Designing methods for explaining solutions stemming from optimization systems

Application to the workforce scheduling and routing problem

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

ROADEF
23 February 2022

Plan

1. Introduction
    - General context and our use case (WSRP)
    - Motivations for explaining solutions and our goals
2. Related works
    - Explanations in operations research literature
    - Regular characteristics of questions
3. Our method for explaining WSRP solutions
    - End-user's questions
    - Overview of the question-to-explanation process
    - Examples of explanations for the end-user
4. Conclusion

Plan

1. Introduction
   - General context and our use case (WSRP)
   - Motivations for explaining solutions and our goals
2. Related works
   - Explanations in operations research literature
   - Regular characteristics of questions
3. Our method for explaining WSRP solutions
   - End-user's questions
   - Overview of the question-to-explanation process
   - Examples of explanations for the end-user
4. Conclusion

General context

General context made of 3 components:

(1) **Optimization problem** modeling a real-world problem;

(2) **Optimization system** for solving the problem;

(3) **Non-expert end-user** using the optimization system.

↪ What are (1), (2) and (3) in our use case?

General context

General context made of 3 components:

(1) **Optimization problem** modeling a real-world problem;

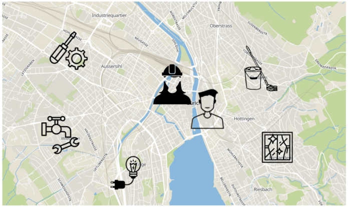
(2) **Optimization system** for solving the problem;

(3) **Non-expert end-user** using the optimization system.

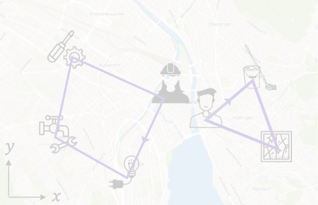
↪ What are (1), (2) and (3) in our use case?

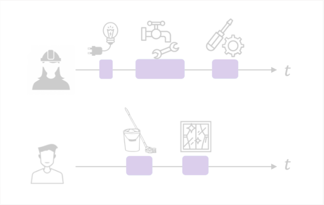
Use case - (1) Optimization problem

Workforce Scheduling and Routing Problem (WSRP):



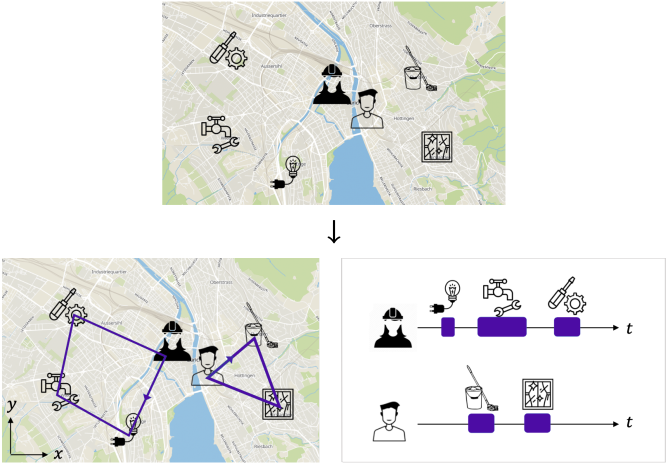
↓





Use case - (1) Optimization problem

Workforce Scheduling and Routing Problem (WSRP):



Our use case - (1) Optimization problem

Instance of the WSRP:

- $\mathcal{E} = \{e_1, \ \ldots, \ e_n\}$

  set of n **mobile employees** e_i characterized by:

  - a skill level;
  - a working time-window;
  - a location.

- $\mathcal{T} = \{t_1, \ \ldots, \ t_m\}$

  set of m **tasks** t_j characterized by:

  - a skill level;
  - an availability time-window;
  - a duration;
  - a location.

Our use case - (1) Optimization problem

Instance of the WSRP:

- $\mathcal{E} = \{e_1, \ \ldots, \ e_n\}$

  set of *n* **mobile employees** e_i characterized by:

  - a skill level;
  - a working time-window;
  - a location.

- $\mathcal{T} = \{t_1, \ \ldots, \ t_m\}$

  set of *m* **tasks** t_j characterized by:

  - a skill level;
  - an availability time-window;
  - a duration;
  - a location.

Our use case - (1) Optimization problem

Integer Programming model of the WSRP:

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

s.t. - employees must work within their time windows;
- tasks must be performed within their time windows;
- employees must be skilled enough to perform the tasks;
- ...

$U_{e_i,(t_j,t_k)} \in \{0,1\}$ whether or not e_i goes from t_j to t_k;

$T_{t_j} \in \mathbb{N}$ start time of t_j.

Our use case - (1) Optimization problem

Integer Programming model of the WSRP:

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

s.t.
- employees must work within their time windows;
- tasks must be performed within their time windows;
- employees must be skilled enough to perform the tasks;
- ...

$U_{e_i,(t_j,t_k)} \in \{0,1\}$ whether or not e_i goes from t_j to t_k;

$T_{t_j} \in \mathbb{N}$ start time of t_j.

Our use case - (1) Optimization problem

Integer Programming model of the WSRP:

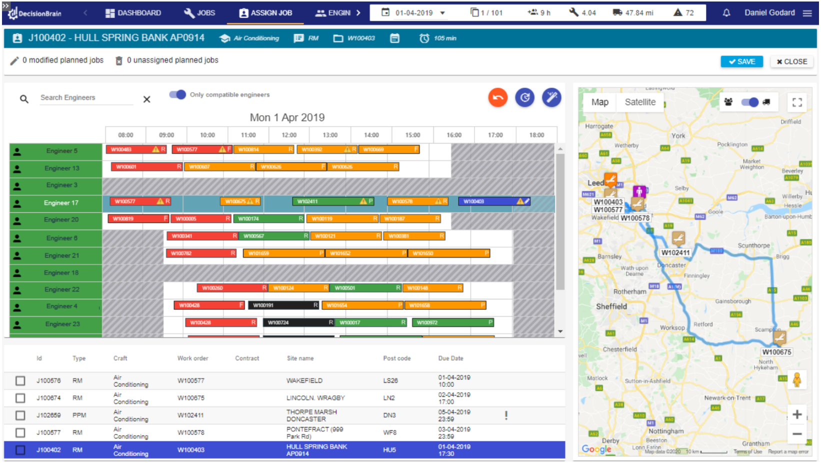
lex max $\big(\text{total working duration}, \; -\text{total traveling duration}\big)$

s.t.
- employees must work within their time windows;
- tasks must be performed within their time windows;
- employees must be skilled enough to perform the tasks;
- ...

$U_{e_i,(t_j,t_k)} \in \{0,1\}$ whether or not e_i goes from t_j to t_k;
$T_{t_j} \in \mathbb{N}$ start time of t_j.

