
4. Generating counterfactual explanations
5. Conclusion

Plan

1. Introduction
2. Related works
3. Mathematical model of the WSRP
4. Generating counterfactual explanations
   - Explanations requested through questions
   - From questions to mathematical programming
   - From mathematical programming to explanations
5. Conclusion

Explanations requested through questions

List of end-user questions:

We propose **16 template questions** about **various topics**:

- **inserting** a task in an employee's route
  e.g. "How to make ⟨employee i^*⟩ perform ⟨task j^*⟩ in addition to the tasks of their route?";
- **swapping** two tasks outside / inside a route;
- **changing the order** of tasks in a route.

We selected these questions as they satisfy some **assumptions**:

- they focus **on one solution**;
- they mention a "***desideratum***";
- they ask for **counterfactual** explanations ("how to ...").

Explanations requested through questions

List of end-user questions:

We propose **16 template questions** about **various topics**:

- **inserting** a task in an employee's route
  e.g. "How to make ⟨employee i^*⟩ perform ⟨task j^*⟩ in addition to the tasks of their route?";
- **swapping** two tasks outside / inside a route;
- **changing the order** of tasks in a route.

We selected these questions as they satisfy some **assumptions**:

- they focus **on one solution**;
- they mention a "***desideratum***";
- they ask for **counterfactual** explanations ("how to ...").

Explanations requested through questions

List of end-user questions:

We propose **16 template questions** about **various topics**:

- **inserting** a task in an employee's route
  e.g. "How to make ⟨employee i^*⟩ perform ⟨task j^*⟩ in addition to the tasks of their route?";
- **swapping** two tasks outside / inside a route;
- **changing the order** of tasks in a route.

We selected these questions as they satisfy some **assumptions**:

- they focus **on one solution**;
- they mention a "***desideratum***";
- they ask for **counterfactual** explanations ("how to ...").

Explanations requested through questions

List of end-user questions:

We propose **16 template questions** about **various topics**:

- **inserting** a task in an employee's route
  e.g. "How to make ⟨employee i^*⟩ perform ⟨task j^*⟩ in addition to the tasks of their route?";
- **swapping** two tasks outside / inside a route;
- **changing the order** of tasks in a route.

We selected these questions as they satisfy some **assumptions**:

- they focus **on one solution**;
- they mention a "***desideratum***";
- they ask for **counterfactual** explanations ("how to ...").

Explanations requested through questions

List of end-user questions:

We propose **16 template questions** about **various topics**:

- **inserting** a task in an employee's route
  e.g. "How to make ⟨employee i^*⟩ perform ⟨task j^*⟩ in addition to the tasks of their route?";
- **swapping** two tasks outside / inside a route;
- **changing the order** of tasks in a route.

We selected these questions as they satisfy some **assumptions**:

- they focus **on one solution**;
- they mention a "***desideratum***";
- they ask for **counterfactual** explanations ("how to ...").

Explanations requested through questions

List of end-user questions:

We propose **16 template questions** about **various topics**:

- **inserting** a task in an employee's route
  e.g. "How to make ⟨employee i^*⟩ perform ⟨task j^*⟩ in addition to the tasks of their route?";
- **swapping** two tasks outside / inside a route;
- **changing the order** of tasks in a route.

We selected these questions as they satisfy some **assumptions**:

- they focus **on one solution**;
- they mention a "***desideratum***";
- they ask for **counterfactual** explanations ("how to ...").

Explanations requested through questions

List of end-user questions:

We propose **16 template questions** about **various topics**:

- **inserting** a task in an employee's route
  e.g. "How to make ⟨employee i^*⟩ perform ⟨task j^*⟩ in addition to the tasks of their route?";
- **swapping** two tasks outside / inside a route;
- **changing the order** of tasks in a route.

We selected these questions as they satisfy some **assumptions**:

- they focus **on one solution**;
- they mention a "***desideratum***";
- they ask for **counterfactual** explanations ("how to ...").

Explanations requested through questions

List of end-user questions:

We propose **16 template questions** about **various topics**:

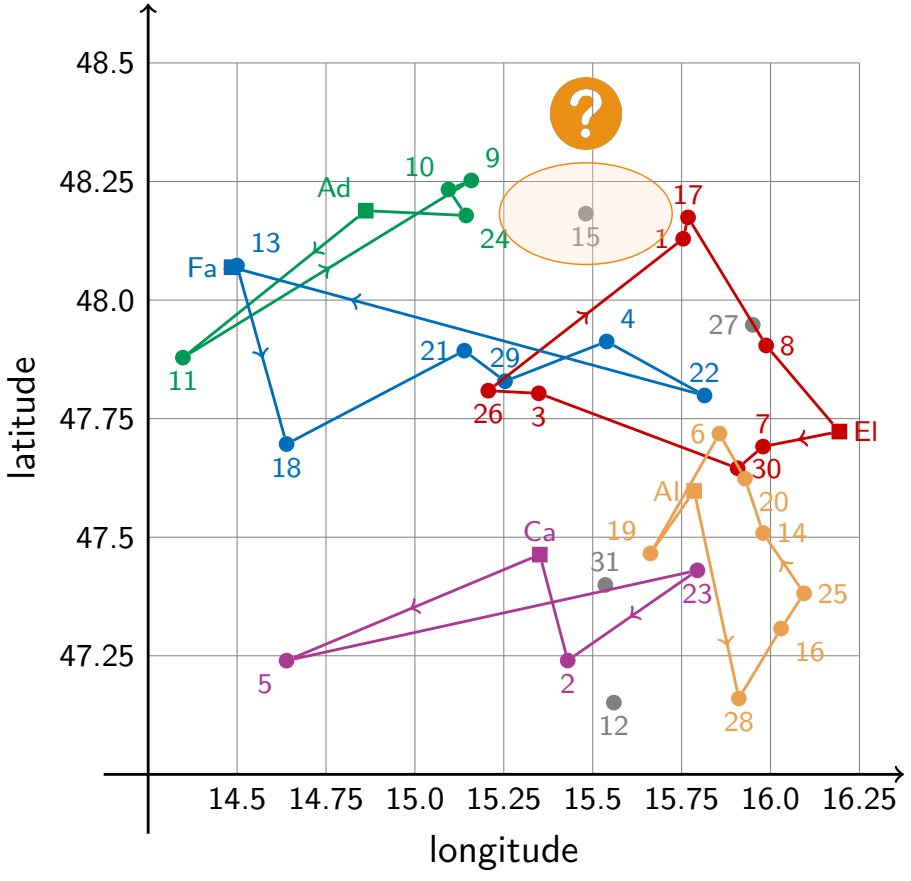
- **inserting** a task in an employee's route
  e.g. "How to make ⟨employee i^*⟩ perform ⟨task j^*⟩ in addition to the tasks of their route?";
- **swapping** two tasks outside / inside a route;
- **changing the order** of tasks in a route.

