Challenges of designing explanation tools for optimization systems FIIA 2021

> Mathieu Lerouge mathieulerouge.github.io

MICS, CentraleSupélec Université Paris-Saclay

7 October 2021





Our observation

End-users of an optimization system (*e.g.* DecisionBrain's) may be **reluctant to accept** the decisions computed by the system sometimes.

Our approach

Providing explanations to end-users as a way to reach some transparency and improve their confidence in the system.



- Context
- Stakes and goals

In the literature

- Explanation in optimization
- Types of questions
- Types of explanations

3 Our explanation tool

- Template questions
- Typology of (question, explanation)
- Graphic User Interface



Context

Stakes and goals

In the literature

- Explanation in optimization
- Types of questions
- Types of explanations

Our explanation tool

- Template questions
- Typology of (question, explanation)
- Graphic User Interface

3 components

- (A) Real-world complex situation that can be modeled as an optimization problem;
- (B) **Optimization system** for solving the problem;
- (C) **Non-expert end-users** using the optimization system (*via* an interface) who need explanations.
- \hookrightarrow What are (A), (B) and (C) in a given use case?

Our use case - (A) Optimization problem

Workforce Scheduling and Routing Problem (WSRP)

Instance:





Solution:

Inputs of the WSRP

•
$$\boldsymbol{\mathcal{E}} = \{\mathcal{E}_1, \ldots, \mathcal{E}_n\}$$

set of *n* **mobile employees** \mathcal{E}_i characterized by:

- a **skill level**, $s_i \in \mathbb{N}$;
- a working time-window, $\llbracket a_i, b_i \rrbracket \subset \mathbb{N}$;
- a location, $I_i \in \mathbb{R}^2$.

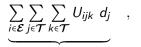
•
$$\mathcal{T} = \{\mathcal{T}_1, \ldots, \mathcal{T}_m\}$$

set of *m* tasks \mathcal{T}_j characterized by:

- a skill level, $s_j \in \mathbb{N}$;
- an availability time-window, $\llbracket a_j, b_j \rrbracket \subset \mathbb{N}$;
- a duration, $d_i \in \mathbb{N}$;
- a **location** $I_j \in \mathbb{R}^2$.

<u>Our use case - (A) Optimization problem</u>

Mathematical formulation of the WSRP



 $\max \left(\sum_{i \in \mathcal{E}} \sum_{j \in \mathcal{T}} \sum_{k \in \mathcal{T}} U_{ijk} d_j , - \sum_{i \in \mathcal{E}} \sum_{j \in \mathcal{T}} \sum_{k \in \mathcal{T}} U_{ijk} \Delta t_{jk}^i \right)$

total working duration total traveling duration

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

- ...

$$\begin{aligned} \mathcal{J}_{ijk} \in \{0,1\} & \text{whether or not } \mathcal{E}_i \text{ goes from } \mathcal{T}_j \text{ to } \mathcal{T}_k; \\ \mathcal{T}_j \in \mathbb{N} & \text{performing time of } \mathcal{T}_j. \end{aligned}$$