WSRP-solving system:

e.g DecisionBrain's Dynamic Scheduler



DecisionBrain | DASHBOARD | JOBS | ASSIGN JOB | ENGIN | 01-04-2019 | 1 / 101 | 9 h | 4.04 | 47.84 mi | 72 | Daniel Godard

J100402 - HULL SPRING BANK AP0914 | Air Conditioning | RM | W100403 | 105 min

0 modified planned jobs | 0 unassigned planned jobs | ✓ SAVE | × CLOSE

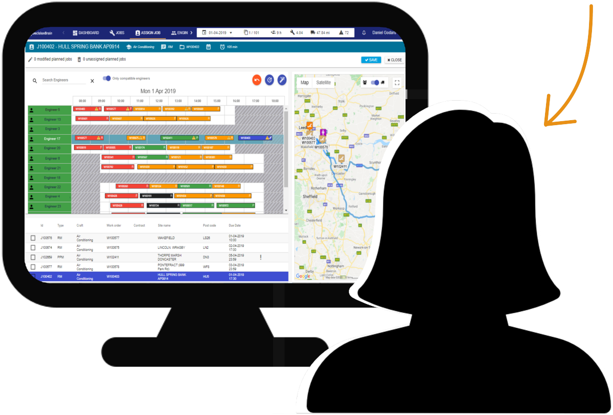
Search Engineers | Only compatible engineers

Mon 1 Apr 2019

	Id	Type	Craft	Work order	Contract	Site name	Post code	Due Date	
□	J100576	RM	Air Conditioning	W100577		WAKEFIELD	LS26	01-04-2019 10:00	
□	J100674	RM	Air Conditioning	W100675		LINCOLN. WRAGBY	LN2	02-04-2019 17:00	
□	J102659	PPM	Air Conditioning	W102411		THORPE MARSH DONCASTER	DN3	05-04-2019 23:59	!
□	J100577	RM	Air Conditioning	W100578		PONTEFRACT (999 Park Rd)	WF8	03-04-2019 23:59	
□	J100402	RM	Air Conditioning	W100403		HULL SPRING BANK AP0914	HU5	01-04-2019 17:30	

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

Planner:



J100402 - HULL SPRING BANK AP0914

0 modified planned jobs 0 unassigned planned jobs

Mon 1 Apr 2019

08:00 09:00 10:00 11:00 12:00 13:00 14:00 15:00 16:00 17:00 18:00

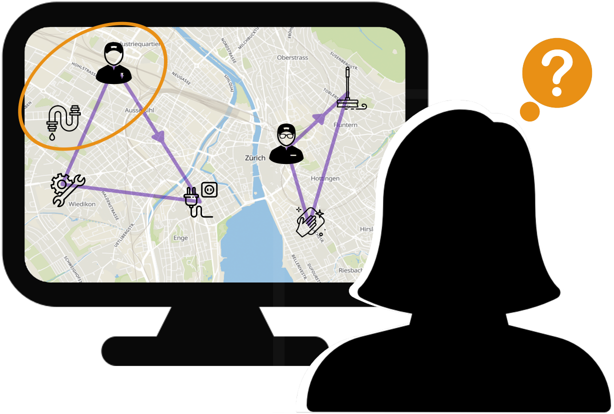
	Id	Type	Craft	Work order	Contract	Site name	Post code	Due Date	
□	J100576	RM	Air Conditioning	W100577		WAKEFIELD	LS26	01-04-2019 10:00	
□	J100674	RM	Air Conditioning	W100675		LINCOLN. WRAGBY	LN2	02-04-2019 17:00	
□	J102659	PPM	Air Conditioning	W102411		THORPE MARSH DONCASTER	DN3	05-04-2019 23:59	!
□	J100577	RM	Air Conditioning	W100578		PONTEFRACT (999 Park Rd)	WF8	03-04-2019 23:59	
□	J100402	RM	Air Conditioning	W100403		HULL SPRING BANK AP0914	HU5	01-04-2019 17:30	

Plan

1. Introduction
   - General context and our use case (WSRP)
   - Motivations for explaining solutions and our goals
2. Related works
   - Explanations in operations research literature
   - Regular characteristics of questions
3. Our method for explaining WSRP solutions
   - End-user's questions
   - Overview of the question-to-explanation process
   - Examples of explanations for the end-user
4. Conclusion

Motivations for explaining solutions

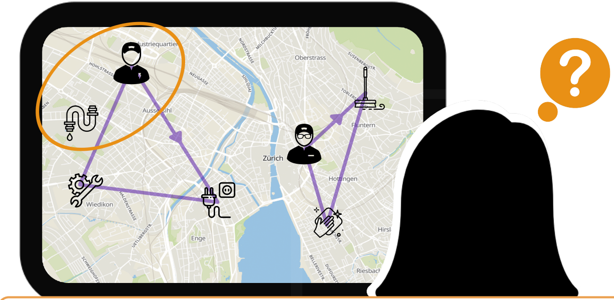
End-user may have **questions and doubts** about a solution.



↪ **Reluctance** to apply the solution, **frustration**, *etc.* ...

Motivations for explaining solutions

End-user may have **questions and doubts** about a solution.

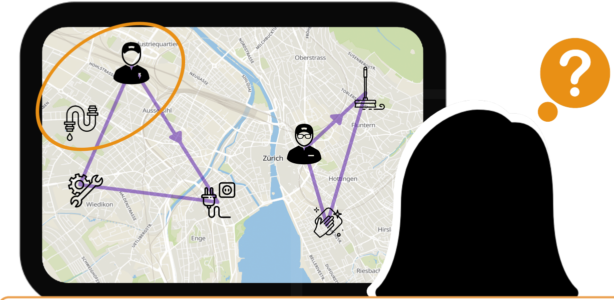


"Why is Adam not performing the plumbing task in addition to the tasks of his planning?"

↪ **Reluctance** to apply the solution, **frustration**, *etc.* ...

Motivations for explaining solutions

End-user may have **questions and doubts** about a solution.



"Why is Adam not performing the plumbing task in addition to the tasks of his planning?"

↪ **Reluctance** to apply the solution, **frustration**, *etc.* ...

Our goals

Designing **methods**, for **explaining (WSRP) solutions**, which:

- enable users to **ask various questions** about a given solution and **get explanations** back;
- are **independent from the solving process**.

↪ **Tackling end-users' issues** (questions, doubts, frustration, reluctance, *etc.*), **improving their trust** in the system and **their confidence** at work.

Our goals

Designing **methods**, for **explaining (WSRP) solutions**, which:

- enable users to **ask various questions** about a given solution and **get explanations** back;
- are **independent from the solving process**.

↪ **Tackling end-users' issues** (questions, doubts, frustration, reluctance, *etc.*), **improving their trust** in the system and **their confidence** at work.

Our goals

Designing **methods**, for **explaining (WSRP) solutions**, which:

- enable users to **ask various questions** about a given solution and **get explanations** back;
- are **independent from the solving process**.

↪ **Tackling end-users' issues** (questions, doubts, frustration, reluctance, *etc.*), **improving their trust** in the system and **their confidence** at work.

Our goals

Designing **methods**, for **explaining (WSRP) solutions**, which:

- enable users to **ask various questions** about a given solution and **get explanations** back;
- are **independent from the solving process**.