We selected these questions as they satisfy some **assumptions**:

- they focus **on one solution**;
- they mention a "***desideratum***";
- they ask for **counterfactual** explanations ("how to ...").

Reminder | Our use case - (3) Non-expert end-user

A problematic situation for a planner:



"How to make Ellen perform task 15 in addition to the tasks of her route?"

↪ If no explanations, then black box feeling...

Plan

1. Introduction
2. Related works
3. Mathematical model of the WSRP
4. Generating counterfactual explanations
   - Explanations requested through questions
   - From questions to mathematical programming
   - From mathematical programming to explanations
5. Conclusion

From questions to mathematical programming

General principle:

Each question q is **mapped** to a **multi-obj. ILP model** which:

- aims at finding **how to alter the instance parameters** s.t.:
  1. we obtain a solution satisfying the *desideratum*;
  2. we minimize the **alterations** (**magnitude** and **number**);
  3. we optimize the solution quality.
- is based on the WSRP model with **adaptations**:
  - focus on only one employee's route;
  - new decision variables, changes in constraints, changes in the objective functions, *etc.*

From questions to mathematical programming

General principle:

Each question q is **mapped** to a **multi-obj. ILP model** which:

- aims at finding **how to alter the instance parameters** s.t.:
  1. we obtain a solution satisfying the *desideratum*;
  2. we minimize the **alterations** (**magnitude** and **number**);
  3. we optimize the solution quality.
- is based on the WSRP model with **adaptations**:
  - focus on only one employee's route;
  - new decision variables, changes in constraints, changes in the objective functions, *etc.*

From questions to mathematical programming

General principle:

Each question q is **mapped** to a **multi-obj. ILP model** which:

- aims at finding **how to alter the instance parameters** s.t.:
  1. we obtain a solution satisfying the *desideratum*;
  2. we minimize the **alterations** (**magnitude** and **number**);
  3. we optimize the solution quality.
- is based on the WSRP model with **adaptations**:
  - focus on only one employee's route;
  - new decision variables, changes in constraints, changes in the objective functions, *etc.*

From questions to mathematical programming

General principle:

Each question q is **mapped** to a **multi-obj. ILP model** which:

- aims at finding **how to alter the instance parameters** s.t.:
  1. we obtain a solution satisfying the *desideratum*;
  2. we minimize the **alterations** (**magnitude** and **number**);
  3. we optimize the solution quality.
- is based on the WSRP model with **adaptations**:
  - focus on only one employee's route;
  - new decision variables, changes in constraints, changes in the objective functions, *etc.*

From questions to mathematical programming

General principle:

Each question q is **mapped** to a **multi-obj. ILP model** which:

- aims at finding **how to alter the instance parameters** s.t.:
  1. we obtain a solution satisfying the *desideratum*;
  2. we minimize the **alterations** (**magnitude** and **number**);
  3. we optimize the solution quality.
- is based on the WSRP model with **adaptations**:
  - focus on only one employee's route;
  - new decision variables, changes in constraints, changes in the objective functions, *etc.*

From questions to mathematical programming

General principle:

Each question q is **mapped** to a **multi-obj. ILP model** which:

- aims at finding **how to alter the instance parameters** s.t.:
  1. we obtain a solution satisfying the *desideratum*;
  2. we minimize the **alterations** (**magnitude** and **number**);
  3. we optimize the solution quality.
- is based on the WSRP model with **adaptations**:
  - focus on only one employee's route;
  - new decision variables, changes in constraints, changes in the objective functions, *etc.*

From questions to mathematical programming

General principle:

Each question q is **mapped** to a **multi-obj. ILP model** which:

- aims at finding **how to alter the instance parameters** s.t.:
  1. we obtain a solution satisfying the *desideratum*;
  2. we minimize the **alterations** (**magnitude** and **number**);
  3. we optimize the solution quality.
- is based on the WSRP model with **adaptations**:
  - focus on only one employee's route;
  - new decision variables, changes in constraints, changes in the objective functions, *etc.*

From questions to mathematical programming

General principle:

Each question q is **mapped** to a **multi-obj. ILP model** which:

- aims at finding **how to alter the instance parameters** s.t.:
  1. we obtain a solution satisfying the *desideratum*;
  2. we minimize the **alterations** (**magnitude** and **number**);
  3. we optimize the solution quality.
- is based on the WSRP model with **adaptations**:
  - focus on only one employee's route;
  - new decision variables, changes in constraints, changes in the objective functions, *etc.*

Multi-objective ILP model on an example of question:

Let q be a question based on "How to make ⟨employee i^*⟩ perform ⟨task j^*⟩ in addition to the tasks of their route?"