Our use case - (B) Optimization system

WSRP-solving systems

Deci	sionBrain	<	DASHBOAR	D 🔧 JOBS	ASSIGN	JOB 🚉 ENGIN	> 🗊 01-04-2	019 🔻 🖺 1 /	101 *#%.9 h	🔦 4.04 🗮 47.84 mi 🔺	72 A Daniel Godard
g J	100402	- HULI	L SPRING BAI	NK AP0914	🤤 Air Conditioning	🖬 RM 亡 V	1100403 🗎 i	🕤 105 min			
° 0 n	nodified pl	anned jo	bs 📋 0 unassi	gned planned jobs							SAVE X CLOS
a	Search I	Engineers	×	Conly com	patible engineers				s @ Ø	Map Satellite	
~					Mon 1	Apr 2019				Map Satellite	
			08:00	09:00 10:00	11:00	12:00 13:00	14:00 15:00	16:00	17:00 18:00	Harrogate	Dettiek
2	Engine	er 5	W100403	R W100577	A.F. W100814	R wnoose 🕰	R wtooseo F			A Wetherby	Pocklington
	Engine	er 13	W100621	R www.	R whore	s F wrona	R	1111			Market Weighton Bever
2	Engine	er 3								Leed and Seb	
	Enginee		W100577	AR	W100675 (A. R	W103411	A P W100575	AR W120403	A	W100403	
	Enginee	r 20	W100819	F W100005	R W100174	R W120119	R wtoots2	R /////		Wakefield W100578	Eaton-upon-
	Engine	er 6		W100341	R w100567	R W100121	R W100981	R			Scutthorpe
	Engine	er 21		W100752	R whonese	P wnoness	P wn	11650 P		Barnsley W10241 Wath upon Diricaster	arigg
<u> </u>	Enginee									Cearra and	aningley
_	Engine			W100426	E w100191		P was			Rotherham	Caneboreugh
_	Engine			W10050	R wig	724 R W10		109872			Retified
		0.20		9						- Worksop	Scampto
	Ы	Туре	Craft	Work order	Contract	Site name	Post code	Due Date		Chestorfield	W100675
	J100576	RM	Air Conditioning	W100577		WAKEFIELD	LS26	01-04-2019 10:00		Matiock Sutton-in Ashfield	
	J100874	RM	Air Conditioning	W100575		LINCOLN, WRAGBY	LN2	02-04-2019 17:00		- WAN	Newark on Trent
	J102659	PPM	Air Conditioning	W102411		THORPE MARSH DONCASTER	DN3	05-04-2019 23:59	1		
	J100577	RM	Air Conditioning	W100578		PONTEFRACT (999 Park Rd)	WF8	03-04-2019 23:59		Derby Beeston	Grantham
	J100402		Air Conditioning	W100403		HULL SPRING BANK AP0914		01-04-2019 17:30		Google Loss Later Map data cooge	0 km Terms of Use Report a map

Figure: Graphic User Interface of DecisionBrain's Dynamic Scheduler

Mathieu Lerouge

Our use case - (C) Non-expert end-users

Planners using WSRP-solving systems



Mathieu Lerouge

Challenges of designing explanation tools for optimization systems 7 October 2021 8 / 22



Context

Stakes and goals

In the literature

- Explanation in optimization
- Types of questions
- Types of explanations

Our explanation tool

- Template questions
- Typology of (question, explanation)
- Graphic User Interface

Stakes

Problematic situation with the optimization system



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

Mathieu Lerouge

Challenges of designing explanation tools for optimization systems 7 October 2021 9 / 22

End-users' issues

- Confusion: in a given solution,
 - presence of unexpected decisions;
 - difference between its quality and the expected one.
- **Frustration**: optimization systems experienced as non-transparent systems / black boxes.
- → **Users loosing confidence** in the optimization system.

Goals

Our proposals

Designing explanation tools, which are **independent** from the solving algorithm, and that enable users:

- to ask questions about a solution (by selecting template questions in a given list) and get explanations back;
- to explore the space of feasible solutions and the space of neighboring instances;
- to identify critical data in the inputs.
- → Tackling users' black-box feeling about the system, increasing users' trust in the system.

Introduction

- Context
- Stakes and goals

In the literature

• Explanation in optimization

- Types of questions
- Types of explanations

Our explanation tool

- Template questions
- Typology of (question, explanation)
- Graphic User Interface

Our observation

Very few works dealing with explanation in optimization

Article	Explanation approach							
Article	Based on	Applied to	Applicable to	Dependance	Questions			
[Ludwig	Solving	Makespan	Specific MSP	Depending on				
et al.,	algorithm	Scheduling	solved via specific	solving	1 type			
2018]	memoriz.	Problem	algorithm	algorithm				
[Čyras et al., 2019]	Abstract Argument.	Makespan Scheduling Problem	Specific problems with binary decision variables	Not depending on solving algorithm	3 types			
[Korikov et al., 2021]	Inverse Optim.	Knapsack, Portfolio	Specific linear problems whose weights in OF are not in constraints	Depending on on solving algorithm	1 type			

[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

Introduction

- Context
- Stakes and goals

In the literature

- Explanation in optimization
- Types of questions
- Types of explanations

Our explanation tool

- Template questions
- Typology of (question, explanation)
- Graphic User Interface

One key type of questions

Contrastive questions



e.g. "Why is Adam not performing the plumbing task in addition to his two other tasks?"

