

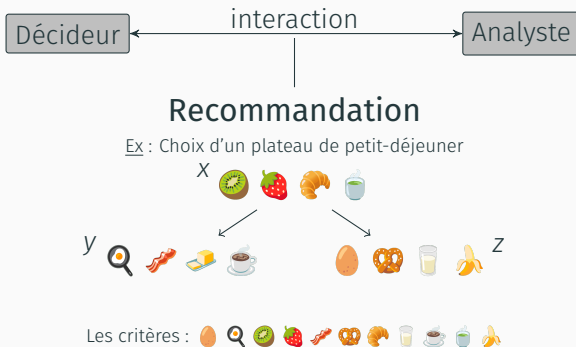
Des explications transitives questionnables au service de l'élicitation de préférences additives

Manuel Amoussou¹, Khaled Belahcene¹,
Nicolas Maudet², Vincent Mousseau¹, Wassila Ouerdane¹

¹MICS, CentraleSupélec, Université Paris-Saclay

²Sorbonne Université, CNRS, Laboratoire d'Informatique de Paris 6

Contexte : AIDE MULTICRITÈRE À LA DÉCISION



La **Recommandation**, comme issue d'un processus d'**Élicitation de préférences**

- Comparaisons par paire d'alternatives de référence

$$\mathbb{PI} = \{r^1 \succ r^2, r^3 \succ r^4, \dots\}$$

- Hypothèse d'un modèle rationnel de représentation des préférences.

👁 Ex : Le modèle additif : $x \succeq y \Leftrightarrow \omega(x) \geq \omega(y)$
 $\omega(\text{🍌} \text{🍓} \text{🍌} \text{🥤}) = \omega[\text{🍌}] + \omega[\text{🍓}] + \omega[\text{🍌}] + \omega[\text{🥤}]$

La **Recommandation**, comme issue d'un processus d'**Élicitation de préférences**

- Comparaisons par paire d'alternatives de référence

$$\mathbb{PI} = \{r^1 \succ r^2, r^3 \succ r^4, \dots\}$$

- Hypothèse d'un modèle rationnel de représentation des préférences.

👁 Ex : Le modèle additif : $x \succeq y \Leftrightarrow \omega(x) \geq \omega(y)$
 $\omega(\text{🍌} \text{🍓} \text{🍊} \text{🍼}) = \omega[\text{🍌}] + \omega[\text{🍓}] + \omega[\text{🍊}] + \omega[\text{🍼}]$

Au programme aujourd'hui,

Explication de la Recommandation ? Non

Explication pour de l'Élicitation ? Oui

Expliquer quoi ? Expliquer Comment ?

Pourquoi questionnables ?

Quid des Explications non questionnables

Explications non questionnables d'une déduction

* Déduction de la relation nécessaire N_{PI}

Ensemble des comparaisons par paire (x, y) vérifiant $\omega(x) \geq \omega(y)$ pour **toute** fonction ω d'évaluation compatible avec PI .

PI (Preference Information) désignant l'ensemble des préférences exprimées par le décideur.

Toute **déduction** de la relation nécessaire **est** (par définition) **non questionnable** sous l'hypothèse de préférences additives.

* Explication¹ non questionnable d'une déduction (x, y)

Ensemble E tel que $(x, y) \in N_E$ avec $E \subseteq N_{PI}$.

Toute explication non questionnable est uniquement composée de déductions (non questionnables) de la PI .

En contester un élément remet donc en cause l'hypothèse de préférences additives.

¹“Expliquer un évènement, c'est fournir des informations sur ses causes.” Miller Tim dans Explanation in Artificial Intelligence: Insights from the Social Sciences, ArXiv,abs/1706.07269,2019.

Explications non questionnables dans la littérature

Preferential reducts² vs. Preference swaps³

$$\mathbb{E} \subseteq \mathbb{PI} \subseteq \mathbb{N}_{\mathbb{PI}}$$

$\mathbb{E}$ de taille minimale.

$$\mathbb{E} \subseteq \mathbb{S}_{\mathbb{PI}} \subseteq \mathbb{N}_{\mathbb{PI}}$$

$\mathbb{E}$ couvrant et de taille minimale.

$\mathbb{S}_{\mathbb{PI}}$: ensemble de déductions de type swaps.