Some preliminary remarks:

- We can work with a **reduced set of tasks** $\mathcal{T}^* = \{j \text{ performed by } i^* \text{ in solution}\} \cup \{j^*\}$.
- We assume that the user wants to alter only **task time parameters** (dt_j, lbt_j, ubt_j) within some acceptable ranges.
- We introduce new **decision variables** including:
  - ΔDT_j, decrease of dt_j, for all $j \in \mathcal{T}^*$;
  - ΔT_{max}, greatest value of time alterations;
  - T_{j^*} split in two and replaced by $T_{j^*}^{lb}$ and $T_{j^*}^{ub}$: earliest start time of j^* when satisfying time constraints before, latest start time of j^* when satisfying time constraints after.

From questions to mathematical programming

Multi-objective ILP model on an example of question:

Let q be a question based on "How to make ⟨employee i^*⟩ perform ⟨task j^*⟩ in addition to the tasks of their route?"

Some preliminary remarks:

- We can work with a **reduced set of tasks** $\mathcal{T}^* = \{j \text{ performed by } i^* \text{ in solution}\} \cup \{j^*\}$.
- We assume that the user wants to alter only **task time parameters** (dt_j, lbt_j, ubt_j) within some acceptable ranges.
- We introduce new **decision variables** including:
  - ΔDT_j, decrease of dt_j, for all $j \in \mathcal{T}^*$;
  - ΔT_{max}, greatest value of time alterations;
  - T_{j^*} split in two and replaced by $T_{j^*}^{lb}$ and $T_{j^*}^{ub}$: earliest start time of j^* when satisfying time constraints before, latest start time of j^* when satisfying time constraints after.

Multi-objective ILP model on an example of question:

Let q be a question based on "How to make ⟨employee i^*⟩ perform ⟨task j^*⟩ in addition to the tasks of their route?"

Some preliminary remarks:

- We can work with a **reduced set of tasks** $\mathcal{T}^* = \{j \text{ performed by } i^* \text{ in solution}\} \cup \{j^*\}$.
- We assume that the user wants to alter only **task time parameters** (dt_j, lbt_j, ubt_j) within some acceptable ranges.
- We introduce new **decision variables** including:
  - ΔDT_j, decrease of dt_j, for all $j \in \mathcal{T}^*$;
  - ΔT_{max}, greatest value of time alterations;
  - T_{j^*} split in two and replaced by $T_{j^*}^{lb}$ and $T_{j^*}^{ub}$: earliest start time of j^* when satisfying time constraints before, latest start time of j^* when satisfying time constraints after.

Multi-objective ILP model on an example of question:

Let q be a question based on "How to make ⟨employee i^*⟩ perform ⟨task j^*⟩ in addition to the tasks of their route?"

Some preliminary remarks:

- We can work with a **reduced set of tasks** $\mathcal{T}^* = \{j \text{ performed by } i^* \text{ in solution}\} \cup \{j^*\}$.
- We assume that the user wants to alter only **task time parameters** (dt_j, lbt_j, ubt_j) within some acceptable ranges.
- We introduce new **decision variables** including:
  - ΔDT_j, decrease of dt_j, for all $j \in \mathcal{T}^*$;
  - ΔT_{max}, greatest value of time alterations;
  - T_{j^*} split in two and replaced by $T_{j^*}^{lb}$ and $T_{j^*}^{ub}$: earliest start time of j^* when satisfying time constraints before, latest start time of j^* when satisfying time constraints after.

Multi-objective ILP model on an example of question:

Let q be a question based on "How to make ⟨employee i^*⟩ perform ⟨task j^*⟩ in addition to the tasks of their route?"

Some preliminary remarks:

- We can work with a **reduced set of tasks** $\mathcal{T}^* = \{j \text{ performed by } i^* \text{ in solution}\} \cup \{j^*\}$.
- We assume that the user wants to alter only **task time parameters** (dt_j, lbt_j, ubt_j) within some acceptable ranges.
- We introduce new **decision variables** including:
  - ΔDT_j, decrease of dt_j, for all $j \in \mathcal{T}^*$;
  - ΔT_{max}, greatest value of time alterations;
  - T_{j^*} split in two and replaced by $T_{j^*}^{lb}$ and $T_{j^*}^{ub}$:
    earliest start time of j^* when satisfying time constraints before,
    latest start time of j^* when satisfying time constraints after.

Multi-objective ILP model on an example of question:

Let q be a question based on "How to make ⟨employee i^*⟩ perform ⟨task j^*⟩ in addition to the tasks of their route?"

Some preliminary remarks:

- We can work with a **reduced set of tasks** $\mathcal{T}^* = \{j \text{ performed by } i^* \text{ in solution}\} \cup \{j^*\}$.
- We assume that the user wants to alter only **task time parameters** (dt_j, lbt_j, ubt_j) within some acceptable ranges.
- We introduce new **decision variables** including:
  - ΔDT_j, decrease of dt_j, for all $j \in \mathcal{T}^*$;
  - ΔT_{max}, greatest value of time alterations;
  - T_{j^*} split in two and replaced by $T_{j^*}^{lb}$ and $T_{j^*}^{ub}$: earliest start time of j^* when satisfying time constraints before, latest start time of j^* when satisfying time constraints after.

Multi-objective ILP model on an example of question:

Let q be a question based on "How to make ⟨employee i^*⟩ perform ⟨task j^*⟩ in addition to the tasks of their route?"