↪ **Tackling end-users' issues** (questions, doubts, frustration, reluctance, *etc.*), **improving their trust** in the system and **their confidence** at work.

Plan

1. Introduction
   - General context and our use case (WSRP)
   - Motivations for explaining solutions and our goals
2. Related works
   - Explanations in operations research literature
   - Regular characteristics of questions
3. Our method for explaining WSRP solutions
   - End-user's questions
   - Overview of the question-to-explanation process
   - Examples of explanations for the end-user
4. Conclusion

Explanations in operations research literature

Few works dealing with explanations in optimization, among them:

Articles	Methods for explaining solutions			
	Applied to	Applicable to	Dependance	Questions
[Ludwig et al., 2018]	Makespan Scheduling Problem	Specific MSP solved via specific algorithm	Depending on solving algorithm	1 type
[Čyras et al., 2019]	Makespan Scheduling Problem	Specific problems with binary decision variables	Not depending on solving algorithm	3 types
[Korikov et al., 2021]	Knapsack, Portfolio	Specific linear problems whose weights in OF are not in constraints	Depending on solving algorithm	1 type

→ We want a method that is **less specific**, **not depending** on solving algorithm and handling **more questions**.

[Ludwig et al., 2018] Explaining Complex Scheduling Decisions
[Čyras et al., 2019] Argumentation for Explainable Scheduling
[Korikov et al., 2021] Counterfactual Explanations for Optimization-Based Decisions in the Context of the GDPR

Explanations in operations research literature

Few works dealing with explanations in optimization, among them:

Articles	Methods for explaining solutions			
	Applied to	Applicable to	Dependance	Questions
[Ludwig et al., 2018]	Makespan Scheduling Problem	Specific MSP solved via specific algorithm	Depending on solving algorithm	1 type
[Čyras et al., 2019]	Makespan Scheduling Problem	Specific problems with binary decision variables	Not depending on solving algorithm	3 types
[Korikov et al., 2021]	Knapsack, Portfolio	Specific linear problems whose weights in OF are not in constraints	Depending on solving algorithm	1 type

↪ We want a method that is **less specific**, **not depending** on solving algorithm and handling **more questions**.

[Ludwig et al., 2018] Explaining Complex Scheduling Decisions
[Čyras et al., 2019] Argumentation for Explainable Scheduling
[Korikov et al., 2021] Counterfactual Explanations for Optimization-Based Decisions in the Context of the GDPR

Explanations in operations research literature

Few works dealing with explanations in optimization, among them:

Articles		Methods for explaining solutions		
	Applied to	Applicable to	Dependance	Questions
[Ludwig et al., 2018]	Makespan Scheduling Problem	Specific MSP solved via specific algorithm	Depending on solving algorithm	1 type
[Čyras et al., 2019]	Makespan Scheduling Problem	Specific problems with binary decision variables	Not depending on solving algorithm	3 types
[Korikov et al., 2021]	Knapsack, Portfolio	Specific linear problems whose weights in OF are not in constraints	Depending on solving algorithm	1 type

↪ We want a method that is **less specific**, **not depending** on solving algorithm and handling **more questions**.

[Ludwig et al., 2018] Explaining Complex Scheduling Decisions
[Čyras et al., 2019] Argumentation for Explainable Scheduling
[Korikov et al., 2021] Counterfactual Explanations for Optimization-Based Decisions in the Context of the GDPR

Explanations in operations research literature

Few works dealing with explanations in optimization, among them:

Articles	Methods for explaining solutions			
	Applied to	Applicable to	Dependance	Questions
[Ludwig et al., 2018]	Makespan Scheduling Problem	Specific MSP solved via specific algorithm	Depending on solving algorithm	1 type
[Čyras et al., 2019]	Makespan Scheduling Problem	Specific problems with binary decision variables	Not depending on solving algorithm	3 types
[Korikov et al., 2021]	Knapsack, Portfolio	Specific linear problems whose weights in OF are not in constraints	Depending on solving algorithm	1 type

↪ We want a method that is **less specific**, **not depending** on solving algorithm and handling **more questions**.

[Ludwig et al., 2018] Explaining Complex Scheduling Decisions
[Čyras et al., 2019] Argumentation for Explainable Scheduling
[Korikov et al., 2021] Counterfactual Explanations for Optimization-Based Decisions in the Context of the GDPR

Explanations in operations research literature

Few works dealing with explanations in optimization, among them:

Articles	Methods for explaining solutions			
	Applied to	Applicable to	Dependance	Questions
[Ludwig et al., 2018]	Makespan Scheduling Problem	Specific MSP solved via specific algorithm	Depending on solving algorithm	1 type
[Čyras et al., 2019]	Makespan Scheduling Problem	Specific problems with binary decision variables	Not depending on solving algorithm	3 types
[Korikov et al., 2021]	Knapsack, Portfolio	Specific linear problems whose weights in OF are not in constraints	Depending on solving algorithm	1 type

↪ We want a method that is **less specific**, **not depending** on solving algorithm and handling **more questions**.

[Ludwig et al., 2018] Explaining Complex Scheduling Decisions
[Čyras et al., 2019] Argumentation for Explainable Scheduling
[Korikov et al., 2021] Counterfactual Explanations for Optimization-Based Decisions in the Context of the GDPR

Explanations in operations research literature

Few works dealing with explanations in optimization, among them:

Articles	Methods for explaining solutions			
	Applied to	Applicable to	Dependance	Questions
[Ludwig et al., 2018]	Makespan Scheduling Problem	Specific MSP solved via specific algorithm	Depending on solving algorithm	1 type
[Čyras et al., 2019]	Makespan Scheduling Problem	Specific problems with binary decision variables	Not depending on solving algorithm	3 types
[Korikov et al., 2021]	Knapsack, Portfolio	Specific linear problems whose weights in OF are not in constraints	Depending on solving algorithm	1 type

↪ We want a method that is **less specific**, **not depending** on solving algorithm and handling **more questions**.

[Ludwig et al., 2018] Explaining Complex Scheduling Decisions
[Čyras et al., 2019] Argumentation for Explainable Scheduling
[Korikov et al., 2021] Counterfactual Explanations for Optimization-Based Decisions in the Context of the GDPR

Plan

1. Introduction
   - General context and our use case (WSRP)
   - Motivations for explaining solutions and our goals
2. Related works
   - Explanations in operations research literature
   - Regular characteristics of questions
3. Our method for explaining WSRP solutions
   - End-user's questions
   - Overview of the question-to-explanation process
   - Examples of explanations for the end-user
4. Conclusion

Regular characteristics of questions

In eXplainable Artificial Intelligence (XAI), **questions** are often:

- **local** *i.e* focusing on a **specific result** generated by the system used (≠ global questions);
- **contrastive** *i.e* having the following **form**:
  "Why $\underbrace{\text{this current result}}_{\text{fact}}$ rather than $\underbrace{\text{that other one}}_{\text{foil (often implicit)}}$?"
- **templates** *i.e.* questions with empty **fields** to fill with data.