Relevance of working with contrastive questions

- they correspond to most of the "Why" questions people ask [Miller, 2019];
- they tend to specify the question, to narrow the set of solutions to exam.

[Miller, 2019] Explanation in artificial intelligence: Insights from the social sciences

Introduction

- Context
- Stakes and goals

In the literature

- Explanation in optimization
- Types of questions
- Types of explanations

Our explanation tool

- Template questions
- Typology of (question, explanation)
- Graphic User Interface

One key type of explanations

Counterfactual explanations

"To get that decision, the inputs should have been this way?"

counterfact in decisions

counterfact in inputs

e.g. "So that Adam performs the plumbing task in addition to his two other tasks, the electric task should be available 15 minutes earlier".

Benefits of employing counterfactual explanations

- they get around the challenge of explaining the functionality or the rationale of complex algorithmic decision-making systems [Wachter et al., 2018];
- they reveal critical data in the inputs.

[Wachter et al., 2018] Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR

Introduction

- Context
- Stakes and goals

In the literature

- Explanation in optimization
- Types of questions
- Types of explanations

Our explanation tool

Template questions

- Typology of (question, explanation)
- Graphic User Interface

Questions about assigning a new task to an employee "Why is the employee $\langle \mathcal{E}_i \rangle$ not performing the task $\langle \mathcal{T}_k \rangle$...

- ... instead of the task $< T_i > ?"$
- ... instead of one of his/her planning's task?"
- ... just after the task $< T_j >$?"
- ... in addition to his/her planning's tasks?"

Questions about changing the order of a task in a planning "Why is the employee $\langle \mathcal{E}_i \rangle$ not performing the task $\langle \mathcal{T}_k \rangle$...

- ... just after the task $< T_j >$?"
- ... in the portion after the task $< T_j >$?"
- ... at another time of his/her planning?"

Introduction

- Context
- Stakes and goals

In the literature

- Explanation in optimization
- Types of questions
- Types of explanations

Our explanation tool

- Template questions
- Typology of (question, explanation)
- Graphic User Interface

3 types

One template

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

Type (QE_1)

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

Type (QE₂)

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

Type (QE₃)

"How to make it possible for Adam to perform the plumbing

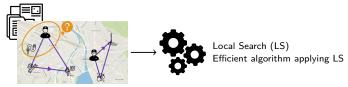
task in addition to his tasks?"

Type (QE_1)

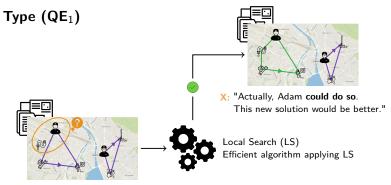


U: "Why is Adam not performing the plumbing task in addition to his tasks - while keeping the order of his planning?"

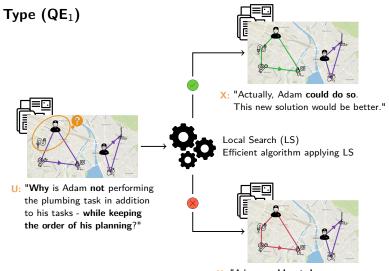
Type (QE_1)



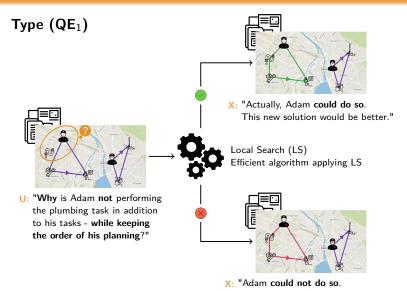
U: "Why is Adam not performing the plumbing task in addition to his tasks - while keeping the order of his planning?"



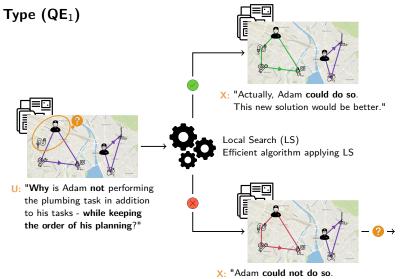
U: "Why is Adam not performing the plumbing task in addition to his tasks - while keeping the order of his planning?"



X: "Adam could not do so. If he would perform the task, at best, he would be at home late by 30min."