Some preliminary remarks:

- We can work with a **reduced set of tasks** $\mathcal{T}^* = \{j \text{ performed by } i^* \text{ in solution}\} \cup \{j^*\}$.
- We assume that the user wants to alter only **task time parameters** (dt_j, lbt_j, ubt_j) within some acceptable ranges.
- We introduce new **decision variables** including:
  - ΔDT_j, decrease of dt_j, for all $j \in \mathcal{T}^*$;
  - ΔT_{max}, greatest value of time alterations;
  - T_{j^*} split in two and replaced by $T_{j^*}^{lb}$ and $T_{j^*}^{ub}$: earliest start time of j^* when satisfying time constraints before, latest start time of j^* when satisfying time constraints after.

Multi-objective ILP model on an example of question:

Let q be a question based on "How to make ⟨employee i^*⟩ perform ⟨task j^*⟩ in addition to the tasks of their route?"

Some preliminary remarks:

- We can work with a **reduced set of tasks** $\mathcal{T}^* = \{j \text{ performed by } i^* \text{ in solution}\} \cup \{j^*\}$.
- We assume that the user wants to alter only **task time parameters** (dt_j, lbt_j, ubt_j) within some acceptable ranges.
- We introduce new **decision variables** including:
  - ΔDT_j, decrease of dt_j, for all $j \in \mathcal{T}^*$;
  - ΔT_{max}, greatest value of time alterations;
  - T_{j^*} split in two and replaced by $T_{j^*}^{lb}$ and $T_{j^*}^{ub}$:
    earliest start time of j^* when satisfying time constraints before,
    latest start time of j^* when satisfying time constraints after.

Multi-objective ILP model on an example of question:

Let q be a question based on "How to make ⟨employee i^*⟩ perform ⟨task j^*⟩ in addition to the tasks of their route?"

Some preliminary remarks:

- We can work with a **reduced set of tasks** $\mathcal{T}^* = \{j \text{ performed by } i^* \text{ in solution}\} \cup \{j^*\}$.
- We assume that the user wants to alter only **task time parameters** (dt_j, lbt_j, ubt_j) within some acceptable ranges.
- We introduce new **decision variables** including:
  - ΔDT_j, decrease of dt_j, for all $j \in \mathcal{T}^*$;
  - ΔT_{max}, greatest value of time alterations;
  - T_{j^*} split in two and replaced by $T_{j^*}^{lb}$ and $T_{j^*}^{ub}$: earliest start time of j^* when satisfying time constraints before, latest start time of j^* when satisfying time constraints after.

Multi-objective ILP model on an example of question:

lex min $\Big(T^{lb}_{j^*} - T^{ub}_{j^*}, \quad \sum_{j \in \mathcal{T}^*} \Delta DT_j, \quad \Delta T_{max}, \quad$ nb of alterations, total traveling duration$\Big)$

s.t.
- adjusted flow constraints
- adjusted occurence constraints
- adjusted tasks availability constraints
- adjusted working hours constraints
- adjusted sequencing constraints
- constraints on altering variables like ΔDT_j

$U_{i^*jk} \in \{0,1\}$	whether or not i^* goes from j to k,	$\forall\ (j,k) \in (\mathcal{T}^*)^2$
$T_j \in \mathbb{N}$	start time of j	$\forall\ j \in \mathcal{T}^*$
$T^{ub}_{j^*}, T^{lb}_{j^*} \in \mathbb{N}$	bounds on the start time of j^*	
$\Delta DT_j \in \mathbb{N}$	decrease of the duration of j	$\forall\ j \in \mathcal{T}^*$
...		
$\Delta T_{max} \in \mathbb{N}$	greatest value of time alterations	

Multi-objective ILP model on an example of question:

$$\text{lex min} \quad \Big(T^{lb}_{j^*} - T^{ub}_{j^*}, \quad \sum_{j \in \mathcal{T}^*} \Delta DT_j, \quad \Delta T_{max}, \quad \text{nb of alterations}, \quad \text{total traveling duration} \Big)$$

s.t.
- adjusted flow constraints
- adjusted occurence constraints
- adjusted tasks availability constraints
- adjusted working hours constraints
- adjusted sequencing constraints
- constraints on altering variables like ΔDT_j

$U_{i^*jk} \in \{0,1\}$	whether or not i^* goes from j to k,	$\forall\ (j,k) \in (\mathcal{T}^*)^2$
$T_j \in \mathbb{N}$	start time of j	$\forall\ j \in \mathcal{T}^*$
$T^{ub}_{j^*}, T^{lb}_{j^*} \in \mathbb{N}$	bounds on the start time of j^*	
$\Delta DT_j \in \mathbb{N}$	decrease of the duration of j	$\forall\ j \in \mathcal{T}^*$
...		
$\Delta T_{max} \in \mathbb{N}$	greatest value of time alterations	

Multi-objective ILP model on an example of question:

lex min $\Big(T^{lb}_{j^*} - T^{ub}_{j^*}, \quad \sum_{j \in \mathcal{T}^*} \Delta DT_j, \quad \Delta T_{max}, \quad$ nb of alterations, total traveling duration$\Big)$

s.t.
- adjusted flow constraints
- adjusted occurence constraints
- adjusted tasks availability constraints
- adjusted working hours constraints
- adjusted sequencing constraints
- constraints on altering variables like ΔDT_j

$U_{i^*jk} \in \{0,1\}$ whether or not i^* goes from j to k, $\forall\ (j,k) \in (\mathcal{T}^*)^2$
$T_j \in \mathbb{N}$ start time of j $\forall\ j \in \mathcal{T}^*$
$T^{ub}_{j^*}, T^{lb}_{j^*} \in \mathbb{N}$ bounds on the start time of j^*
$\Delta DT_j \in \mathbb{N}$ decrease of the duration of j $\forall\ j \in \mathcal{T}^*$
...
$\Delta T_{max} \in \mathbb{N}$ greatest value of time alterations