↪ **In our work**, the end-user's questions will be **local contrastive templates**.

Regular characteristics of questions

In eXplainable Artificial Intelligence (XAI), **questions** are often:

- **local** *i.e* focusing on a **specific result** generated by the system used (≠ global questions);
- **contrastive** *i.e* having the following **form**:
  "Why $\underbrace{\text{this current result}}_{\text{fact}}$ rather than $\underbrace{\text{that other one}}_{\text{foil (often implicit)}}$?"
- **templates** *i.e.* questions with empty **fields** to fill with data.

↪ **In our work**, the end-user's questions will be **local contrastive templates**.

Regular characteristics of questions

In eXplainable Artificial Intelligence (XAI), **questions** are often:

- **local** *i.e* focusing on a **specific result** generated by the system used ($\neq$ global questions);
- **contrastive** *i.e* having the following **form**:
  "Why $\underbrace{\text{this current result}}_{\text{fact}}$ rather than $\underbrace{\text{that other one}}_{\text{foil (often implicit)}}$?"
- **templates** *i.e.* questions with empty **fields** to fill with data.

↪ **In our work**, the end-user's questions will be **local contrastive templates**.

Regular characteristics of questions

In eXplainable Artificial Intelligence (XAI), **questions** are often:

- **local** *i.e* focusing on a **specific result** generated by the system used (≠ global questions);
- **contrastive** *i.e* having the following **form**:
  "Why $\underbrace{\text{this current result}}_{\text{fact}}$ rather than $\underbrace{\text{that other one}}_{\text{foil (often implicit)}}$?"
- **templates** *i.e.* questions with empty **fields** to fill with data.

↪ **In our work**, the end-user's questions will be **local contrastive templates**.

Regular characteristics of questions

In eXplainable Artificial Intelligence (XAI), **questions** are often:

- **local** *i.e* focusing on a **specific result** generated by the system used ($\neq$ global questions);
- **contrastive** *i.e* having the following **form**:

  "Why $\underbrace{\text{this current result}}_{\text{fact}}$ rather than $\underbrace{\text{that other one}}_{\text{foil (often implicit)}}$?"
- **templates** *i.e.* questions with empty **fields** to fill with data.

↪ **In our work**, the end-user's questions will be **local contrastive templates**.

Plan

1. Introduction
   - General context and our use case (WSRP)
   - Motivations for explaining solutions and our goals
2. Related works
   - Explanations in operations research literature
   - Regular characteristics of questions
3. Our method for explaining WSRP solutions
   - End-user's questions
   - Overview of the question-to-explanation process
   - Examples of explanations for the end-user
4. Conclusion

End-user's questions

List of end-user's questions:

15 templates q including questions about:

- an **insertion of a task** in plannings;
  e.g. "Why is $\langle e_i \rangle$ not performing $\langle t_j \rangle$ ...
  - ... just after $\langle t_k \rangle$?"
  - ... between two consecutive tasks of their planning?"
  - ... in addition to the tasks of their planning?"
- a **swap of tasks** in/out of plannings;
- a **change of order** of tasks in plannings.

End-user's questions

List of end-user's questions:

15 templates q including questions about:

- an **insertion of a task** in plannings;
  e.g. "Why is $\langle e_i \rangle$ not performing $\langle t_j \rangle$ ...
  - ... just after $\langle t_k \rangle$?"
  - ... between two consecutive tasks of their planning?"
  - ... in addition to the tasks of their planning?"
- a **swap of tasks** in/out of plannings;
- a **change of order** of tasks in plannings.

End-user's questions

List of end-user's questions:

15 templates q including questions about:

- an **insertion of a task** in plannings;
  e.g. "Why is $\langle e_i \rangle$ not performing $\langle t_j \rangle$ ...
  - ... just after $\langle t_k \rangle$?"
  - ... between two consecutive tasks of their planning?"
  - ... in addition to the tasks of their planning?"
- a **swap of tasks** in/out of plannings;
- a **change of order** of tasks in plannings.

End-user's questions

Transformations suggested in questions:

Each question q suggests to **transform** the given solution:

- by **inserting** a task;
- by **swapping** tasks;
- by **changing the order** of tasks in a planning.

Neighboring solutions induced by questions:

Each question q induces a **set of neighboring solutions** $\mathcal{N}(q)$

↪ We can **exploit** $\mathcal{N}(q)$ for answering to q.

End-user's questions

Transformations suggested in questions:

Each question q suggests to **transform** the given solution:

- by **inserting** a task;
- by **swapping** tasks;
- by **changing the order** of tasks in a planning.

Neighboring solutions induced by questions:

Each question q induces a **set of neighboring solutions** $\mathcal{N}(q)$

↪ We can **exploit** $\mathcal{N}(q)$ for answering to q.

End-user's questions

Transformations suggested in questions:

Each question q suggests to **transform** the given solution:

- by **inserting** a task;
- by **swapping** tasks;
- by **changing the order** of tasks in a planning.

Neighboring solutions induced by questions:

Each question q induces a **set of neighboring solutions** $\mathcal{N}(q)$

$\hookrightarrow$ We can **exploit** $\mathcal{N}(q)$ for answering to q.

Plan

1. Introduction
   - General context and our use case (WSRP)
   - Motivations for explaining solutions and our goals
2. Related works
   - Explanations in operations research literature
   - Regular characteristics of questions
3. Our method for explaining WSRP solutions
   - End-user's questions
   - Overview of the question-to-explanation process
   - Examples of explanations for the end-user
4. Conclusion

Overview of the question-to-explanation process

Let q be an end-user's contrastive **question**:

"Why is this fact rather than that foil?"

Answering q can lead to **two possible cases**.

- **Negative case** ≃ "the foil is not possible/interesting" (with arguments using quantities);
- **Positive case** ≃ "the foil is not observed in the given solution, but it can be observed and it improves the solution" (with new feasible solution).

↪ We are mainly interested in explaining **negative cases**.

Overview of the question-to-explanation process

Let q be an end-user's contrastive **question**:

"Why is this fact rather than that foil?"

Answering q can lead to **two possible cases**.

- **Negative case** ≃ "the foil is not possible/interesting" (with arguments using quantities);
- **Positive case** ≃ "the foil is not observed in the given solution, but it can be observed and it improves the solution" (with new feasible solution).

↪ We are mainly interested in explaining **negative cases**.

Overview of the question-to-explanation process

Let q be an end-user's contrastive **question**:

"Why is this fact rather than that foil?"

Answering q can lead to **two possible cases**.

- **Negative case** ≃ "the foil is not possible/interesting" (with arguments using quantities);
- **Positive case** ≃ "the foil is not observed in the given solution, but it can be observed and it improves the solution" (with new feasible solution).

↪ We are mainly interested in explaining **negative cases**.

Overview of the question-to-explanation process

Let q be an end-user's contrastive **question**:

"Why is this fact rather than that foil?"