 Solutions space exploration
 If he would perform the task, at best, he would be at home late by 30min."



If he would perform the task, at best, he would be at home late by 30min."

Solutions space exploration

Type (QE₂)



U: "Why is Adam not performing the plumbing task in addition to his tasks?"

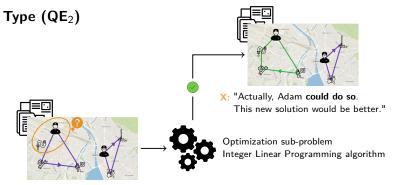


Type (QE₂)

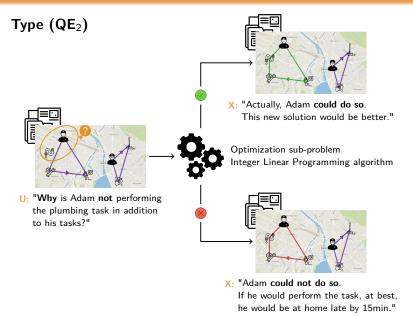


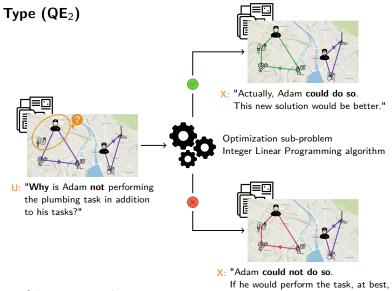
Optimization sub-problem Integer Linear Programming algorithm

U: "Why is Adam not performing the plumbing task in addition to his tasks?"



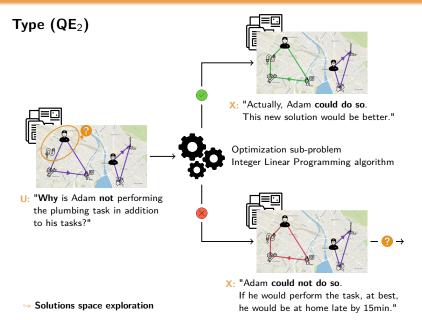
U: "Why is Adam not performing the plumbing task in addition to his tasks?"



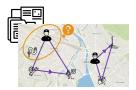


→ Solutions space exploration

he would be at home late by 15min."



Type (QE₃)



U: "How to make it possible for Adam to perform the plumbing task in addition to his tasks?"

Type (QE_3)

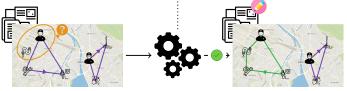
Optimization sub-problem with slacks Integer Linear Programming algorithm



U: "How to make it possible for Adam to perform the plumbing task in addition to his tasks?"

Type (QE₃)

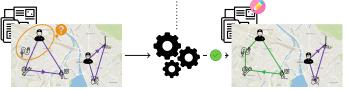
Optimization sub-problem with slacks Integer Linear Programming algorithm



U: "How to make it possible for Adam to perform the plumbing task in addition to his tasks?" X: "So that Adam performs it, the electric task should be available 15 minutes earlier."

Type (QE₃)

Optimization sub-problem with slacks Integer Linear Programming algorithm



U: "How to make it possible for Adam to perform the plumbing task in addition to his tasks?" X: "So that Adam performs it, the electric task should be available 15 minutes earlier."

→ Instances space exploration

Introduction

- Context
- Stakes and goals

In the literature

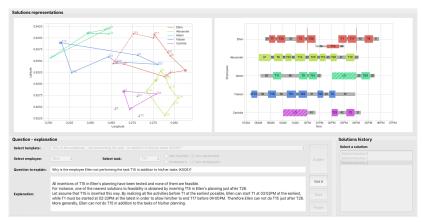
- Explanation in optimization
- Types of questions
- Types of explanations

Our explanation tool

- Template questions
- Typology of (question, explanation)
- Graphic User Interface

Graphic User Interface

A first GUI prototype of WSRP explanation tool



Mathieu Lerouge

Challenges

- How to deal with less local / more global users' questions?
 e.g. "Why Adam is working much less than Ellen?"
- How much **generic** our approach is? How to transpose it to other optimization problems?
- How to structure the exploration of solutions and instances?
- How to make the interaction with the user closer to a dialog?

[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.

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. [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.

[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.