Multi-objective ILP model on an example of question:

lex min $\Big(T^{lb}_{j^*} - T^{ub}_{j^*}, \quad \sum_{j \in \mathcal{T}^*} \Delta DT_j, \quad \Delta T_{max}, \quad$ nb of alterations, total traveling duration$\Big)$

s.t.
- adjusted flow constraints
- adjusted occurence constraints
- adjusted tasks availability constraints
- adjusted working hours constraints
- adjusted sequencing constraints
- constraints on altering variables like ΔDT_j

$U_{i^*jk} \in \{0,1\}$	whether or not i^* goes from j to k,	$\forall\ (j,k) \in (\mathcal{T}^*)^2$
$T_j \in \mathbb{N}$	start time of j	$\forall\ j \in \mathcal{T}^*$
$T^{ub}_{j^*}, T^{lb}_{j^*} \in \mathbb{N}$	bounds on the start time of j^*	
$\Delta DT_j \in \mathbb{N}$	decrease of the duration of j	$\forall\ j \in \mathcal{T}^*$
...		
$\Delta T_{max} \in \mathbb{N}$	greatest value of time alterations	

Multi-objective ILP model on an example of question:

lex min $\Big(T^{lb}_{j^*} - T^{ub}_{j^*}, \quad \sum_{j \in \mathcal{T}^*} \Delta DT_j, \quad \Delta T_{max}, \quad$ nb of alterations, total traveling duration$\Big)$

s.t.
- adjusted flow constraints
- adjusted occurence constraints
- adjusted tasks availability constraints
- adjusted working hours constraints
- adjusted sequencing constraints
- constraints on altering variables like ΔDT_j

$U_{i^*jk} \in \{0,1\}$	whether or not i^* goes from j to k,	$\forall\ (j,k) \in (\mathcal{T}^*)^2$
$T_j \in \mathbb{N}$	start time of j	$\forall\ j \in \mathcal{T}^*$
$T^{ub}_{j^*}, T^{lb}_{j^*} \in \mathbb{N}$	bounds on the start time of j^*	
$\Delta DT_j \in \mathbb{N}$	decrease of the duration of j	$\forall\ j \in \mathcal{T}^*$
...		
$\Delta T_{max} \in \mathbb{N}$	greatest value of time alterations	

Sequencing constraints:

$$\forall\ (j,k) \in \mathcal{T}^2,\ j \neq k,$$

$$T_j + \underbrace{dt_j}_{\substack{\text{duration}\\ \text{of } j}} + \sum_{i \in \mathcal{E}} U_{ijk}\ \underbrace{tr_{jk}}_{\substack{\text{travel}\\ \text{duration}\\ j \to k}} \leq T_k + \left(1 - \sum_{i \in \mathcal{E}} U_{ijk}\right) \underbrace{ubt_j}_{\substack{\text{upper}\\ \text{bound}\\ \text{of } j}}$$

↪ If there is an employee performing j then k,

$$T_j + dt_j + 1 \times tr_{jk} \leq T_k + 0 \times ubt_j$$

↪ Otherwise,

$$T_j + dt_j + 0 \times tr_{jk} \leq T_k + 1 \times ubt_j$$

Sequencing constraints with altering variables:

3 groups of constraints instead of 1:

- $\forall\ (j,k) \in (\mathcal{T}^* \setminus \{j^*\})^2,\ j \neq k,$

$$T_j + dt_j - \Delta DT_j + U_{i^*jk}\ tr_{jk} \leq T_k + \left(1 - U_{i^*jk}\right) ubt_j$$

- $\forall\ k \in \mathcal{T}^* \setminus \{j^*\},$

$$T_{j^*}^{ub} + dt_{j^*} - \Delta DT_{j^*} + U_{i^*j^*k}\ tr_{j^*k} \leq T_k + \left(1 - U_{i^*j^*k}\right) ubt_{j^*}$$

- $\forall\ j \in \mathcal{T}^* \setminus \{j^*\},$

$$T_j + dt_j - \Delta DT_j + U_{i^*jj^*}\ tr_{jj^*} \leq T_{j^*}^{lb} + \left(1 - U_{i^*jj^*}\right) ubt_j$$

Sequencing constraints with altering variables:

3 groups of constraints instead of 1:

- $\forall\ (j,k) \in (\mathcal{T}^* \setminus \{j^*\})^2,\ j \neq k,$
$$T_j + dt_j - \Delta DT_j + U_{i^*jk}\ tr_{jk} \leq T_k + \Big(1 - U_{i^*jk}\Big) ubt_j$$
- $\forall\ k \in \mathcal{T}^* \setminus \{j^*\},$
$$T^{ub}_{j^*} + dt_{j^*} - \Delta DT_{j^*} + U_{i^*j^*k}\ tr_{j^*k} \leq T_k + \Big(1 - U_{i^*j^*k}\Big) ubt_{j^*}$$
- $\forall\ j \in \mathcal{T}^* \setminus \{j^*\},$
$$T_j + dt_j - \Delta DT_j + U_{i^*jj^*}\ tr_{jj^*} \leq T^{lb}_{j^*} + \Big(1 - U_{i^*jj^*}\Big) ubt_j$$

Sequencing constraints with altering variables:

3 groups of constraints instead of 1:

- $\forall\ (j,k) \in (\mathcal{T}^* \setminus \{j^*\})^2,\ j \neq k,$

$$T_j + dt_j - \Delta DT_j + U_{i^*jk}\ tr_{jk} \leq T_k + \Big(1 - U_{i^*jk}\Big) ubt_j$$