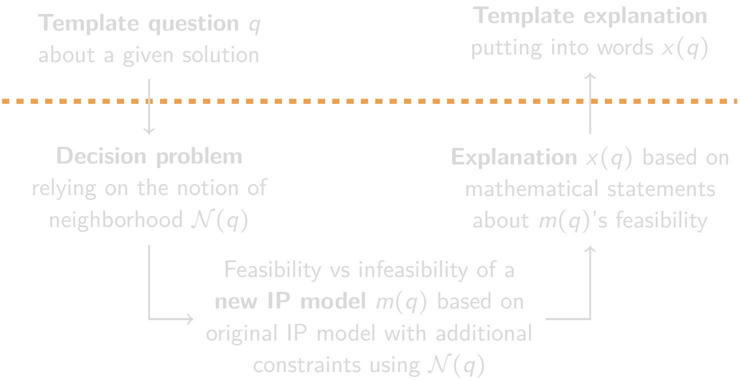
Answering q can lead to **two possible cases**.

- **Negative case** ≃ "the foil is not possible/interesting" (with arguments using quantities);
- **Positive case** ≃ "the foil is not observed in the given solution, but it can be observed and it improves the solution" (with new feasible solution).

↪ We are mainly interested in explaining **negative cases**.

Overview of the question-to-explanation process

End-user's side



Template question q about a given solution

Template explanation putting into words $x(q)$

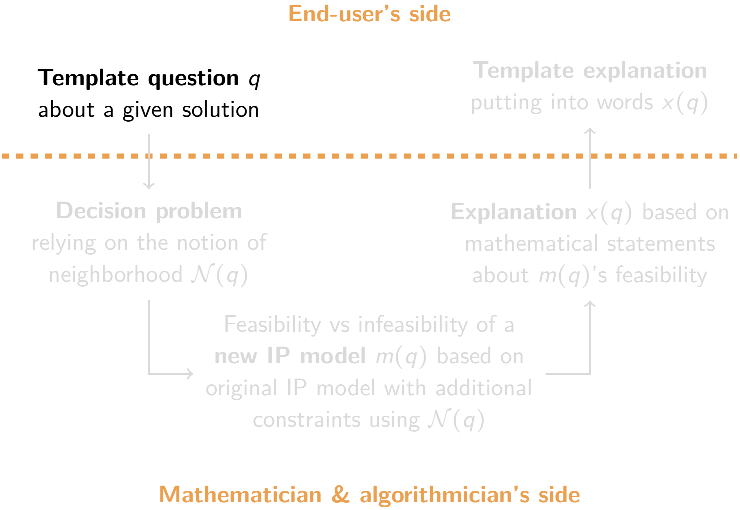
Decision problem relying on the notion of neighborhood $\mathcal{N}(q)$

Explanation $x(q)$ based on mathematical statements about $m(q)$'s feasibility

Feasibility vs infeasibility of a **new IP model** $m(q)$ based on original IP model with additional constraints using $\mathcal{N}(q)$

Mathematician & algorithmician's side

Overview of the question-to-explanation process



End-user's side

Template question q
about a given solution

Template explanation
putting into words $x(q)$

Decision problem
relying on the notion of
neighborhood $\mathcal{N}(q)$

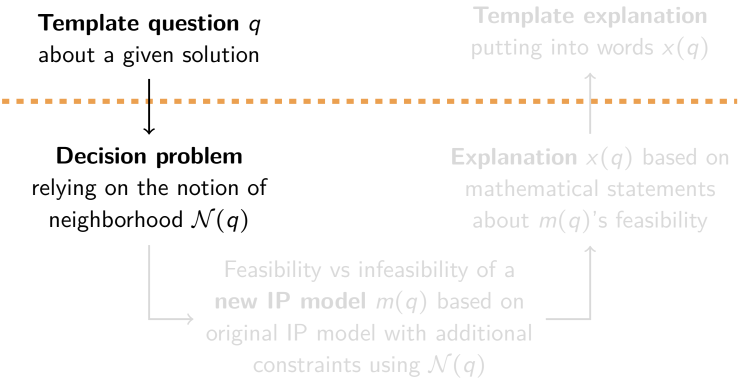
Explanation $x(q)$ based on
mathematical statements
about $m(q)$'s feasibility

Feasibility vs infeasibility of a
new IP model $m(q)$ based on
original IP model with additional
constraints using $\mathcal{N}(q)$

Mathematician & algorithmician's side

Overview of the question-to-explanation process

End-user's side



Template question q
about a given solution

Decision problem
relying on the notion of
neighborhood $\mathcal{N}(q)$

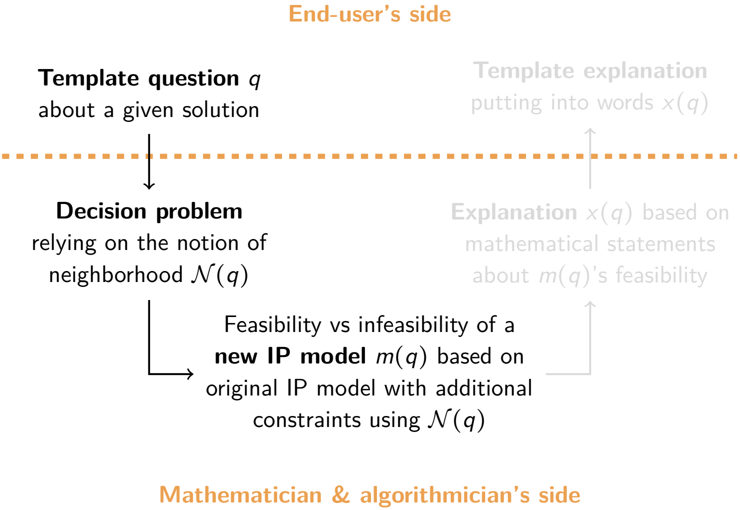
Feasibility vs infeasibility of a
new IP model $m(q)$ based on
original IP model with additional
constraints using $\mathcal{N}(q)$

Explanation $x(q)$ based on
mathematical statements
about $m(q)$'s feasibility

Template explanation
putting into words $x(q)$

Mathematician & algorithmician's side

Overview of the question-to-explanation process



End-user's side

Template question q
about a given solution

Template explanation
putting into words $x(q)$

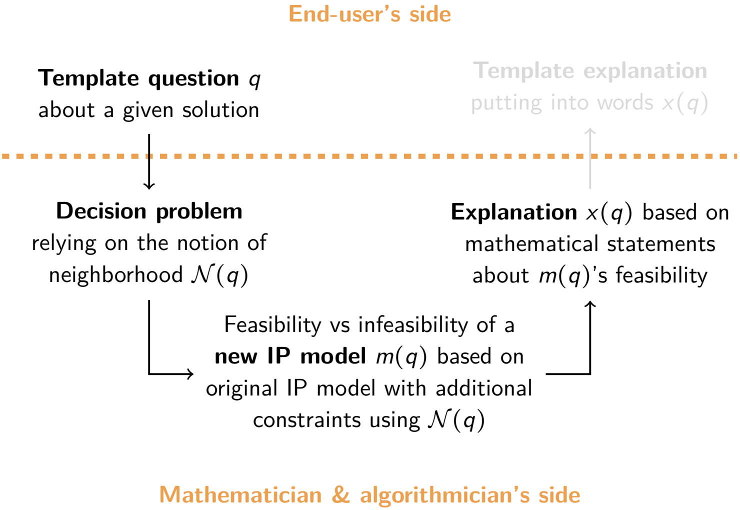
Decision problem
relying on the notion of
neighborhood $\mathcal{N}(q)$