Caractéristiques	PR	PS
Permet d'expliquer toutes les comparaisons déduites	OUI	NON
Explicite les mécanismes en jeu dans la déduction	NON	OUI

²Kadzinski Milosz et al. : *Preferential reducts and constructs in robust multiple criteria ranking and sorting*. Operations Research-Spektrum, 36 :1021-1053, octobre 2014.

³Belahcene Khaled et al. : *Explaining robust additive utility models by preference swaps*. Theory and Decision, 82(2) : 151-183, 2017.

Notre proposition : Explications transitives questionnables

Preferential reducts ET Preference swaps ET

au moins un *swap* n'appartenant pas à $\mathbb{S}_{PI}$, ensemble de déductions de type *swaps*.

Notre intention : Nous servir de l'explication de déductions pour collecter de l'information préférentielle “sûre” qui permettra d'aboutir plus rapidement à la recommandation.

Illustration : Description de la situation d'aide à la décision

	a	b	c	d	e	f
x^{56}	1	1	1	0	0	0
x^{26}	0	1	1	0	1	0
x^{25}	0	1	1	0	0	1
x^{42}	1	0	1	0	1	0
x^{11}	0	0	1	0	1	1
x^{50}	1	1	0	0	1	0
x^{37}	1	0	0	1	0	1
x^7	0	0	0	1	1	1

Description de l'ensemble des alternatives A.

Type de préférences : Additive
Linear Order^a

Type de recommandation : 1—best

^a Fishburn Peter et al.: *Subset Comparisons for Additive Linear Orders*. Mathematics of Operations Research, 27 : 227-243, 2002.

Recherche d'hôtels pour
chercheurs-participants de la PFIA
2024



- ☐ **a** : abordable (1) ou pas (0)
- ☐ **b** : petit-déjeuner inclus (1) ou pas (0)
- ☐ **c** : proche du site de conférences (1) ou pas (0)
- ☐ **d** : ...
- ☐ ...

Illustration : Impact de l'information préférentielle

Proportion de modèles soutenant la recommandation de chacune des alternatives A

En l'absence d'information préférentielle

$$|\Omega| = 124\,187 \times 6! = 89\,414\,640$$

x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7
11.2%	9.9%	13.4%	10.5%	9.6%	12.5%	17.7%	15.2%

Illustration : Impact de l'information préférentielle

Proportion de modèles soutenant la recommandation de chacune des alternatives A

En l'absence d'information préférentielle

$$|\Omega| = 124\,187 \times 6! = 89\,414\,640$$

x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7
11.2%	9.9%	13.4%	10.5%	9.6%	12.5%	17.7%	15.2%

Collecte d'information préférentielle

$$\mathbb{PI}^0 = \{r^{43} \succ r^{24} \succ x^{50} \succ x^7 \succ r^{17}\}$$

$|\Omega|$ passe à 1 992 852

x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7
39.2%	34.6%	6.4%	17.4%	2.4%	0.0%	0.0%	0.0%

Description de l'ensemble A^R des alternatives de référence.

	a	b	c	d	e	f
r^{43}	1	0	1	0	1	1
r^{24}	0	1	1	0	0	0
x^{50}	1	1	0	0	1	0
x^7	0	0	0	1	1	1
r^{17}	0	1	0	0	0	1

L'ensemble des déductions sur A étant donnée

$\mathbb{PI}^0$ est :

$$N_{\mathbb{PI}^0}^A = \{(x^{56}, \underline{x^{50}}), (x^{56}, \underline{x^{37}}), (x^{56}, \underline{x^7}), (x^{26}, \underline{x^{50}}), (x^{26}, \underline{x^7}), (x^{25}, \underline{x^{50}}), (x^{25}, \underline{x^7}), (x^{50}, \underline{x^7})\}$$

x^{50}, x^{37}, x^7 disqualifiés !

Parce que :

x^{50}, x^{37}, x^7 disqualifiés !

Parce que :

$$[x^{56}] \underbrace{111000 \succ 110010}_{[c, e]} [x^{50}]$$

x^{50}, x^{37}, x^7 disqualifiés !

Parce que :

$$[x^{56}] \underbrace{111000 \succ 110010}_{[c, e]} [x^{50}]$$

$$[x^{26}] \underbrace{011010 \succ 011000}_{[e, \emptyset]} \overbrace{[r^{24}] \succ [x^{50}] \succ [x^7]}^{\in PI^0}$$

x^{50}, x^{37}, x^7 disqualifiés !