- $\forall\ k \in \mathcal{T}^* \setminus \{j^*\},$

$$T^{ub}_{j^*} + dt_{j^*} - \Delta DT_{j^*} + U_{i^*j^*k}\ tr_{j^*k} \leq T_k + \Big(1 - U_{i^*j^*k}\Big) ubt_{j^*}$$

- $\forall\ j \in \mathcal{T}^* \setminus \{j^*\},$

$$T_j + dt_j - \Delta DT_j + U_{i^*jj^*}\ tr_{jj^*} \leq T^{lb}_{j^*} + \Big(1 - U_{i^*jj^*}\Big) ubt_j$$

Sequencing constraints with altering variables:

3 groups of constraints instead of 1:

- $\forall\ (j,k) \in (\mathcal{T}^* \setminus \{j^*\})^2,\ j \neq k,$

$$T_j + dt_j - \Delta DT_j + U_{i^*jk}\ tr_{jk} \leq T_k + \Big(1 - U_{i^*jk}\Big) ubt_j$$

- $\forall\ k \in \mathcal{T}^* \setminus \{j^*\},$

$$T^{ub}_{j^*} + dt_{j^*} - \Delta DT_{j^*} + U_{i^*j^*k}\ tr_{j^*k} \leq T_k + \Big(1 - U_{i^*j^*k}\Big) ubt_{j^*}$$

- $\forall\ j \in \mathcal{T}^* \setminus \{j^*\},$

$$T_j + dt_j - \Delta DT_j + U_{i^*jj^*}\ tr_{jj^*} \leq T^{lb}_{j^*} + \Big(1 - U_{i^*jj^*}\Big) ubt_j$$

From questions to mathematical programming

Multi-objective ILP model on an example of question:

lex min $\big(T^{lb}_{j^*} - T^{ub}_{j^*}, \quad \sum_{j \in \mathcal{T}^*} \Delta DT_j, \quad \Delta T_{max}, \quad$ nb of alterations, total traveling duration$\big)$

s.t.
- adjusted flow constraints
- adjusted occurence constraints
- adjusted tasks availability constraints
- adjusted working hours constraints
- adjusted sequencing constraints
- constraints on altering variables like ΔDT_j

$U_{i^*jk} \in \{0,1\}$	whether or not i^* goes from j to k,	$\forall\ (j,k) \in (\mathcal{T}^*)^2$
$T_j \in \mathbb{N}$	start time of j	$\forall\ j \in \mathcal{T}^*$
$T^{ub}_{j^*}, T^{lb}_{j^*} \in \mathbb{N}$	bounds on the start time of j^*	
$\Delta DT_j \in \mathbb{N}$	decrease of the duration of j	$\forall\ j \in \mathcal{T}^*$
...		
$\Delta T_{max} \in \mathbb{N}$	greatest value of time alterations	

Plan

1. Introduction
2. Related works
3. Mathematical model of the WSRP
4. Generating counterfactual explanations
   - Explanations requested through questions
   - From questions to mathematical programming
   - From mathematical programming to explanations
5. Conclusion

From mathematical programming to explanations

3 cases of explanations:

Assume that the multi-objective ILP model has been solved.

1. $T_{j^*}^{lb} - T_{j^*}^{ub} = 0$ and $\Delta T_{max} > 0$
   ↪ The *desideratum* is obtained **thanks to alterations**.

2. $T_{j^*}^{lb} - T_{j^*}^{ub} = 0$ and $\Delta T_{max} = 0$
   ↪ The *desideratum* is obtained **without any alterations**.

3. $T_{j^*}^{lb} - T_{j^*}^{ub} > 0$
   ↪ Alterations ranges are **not enough** to get the *desideratum*.

From mathematical programming to explanations

3 cases of explanations:

Assume that the multi-objective ILP model has been solved.

1. $T_{j^*}^{lb} - T_{j^*}^{ub} = 0$ and $\Delta T_{max} > 0$
   ↪ The *desideratum* is obtained **thanks to alterations**.
2. $T_{j^*}^{lb} - T_{j^*}^{ub} = 0$ and $\Delta T_{max} = 0$
   ↪ The *desideratum* is obtained **without any alterations**.
3. $T_{j^*}^{lb} - T_{j^*}^{ub} > 0$
   ↪ Alterations ranges are **not enough** to get the *desideratum*.

3 cases of explanations:

Assume that the multi-objective ILP model has been solved.

1. $T_{j^*}^{lb} - T_{j^*}^{ub} = 0$ and $\Delta T_{max} > 0$
   ↪ The *desideratum* is obtained **thanks to alterations**.

2. $T_{j^*}^{lb} - T_{j^*}^{ub} = 0$ and $\Delta T_{max} = 0$
   ↪ The *desideratum* is obtained **without any alterations**.

3. $T_{j^*}^{lb} - T_{j^*}^{ub} > 0$
   ↪ Alterations ranges are **not enough** to get the *desideratum*.

3 cases of explanations:

Assume that the multi-objective ILP model has been solved.

1. $T_{j^*}^{lb} - T_{j^*}^{ub} = 0$ and $\Delta T_{max} > 0$
   ↪ The *desideratum* is obtained **thanks to alterations**.
2. $T_{j^*}^{lb} - T_{j^*}^{ub} = 0$ and $\Delta T_{max} = 0$
   ↪ The *desideratum* is obtained **without any alterations**.
3. $T_{j^*}^{lb} - T_{j^*}^{ub} > 0$
   ↪ Alterations ranges are **not enough** to get the *desideratum*.

3 cases of explanations:

Assume that the multi-objective ILP model has been solved.