Explanation $x(q)$ based on
mathematical statements
about $m(q)$'s feasibility

Feasibility vs infeasibility of a
new IP model $m(q)$ based on
original IP model with additional
constraints using $\mathcal{N}(q)$

Mathematician & algorithmician's side

Overview of the question-to-explanation process



End-user's side

Template question q
about a given solution

Template explanation
putting into words $x(q)$

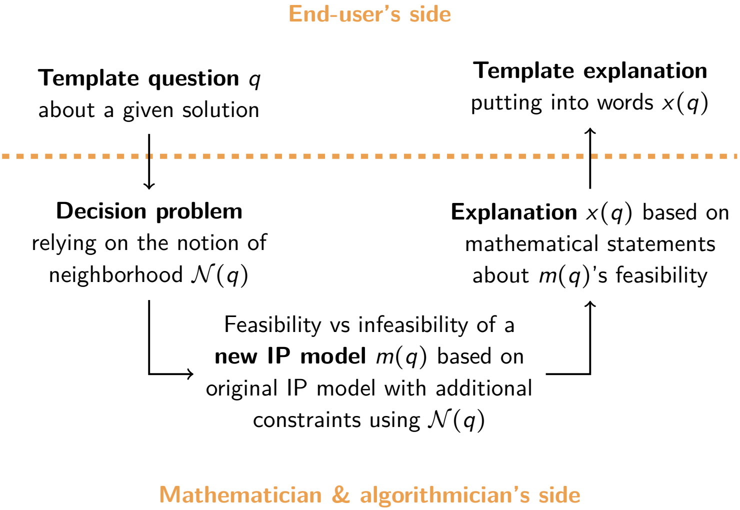
Decision problem
relying on the notion of
neighborhood $\mathcal{N}(q)$

Explanation $x(q)$ based on
mathematical statements
about $m(q)$'s feasibility

Feasibility vs infeasibility of a
new IP model $m(q)$ based on
original IP model with additional
constraints using $\mathcal{N}(q)$

Mathematician & algorithmician's side

Overview of the question-to-explanation process



End-user's side

Template question q about a given solution

Template explanation putting into words $x(q)$

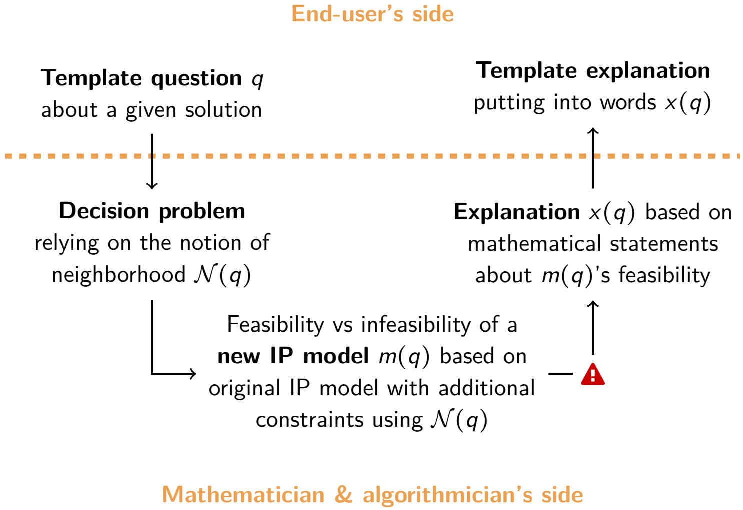
Decision problem relying on the notion of neighborhood $\mathcal{N}(q)$

Explanation $x(q)$ based on mathematical statements about $m(q)$'s feasibility

Feasibility vs infeasibility of a **new IP model** $m(q)$ based on original IP model with additional constraints using $\mathcal{N}(q)$

Mathematician & algorithmician's side

Overview of the question-to-explanation process



End-user's side

Template question q
about a given solution

Template explanation
putting into words $x(q)$

Decision problem
relying on the notion of
neighborhood $\mathcal{N}(q)$

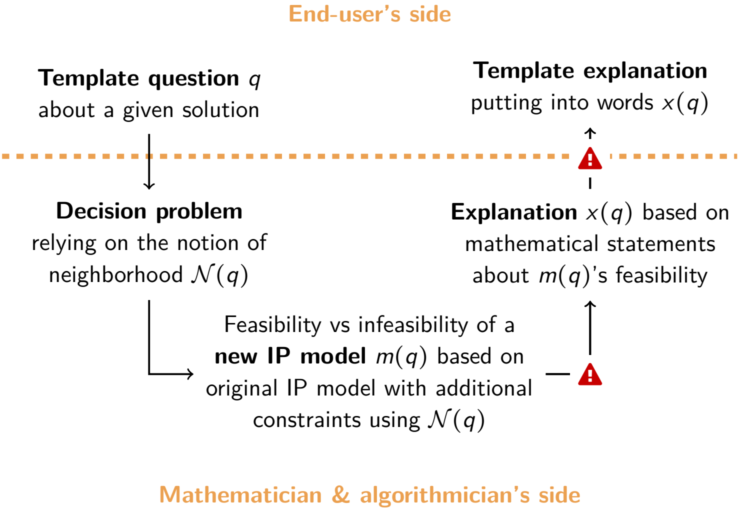
Explanation $x(q)$ based on
mathematical statements
about $m(q)$'s feasibility

Feasibility vs infeasibility of a
new IP model $m(q)$ based on
original IP model with additional
constraints using $\mathcal{N}(q)$

⚠

Mathematician & algorithmician's side

Overview of the question-to-explanation process



End-user's side

Template question q
about a given solution

Template explanation
putting into words $x(q)$

Decision problem
relying on the notion of
neighborhood $\mathcal{N}(q)$

Explanation $x(q)$ based on
mathematical statements
about $m(q)$'s feasibility

Feasibility vs infeasibility of a
new IP model $m(q)$ based on
original IP model with additional
constraints using $\mathcal{N}(q)$

Mathematician & algorithmician's side

Plan

1. Introduction
   - General context and our use case (WSRP)
   - Motivations for explaining solutions and our goals
2. Related works
   - Explanations in operations research literature
   - Regular characteristics of questions
3. Our method for explaining WSRP solutions
   - End-user's questions
   - Overview of the question-to-explanation process
   - Examples of explanations for the end-user
4. Conclusion

Examples of explanations for the end-user

Question q such that $x(q)$ is "easy" to compute and express:

- q: "Why is Adam not performing the plumbing task **just after the mechanical one**?"
- $x(q)$ expressed as:
  "**If so**, Adam would start the plumbing task at the earliest at 3:30PM **while** he must start it at the latest at 2:30PM so that he can be back at home by 6:00PM.
  Hence Adam is not performing the plumbing task just after the mechanical one."