Parce que :

$$[x^{56}] \underbrace{111000 \succ 110010}_{[c, e]} [x^{50}]$$

$$[x^{26}] \underbrace{011010 \succ 011000}_{[e, \emptyset]} \overbrace{[r^{24}] \succ [x^{50}] \succ [x^7]}^{\in PI^0}$$

$$[x^{56}]111000 \succ 100101[x^{37}]$$

$\equiv$

$$[r^{24}]011000 \succ 000101[f^5]$$

x^{50}, x^{37}, x^7 disqualifiés !

Parce que :

$$[x^{56}] \underbrace{111000 \succ 110010}_{[c, e]} [x^{50}]$$

$$[x^{26}] \underbrace{011010 \succ 011000}_{[e, \emptyset]} \overbrace{[r^{24}] \succ [x^{50}] \succ [x^7]}^{\in PI^0}$$

$$[x^{56}] 111000 \succ 100101 [x^{37}]$$

$\equiv$

$$[r^{24}] 011000 \succ 000101 [f^5]$$

et

$$\overbrace{[r^{24}] \succ [x^{50}] \succ [x^7]}^{\in PI^0} \underbrace{000111 \succ 000101}_{[e, \emptyset]} [f^5]$$

Illustration : Explications et collecte d'information (Étape 1 / 4)

Explication de $(x^{50}, x^7) \in \mathbb{N}_{\mathbb{P}\mathbb{I}^0}$

$$[x^{50}] \underbrace{110010 \succ 010110}_{[a, d]} := \underbrace{010110 \succ 000111}_{[b, f]} [x^7]$$

Avant :

x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7
39.2%	34.6%	6.4%	17.4%	2.4%	0.0%	0.0%	0.0%

Après (potentiels) :

Décideur	[a, d]	[b, f]	Conséquences							
	✓	✓	$x^{56} : 47.6\%$	$x^{26} : 29.5\%$	$x^{25} : 2.8\%$	$x^{42} : 20.1\%$	$x^{11} : 0.0\%$	$x^{50} : 0.0\%$	$x^{37} : 0.0\%$	$x^7 : 0.0\%$
	✓	✗	$x^{56} : 35.5\%$	$x^{26} : 0.0\%$	$x^{25} : 11.0\%$	$x^{42} : 42.5\%$	$x^{11} : 10.9\%$	$x^{50} : 0.0\%$	$x^{37} : 0.0\%$	$x^7 : 0.0\%$
	✗	✓	$x^{56} : 35.0\%$	$x^{26} : 55.4\%$	$x^{25} : 6.7\%$	$x^{42} : 2.9\%$	$x^{11} : 0.0\%$	$x^{50} : 0.0\%$	$x^{37} : 0.0\%$	$x^7 : 0.0\%$
	✗	✗	$x^{56} : 0.0\%$	$x^{26} : 0.0\%$	$x^{25} : 0.0\%$	$x^{42} : 0.0\%$	$x^{11} : 0.0\%$	$x^{50} : 0.0\%$	$x^{37} : 0.0\%$	$x^7 : 0.0\%$

Illustration : Explications et collecte d'information (Étape 1 / 4)

Explication de $(x^{50}, x^7) \in \mathbb{N}_{\mathbb{PI}^0}$

$$[x^{50}] \underbrace{110010}_{[a, d]} \succ \underbrace{010110}_{[b, f]} := \underbrace{010110}_{[b, f]} \succ \underbrace{000111}_{[x^7]} [x^7]$$

Avant :

x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7
39.2%	34.6%	6.4%	17.4%	2.4%	0.0%	0.0%	0.0%

Après (potentiels) :

Décideur	[a, d]	[b, f]	Conséquences							
	✓	✓	$x^{56} : 47.6\%$	$x^{26} : 29.5\%$	$x^{25} : 2.8\%$	$x^{42} : 20.1\%$	$x^{11} : 0.0\%$	$x^{50} : 0.0\%$	$x^{37} : 0.0\%$	$x^7 : 0.0\%$
	✓	✗	$x^{56} : 35.5\%$	$x^{26} : 0.0\%$	$x^{25} : 11.0\%$	$x^{42} : 42.5\%$	$x^{11} : 10.9\%$	$x^{50} : 0.0\%$	$x^{37} : 0.0\%$	$x^7 : 0.0\%$
	✗	✓	$x^{56} : 35.0\%$	$x^{26} : 55.4\%$	$x^{25} : 6.7\%$	$x^{42} : 2.9\%$	$x^{11} : 0.0\%$	$x^{50} : 0.0\%$	$x^{37} : 0.0\%$	$x^7 : 0.0\%$
	✗	✗	$x^{56} : 0.0\%$	$x^{26} : 0.0\%$	$x^{25} : 0.0\%$	$x^{42} : 0.0\%$	$x^{11} : 0.0\%$	$x^{50} : 0.0\%$	$x^{37} : 0.0\%$	$x^7 : 0.0\%$