1. $T_{j^*}^{lb} - T_{j^*}^{ub} = 0$ and $\Delta T_{max} > 0$
   ↪ The *desideratum* is obtained **thanks to alterations**.

2. $T_{j^*}^{lb} - T_{j^*}^{ub} = 0$ and $\Delta T_{max} = 0$
   ↪ The *desideratum* is obtained **without any alterations**.

3. $T_{j^*}^{lb} - T_{j^*}^{ub} > 0$
   ↪ Alterations ranges are **not enough** to get the *desideratum*.

From mathematical programming to explanations

Explanation text in case 1.:

"By ⟨applying the alterations of the instance parameters given by positive values of altering variables like ΔDT_j⟩,
⟨the *desideratum*⟩ would be possible;
in this case, the solution would be ⟨the one deduced from the ILP optimal result⟩."

From mathematical programming to explanations

Explanation text in case 1.:

"By ⟨applying the alterations of the instance parameters given by positive values of altering variables like ΔDT_j⟩,
⟨the *desideratum*⟩ would be possible;
in this case, the solution would be ⟨the one deduced from the ILP optimal result⟩."

From mathematical programming to explanations

Explanation text in case 1.:

"By ⟨applying the alterations of the instance parameters given by positive values of altering variables like ΔDT_j⟩,
⟨the *desideratum*⟩ would be possible;
in this case, the solution would be ⟨the one deduced from the ILP optimal result⟩."

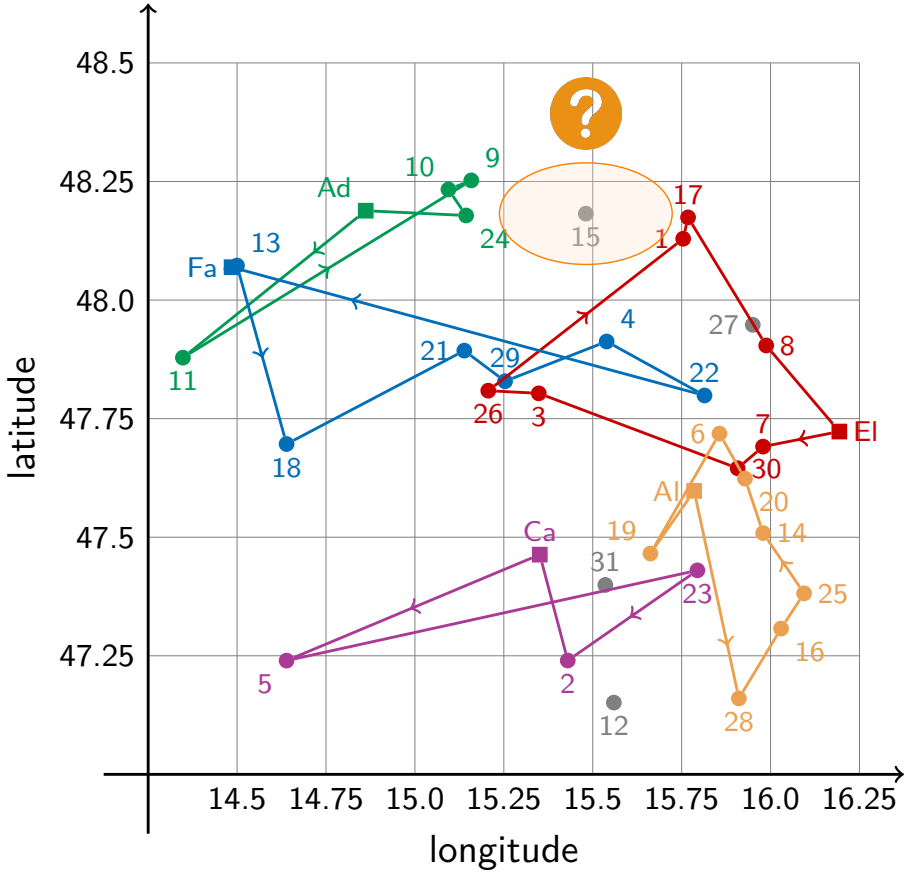
From mathematical programming to explanations

Explanation text in case 1.:

"By ⟨applying the alterations of the instance parameters given by positive values of altering variables like ΔDT_j⟩,
⟨the *desideratum*⟩ would be possible;
in this case, the solution would be ⟨the one deduced from the ILP optimal result⟩."

Reminder | Our use case - (3) Non-expert end-user

A problematic situation for a planner:



"How to make Ellen perform task 15 in addition to the tasks of her route?"

↪ If no explanations, then black box feeling...

From mathematical programming to explanations

Explanation text in case 1.:

"By ⟨applying the alterations of the instance parameters given by positive values of altering variables like ΔDT_j⟩,
⟨the *desideratum*⟩ would be possible;
in this case, the solution would be ⟨the one deduced from the ILP optimal result⟩."

Explanation text in case 1.:

"By ⟨applying the alterations of the instance parameters given by positive values of altering variables like ΔDT_j⟩,
⟨the *desideratum*⟩ would be possible;
in this case, the solution would be ⟨the one deduced from the ILP optimal result⟩."

Explanation text in case 1.:

"By changing the opening time of task 17 to 12:29PM (instead of 12:30PM in the current input data),
⟨the *desideratum*⟩ would be possible;
in this case, the solution would be ⟨the one deduced from the ILP optimal result⟩."

From mathematical programming to explanations

Explanation text in case 1.:

"By changing the opening time of task 17 to 12:29PM (instead of 12:30PM in the current input data),
⟨the *desideratum*⟩ would be possible;
in this case, the solution would be ⟨the one deduced from the ILP optimal result⟩."