Examples of explanations for the end-user

Question q such that $x(q)$ is "easy" to compute and express:

- q: "Why is Adam not performing the plumbing task **just after the mechanical one**?"
- $x(q)$ expressed as:
  "**If so**, Adam would start the plumbing task at the earliest at 3:30PM **while** he must start it at the latest at 2:30PM so that he can be back at home by 6:00PM.
  Hence Adam is not performing the plumbing task just after the mechanical one."

Examples of explanations for the end-user

Question q such that $x(q)$ is "hard" to compute and express:

- q: "Why is Adam not performing the plumbing task **in addition to his tasks**?"
- $x(q)$ expressed as: "**Making** Adam perform the plumbing task in addition to his tasks **would not produce any solution that is feasible**. **In the best scenario**, Adam does the plumbing, mechanical and electricity tasks in this order. **But even in this scenario**, Adam would be at electricity task at the earliest at 3:40PM **while** he must start it at the latest at 3:30PM so that he can be back at home by 6:00PM. **Hence** Adam is not performing the plumbing task in addition to his tasks."

Examples of explanations for the end-user

Question q such that $x(q)$ is "hard" to compute and express:

- q: "Why is Adam not performing the plumbing task **in addition to his tasks**?"
- $x(q)$ expressed as:
  "**Making** Adam perform the plumbing task in addition to his tasks **would not produce any solution that is feasible**.
  In the best scenario, Adam does the plumbing, mechanical and electricity tasks in this order. **But even in this scenario**, Adam would be at electricity task at the earliest at 3:40PM **while** he must start it at the latest at 3:30PM so that he can be back at home by 6:00PM.
  Hence Adam is not performing the plumbing task in addition to his tasks."

Conclusion

Achieved work:

- List of ends users' questions that are local contrastive templates in a WSRP context;
- Method for explaining (WSRP) solutions starting with an end-user's question q ending with an explanation $x(q)$, knowing that:
  - some $x(q)$ are more or less "tough to compute";
  - some $x(q)$ are more or less "tough to put into words".

Conclusion

Forthcoming challenges:

- How to deal with **less restricted** end-users' **questions**?
  e.g. "Why is Adam working much less than Ellen?"
- How much **generic** our **method** is?
  Can we transpose it to other optimization problems?
- How to **structure the exploration** of solutions?
- How to make interactions with end-users closer to a **dialog**?

References I

[Cashmore et al., 2019] Cashmore, M., Collins, A., Krarup, B., Krivic, S., Magazzeni, D., and Smith, D. (2019).
Towards explainable ai planning as a service.

[Čyras et al., 2019] Čyras, K., Letsios, D., Misener, R., and Toni, F. (2019).
Argumentation for explainable scheduling.
Proceedings of the AAAI Conference on Artificial Intelligence.

[Korikov et al., 2021] Korikov, A., Shleyfman, A., and Beck, J. C. (2021).
Counterfactual explanations for optimization-based decisions in the context of the gdpr.

References II

In Zhou, Z.-H., editor, *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21*, pages 4097–4103. International Joint Conferences on Artificial Intelligence 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.

[Lipton, 1990] Lipton, P. (1990). Contrastive explanation. *Royal Institute of Philosophy Supplement*, 27:247–266.

References III

[Ludwig et al., 2018] Ludwig, J., Kalton, A., and Stottler, R. (2018).
Explaining complex scheduling decisions.
In *IUI Workshops*.

[Miller, 2019] Miller, T. (2019).
Explanation in artificial intelligence: Insights from the social sciences.
Artificial Intelligence, 267:1–38.

[Wick and Thompson, 1992] Wick, M. R. and Thompson, W. B. (1992).
Reconstructive expert system explanation.
Artificial Intelligence, 54(1–2):33–70.

First regular characteristic of questions

Local questions:

In explainable artificial intelligence,
e.g. [Wick and Thompson, 1992]:

- **Local** questions focus on a **specific result** generated by the system used;
- ≠ **global** ones which relate to the **system's functioning**.

↪ **In our work**, the leend-user's questions are **local**:
e.g. "Why is Adam not performing the plumbing task in addition to the tasks of his planning?" **in the given solution**.

[Wick and Thompson, 1992] Reconstructive Expert System Explanation

Second regular characteristic of questions

Contrastive questions:

In social sciences [Lipton, 1990]:

- Questions having the following **form**:

  "Why $\underbrace{\text{this current result}}_{\text{fact}}$ rather than $\underbrace{\text{that other one}}_{\text{foil (often implicit)}}$?"

- It is **relevant** to work with such questions as they correspond to most of the "Why" questions people ask [Miller, 2019].

↪ **In our work**, the end-user's questions are **contrastive**:
e.g. "Why is Adam not performing the plumbing task in addition to the tasks of his planning?"

[Lipton, 1990] Contrastive explanation
[Miller, 2019] Explanation in artificial intelligence: Insights from the social sciences

Third regular characteristic of questions

Template questions:

In explainable artificial intelligence planning,
e.g. [Cashmore et al., 2019] and [Krarup et al., 2021]:

- Questions with empty **fields** to fill with data from the result.
- It supposes to set a **list of end-users' questions**.

↪ **In our work**, the end-user's questions are templates *e.g.* "Why is $\langle \boldsymbol{e_i} \rangle$ not performing $\langle \boldsymbol{t_j} \rangle$ in addition to the tasks of his planning?" with $e_i \leftarrow$ Adam and $t_j \leftarrow$ the plumbing task.

[Cashmore et al., 2019] Towards explainable AI planning as a service
[Krarup et al., 2021] Contrastive explanations of plans through model restrictions

Regular characteristics of questions

e.g. q: "Why is Adam not performing the plumbing task
in addition to the tasks of his planning?"

This question q is:

- **local** - q supposes implicitely "**in the given solution**";
- **contrastive** - the fact of q is

  "Why is Adam not performing the plumbing task
  in addition to the tasks of his planning?"
- **template** - q is equivalent to:

  "Why is $\langle \boldsymbol{e_i} \rangle$ not performing $\langle \boldsymbol{t_j} \rangle$
  in addition to the tasks of their planning?"

  with $e_i \leftarrow$ Adam and $t_j \leftarrow$ the plumbing task.

Regular characteristics of questions

e.g. q: "Why is Adam not performing the plumbing task
in addition to the tasks of his planning?"

This question q is:

- **local** - q supposes implicitely "**in the given solution**";
- **contrastive** - the fact of q is

  "Why is Adam not performing the plumbing task
  in addition to the tasks of his planning?"
- **template** - q is equivalent to:

  "Why is $\langle \boldsymbol{e_i} \rangle$ not performing $\langle \boldsymbol{t_j} \rangle$
  in addition to the tasks of their planning?"

  with $e_i \leftarrow$ Adam and $t_j \leftarrow$ the plumbing task.

Regular characteristics of questions

e.g. q: "Why is Adam not performing the plumbing task in addition to the tasks of his planning?"