Mise à jour : $\mathbb{PI}^1 = \mathbb{PI}^0 \cup \{[a, d], [b, f]\}$

On remarque que : $(x^{26}, x^{11}) \in \mathbb{N}_{\mathbb{PI}^1}^A$

Illustration : Explications et collecte d'information (Étape 2 / 4)

Explication de $(x^{56}, x^7) \in \mathbb{N}_{\text{PT}1}$

$$[x^{56}] \underbrace{111000}_{[a, f]} \succ \underbrace{011001}_{[b, e]} := \underbrace{011001}_{[b, e]} \succ \underbrace{001011}_{[c, d]} := \underbrace{001011}_{[c, d]} \succ \underbrace{000111}_{[c, d]} [x^7]$$

Avant :

x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7
47.6%	29.5%	2.8%	20.1%	0.0%	0.0%	0.0%	0.0%

Après (potentiels) :

Décideur	[a, f]	[b, e]	[c, d]	Conséquences			
	✓	✓	✓	$x^{56} : 86.6\%$	$x^{26} : 13.4\%$	$x^{25} : 0.0\%$	$x^{42} : 0.0\%$
	✓	✗	✓	$x^{56} : 0.0\%$	$x^{26} : 33.1\%$	$x^{25} : 0.0\%$	$x^{42} : 66.9\%$
	✗	✓	✓	$x^{56} : 0.0\%$	$x^{26} : 57.9\%$	$x^{25} : 42.1\%$	$x^{42} : 0\%$
	✗	✗	✓	$x^{56} : 0.0\%$	$x^{26} : 100\%$	$x^{25} : 0.0\%$	$x^{42} : 0.0\%$

Illustration : Explications et collecte d'information (Étape 2 / 4)

Explication de $(x^{56}, x^7) \in \mathbb{N}_{\text{PT1}}$

$$[x^{56}] \underbrace{111000}_{[a, f]} \succ \underbrace{011001}_{[b, e]} := \underbrace{011001}_{[b, e]} \succ \underbrace{001011}_{[c, d]} := \underbrace{001011}_{[c, d]} \succ \underbrace{000111}_{[c, d]} [x^7]$$

Avant :

x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7
47.6%	29.5%	2.8%	20.1%	0.0%	0.0%	0.0%	0.0%

Après (potentiels) :

Décideur	[a, f]	[b, e]	[c, d]	Conséquences			
	✓	✓	✓	$x^{56} : 86.6\%$	$x^{26} : 13.4\%$	$x^{25} : 0.0\%$	$x^{42} : 0.0\%$
	✓	✗	✓	$x^{56} : 0.0\%$	$x^{26} : 33.1\%$	$x^{25} : 0.0\%$	$x^{42} : 66.9\%$
	✗	✓	✓	$x^{56} : 0.0\%$	$x^{26} : 57.9\%$	$x^{25} : 42.1\%$	$x^{42} : 0\%$
	✗	✗	✓	$x^{56} : 0.0\%$	$x^{26} : 100\%$	$x^{25} : 0.0\%$	$x^{42} : 0.0\%$

Illustration : Explications et collecte d'information (Étape 3 / 4)

Explication de $(x^{42}, x^7) \in \mathbb{N}_{\mathbb{PI}^2}$

$$[x^{42}] \underbrace{101010 \succ 011010}_{[a, b]} := \underbrace{011010 \succ 011000}_{[d, \emptyset]} \overbrace{[r^{24}] \succ [x^{50}] \succ [x^7]}^{\in \mathbb{PI}^2}$$

Avant :

x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7
86.6%	13.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Après (potentiels) :

Décideur	[a, b]	Conséquences	
	✓	$x^{56} : 100\%$	$x^{26} : 0.0\%$
	✗	$x^{56} : 53.9\%$	$x^{26} : 46.1\%$

Illustration : Explications et collecte d'information (Étape 3 / 4)