Explanation text in case 1.:

"By changing the opening time of task 17 to 12:29PM (instead of 12:30PM in the current input data),
making Ellen perform the task 15 in addition to her already-performed tasks would be possible,
in this case, the solution would be ⟨the one deduced from the ILP optimal result⟩."

Explanation text in case 1.:

"By changing the opening time of task 17 to 12:29PM (instead of 12:30PM in the current input data),
making Ellen perform the task 15 in addition to her already-performed tasks would be possible,
in this case, the solution would be ⟨the one deduced from the ILP optimal result⟩."

Explanation text in case 1.:

"By changing the opening time of task 17 to 12:29PM (instead of 12:30PM in the current input data),
making Ellen perform the task 15 in addition to her already-performed tasks would be possible,
in this case, the solution would be the one obtained by changing Ellen's sequence of performed tasks to [30, 7, 8, 1, 17, 15, 26, 3]."

Plan

1. Introduction
2. Related works
3. Mathematical model of the WSRP
4. Generating counterfactual explanations
5. Conclusion

Conclusion

Achieved work in this article:

Approach for generating **counterfactual explanations** that:

- is thought for an end-user of a system solving a **WSRP**;
- starts from user **questions** about various *desiderata*;
- is based on **mathematical programming**;
- ends on explanations given as **texts**.

Related work:

Approach for generating **contrastive explanations**:

[Lerouge et al., 2022] Explaining solutions stemming from optimization systems solving the Workforce Scheduling and Routing Problem to their end-users (Working paper).

Conclusion

Achieved work in this article:

Approach for generating **counterfactual explanations** that:

- is thought for an end-user of a system solving a **WSRP**;
- starts from user **questions** about various *desiderata*;
- is based on **mathematical programming**;
- ends on explanations given as **texts**.

Related work:

Approach for generating **contrastive explanations**:

> [Lerouge et al., 2022] Explaining solutions stemming from optimization systems solving the Workforce Scheduling and Routing Problem to their end-users (Working paper).

Conclusion

Work in progress:

- Perform an exhaustive study for assessing **computational efficiency**.
- Evaluate how explanations influence **end-users' trust**.
- Design an explanation system handling **various types of explanations** (contrastive, scenario and counterfactual).

Perspectives:

How much **generic** is our **approach**? Can we transpose it to other optimization problems?

Conclusion

Work in progress:

- Perform an exhaustive study for assessing **computational efficiency**.
- Evaluate how explanations influence **end-users' trust**.
- Design an explanation system handling **various types of explanations** (contrastive, scenario and counterfactual).

Perspectives:

How much **generic** is our **approach**? Can we transpose it to other optimization problems?

Conclusion

Work in progress:

- Perform an exhaustive study for assessing **computational efficiency**.
- Evaluate how explanations influence **end-users' trust**.
- Design an explanation system handling **various types of explanations** (contrastive, scenario and counterfactual).

Perspectives:

How much **generic** is our **approach**? Can we transpose it to other optimization problems?

Conclusion

Work in progress:

- Perform an exhaustive study for assessing **computational efficiency**.
- Evaluate how explanations influence **end-users' trust**.
- Design an explanation system handling **various types of explanations** (contrastive, scenario and counterfactual).

Perspectives:

How much **generic** is our **approach**? Can we transpose it to other optimization problems?

Conclusion

Work in progress:

- Perform an exhaustive study for assessing **computational efficiency**.
- Evaluate how explanations influence **end-users' trust**.
- Design an explanation system handling **various types of explanations** (contrastive, scenario and counterfactual).

Perspectives:

How much **generic** is our **approach**? Can we transpose it to other optimization problems?

Conclusion

Thank you for your attention!

References I

[Barredo Arrieta et al., 2020] Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., and Herrera, F. (2020).
Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI.
Information fusion, 58:82–115.

[Chakraborti et al., 2020] **Chakraborti, T., Sreedharan, S., and Kambhampati, S. (2020).**
The emerging landscape of Explainable AI Planning and Decision Making.

References II

In *Proceedings of the twenty-ninth International Joint Conference on Artificial Intelligence*, pages 4803–4811. IJCAI Organization.

[Junker, 2004] **Junker, U. (2004).**
Quickxplain: Preferred explanations and relaxations for over-constrained problems.
In *Proceedings of the ninetieth Association for the Advancement of Artificial Intelligence Conference on Artificial Intelligence*, page 167–172. AAAI Press.

References III

[Korikov et al., 2021] **Korikov, A., Shleyfman, A., and Beck, C. (2021).**
Counterfactual explanations for optimization-based decisions in the context of the GDPR.
In *Proceedings of the thirtieth International Joint Conference on Artificial Intelligence*, pages 4097 – 4103. IJCAI Organization.

[Krarup et al., 2021] **Krarup, B., Krivic, S., Magazzeni, D., Long, D., Cashmore, M., and Smith, D. E. (2021).**
Contrastive explanations of plans through model restrictions.

References IV

[Lerouge et al., 2022] **Lerouge, M., Gicquel, C., Mousseau, V., and Ouerdane, W. (2022).**
Explaining solutions stemming from optimization systems solving the Workforce Scheduling and Routing Problem to their end-users.
Working paper
https://hal.archives-ouvertes.fr/hal-03795653.

[Wachter et al., 2018] **Wachter, S., Mittelstadt, B., and Russell, C. (2018).**
Counterfactual explanations without opening the black box: Automated decisions and the GDPR.
Harvard Journal of Law & Technology, 31.

References V

[Wick and Thompson, 1992] **Wick, M. and Thompson, W. (1992). Reconstructive Expert System explanation.** *Artificial Intelligence*, 54:33–70.