This question q is:

- **local** - q supposes implicitely **"in the given solution"**;
- **contrastive** - the fact of q is

  "Why is Adam not performing the plumbing task in addition to the tasks of his planning?"
- **template** - q is equivalent to:

  "Why is $\langle \boldsymbol{e_i} \rangle$ not performing $\langle \boldsymbol{t_j} \rangle$ in addition to the tasks of their planning?"

  with $e_i \leftarrow$ Adam and $t_j \leftarrow$ the plumbing task.

Regular characteristics of questions

e.g. q: "Why is Adam not performing the plumbing task
in addition to the tasks of his planning?"

This question q is:

- **local** - q supposes implicitely "**in the given solution**";
- **contrastive** - the fact of q is

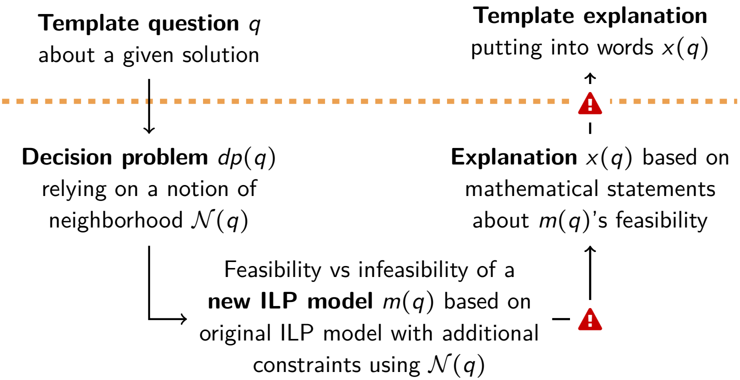
  "Why is Adam not performing the plumbing task
  in addition to the tasks of his planning?"
- **template** - q is equivalent to:

  "Why is $\langle \boldsymbol{e_i} \rangle$ not performing $\langle \boldsymbol{t_j} \rangle$
  in addition to the tasks of their planning?"

  with $e_i \leftarrow$ Adam and $t_j \leftarrow$ the plumbing task.

[Recall] Overview of our method

End-user's side



Template question q
about a given solution

Template explanation
putting into words $x(q)$

⚠

Decision problem $dp(q)$
relying on a notion of
neighborhood $\mathcal{N}(q)$

Explanation $x(q)$ based on
mathematical statements
about $m(q)$'s feasibility

Feasibility vs infeasibility of a
new ILP model $m(q)$ based on
original ILP model with additional
constraints using $\mathcal{N}(q)$

⚠

Mathematician & algorithmician's side

Categories of questions

First aspect partitioning questions q:

Do we have a **polynomial algorithm for computing** $x(q)$?

- For q_1 "Why is Adam not performing the plumbing task **just after the mechanical one**?",

  ✓ with a linear algorithm based on local search techniques;
- For q_2 "Why is Adam not performing the plumbing task **in addition to his tasks**?",

  ✗ thus resort to non-polynomial IP solving process.

  💡 We make sure that the IP model to solve is **"small enough"** to be computed in real time **in practice**.

Categories of questions

Second aspect partitioning questions q:

Do we manage to **put into words** the content of $x(q)$?

- For q_1 "Why is Adam not performing the plumbing task **just after the mechanical one**?",

  ✓ as neighboring solutions in $\mathcal{N}(q_1)$ are **similar enough**.
- For q_2 "Why is Adam not performing the plumbing task **in addition to his tasks**?",

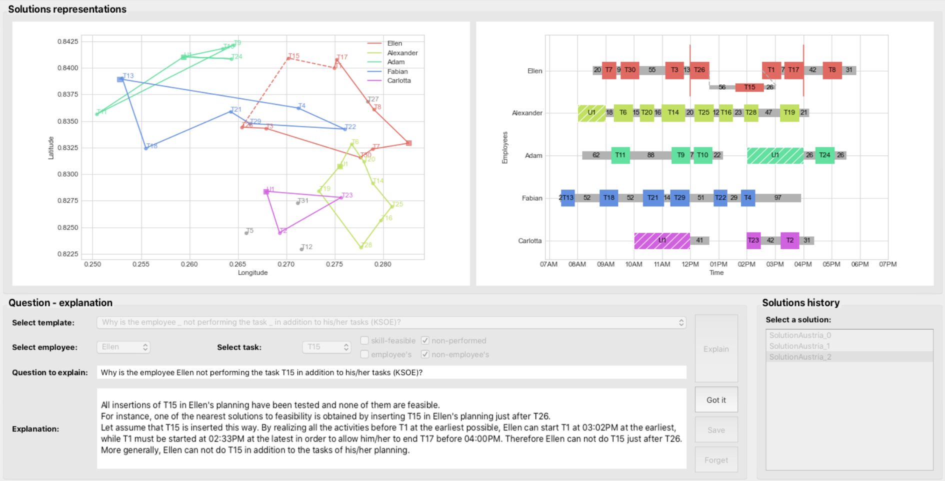
  ✗ as neighboring solutions in $\mathcal{N}(q_2)$ are **too different**.

  💡 We resort to **convincing examples** to get around.

Categories of questions

Categories of questions q	**Do we have a polynomial alg. for computing $x(q)$?**	**Do we manage to put into words $x(q)$?**
Category 1	✓ linear algorithm	✓ solutions in $\mathcal{N}(q)$ are "similar enough"
e.g. q: "Why is Adam not performing the plumbing task **just after the mechanical one**?"		
Category 2	✓ linear algorithm	✗ solutions in $\mathcal{N}(q)$ are "too different" 💡 convincing example in $\mathcal{N}(q)$
e.g. q: "Why is Adam not performing the plumbing task **between two consecutive tasks**?"		
Category 3	✗ non-polynomial alg. solving IP model 💡 small enough IP model	✗ solutions in $\mathcal{N}(q)$ are "too different" 💡 convincing example in $\mathcal{N}(q)$
e.g. q: "Why is Adam not performing the plumbing task **in addition to his tasks**?"		

Graphic User Interface prototype



Question - explanation

Select template: Why is the employee _ not performing the task _ in addition to his/her tasks (KSOE)?

Select employee: Ellen **Select task:** T15

□ skill-feasible ☑ non-performed
□ employee's ☑ non-employee's

Question to explain: Why is the employee Ellen not performing the task T15 in addition to his/her tasks (KSOE)?

Explanation:

All insertions of T15 in Ellen's planning have been tested and none of them are feasible.
For instance, one of the nearest solutions to feasibility is obtained by inserting T15 in Ellen's planning just after T26.
Let assume that T15 is inserted this way. By realizing all the activities before T1 at the earliest possible, Ellen can start T1 at 03:02PM at the earliest, while T1 must be started at 02:33PM at the latest in order to allow him/her to end T17 before 04:00PM. Therefore Ellen can not do T15 just after T26.
More generally, Ellen can not do T15 in addition to the tasks of his/her planning.

Solutions history

Select a solution:

SolutionAustria_0
SolutionAustria_1
SolutionAustria_2