Explication de $(x^{42}, x^7) \in \mathbb{N}_{\text{PI}^2}$

$$[x^{42}] \underbrace{101010 \succ 011010}_{[a, b]} := \underbrace{011010 \succ 011000}_{[d, \emptyset]} \overbrace{[r^{24}] \succ [x^{50}] \succ [x^7]}^{\in \text{PI}^2}$$

Avant :

x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7
86.6%	13.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Après (potentiels) :

Décideur	[a, b]	Conséquences	
	✓	$x^{56} : 100\%$	$x^{26} : 0.0\%$
☞	✗	$x^{56} : 53.9\%$	$x^{26} : 46.1\%$

Illustration : Explications et collecte d'information (Étape 4 / 4)

Explication de $(x^{56}, x^{37}) \in \mathbb{N}_{\mathbb{P}T^3}$

$$[x^{56}] \underbrace{111000 \succ 110010}_{[c, e]} \overbrace{[x^{50}] \succ [x^7]}^{\in \mathbb{P}T^3} \underbrace{000111 \succ 100101}_{[e, a]} [x^{37}]$$

Avant :

x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7
53.9%	46.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Après (potentiels) :

Décideur	[c, e]	[e, a]	Conséquences	
	✓	✓	$x^{56} : 0.0\%$	$x^{26} : 100\%$
	✓	✗	$x^{56} : 100\%$	$x^{26} : 0.0\%$

Illustration : Explications et collecte d'information (Étape 4 / 4)

Explication de $(x^{56}, x^{37}) \in \mathbb{N}_{\mathbb{P}I^3}$

$$[x^{56}] \underbrace{111000 \succ 110010}_{[c, e]} \overbrace{[x^{50}] \succ [x^7]}^{\in \mathbb{P}I^3} \underbrace{000111 \succ 100101}_{[e, a]} [x^{37}]$$

Avant :

x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7
53.9%	46.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Après (potentiels) :

Décideur	[c, e]	[e, a]	Conséquences	
	✓	✓	$x^{56} : 0.0\%$	$x^{26} : 100\%$
	✓	✗	$x^{56} : 100\%$	$x^{26} : 0.0\%$

L'alternative à recommander est donc : x^{56}

Illustration : le résumé des interactions

Évolution de l'éllicitation								$\mathbb{PI}^0$ & Réactions du décideur
x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7	$\mathbb{PI}^0 = \{r^{43} \succ r^{24} \succ x^{50} \succ x^7 \succ r^{17}\}$
11.2%	9.9%	13.4%	10.5%	9.6%	12.5%	17.7%	15.2%	

Illustration : le résumé des interactions

Évolution de l'éllicitation								$\mathbb{PI}^0$ & Réactions du décideur
x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7	$\mathbb{PI}^0 = \{r^{43} \succ r^{24} \succ x^{50} \succ x^7 \succ r^{17}\}$
11.2%	9.9%	13.4%	10.5%	9.6%	12.5%	17.7%	15.2%	
x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7	[a, d] : ✓ [b, f] : ✓
39.2%	34.6%	6.4%	17.4%	2.4%	0.0%	0.0%	0.0%	

Illustration : le résumé des interactions

Évolution de l'éllicitation								$\mathbb{PT}^0$ & Réactions du décideur	
x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7	$\mathbb{PT}^0 = \{r^{43} \succ r^{24} \succ x^{50} \succ x^7 \succ r^{17}\}$	
11.2%	9.9%	13.4%	10.5%	9.6%	12.5%	17.7%	15.2%		
x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7	[a, d] : ✓	[b, f] : ✓
39.2%	34.6%	6.4%	17.4%	2.4%	0.0%	0.0%	0.0%		
x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7	[a, f] : ✓ [b, e] : ✓ [c, d] : ✓	
47.6%	29.5%	2.8%	20.1%	0.0%	0.0%	0.0%	0.0%		

Illustration : le résumé des interactions

Évolution de l'éllicitation								$\mathbb{PI}^0$ & Réactions du décideur	
x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7	$\mathbb{PI}^0 = \{r^{43} \succ r^{24} \succ x^{50} \succ x^7 \succ r^{17}\}$	
11.2%	9.9%	13.4%	10.5%	9.6%	12.5%	17.7%	15.2%		
x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7	[a, d] : ✓ [b, f] : ✓	
39.2%	34.6%	6.4%	17.4%	2.4%	0.0%	0.0%	0.0%		
x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7	[a, f] : ✓ [b, e] : ✓ [c, d] : ✓	
47.6%	29.5%	2.8%	20.1%	0.0%	0.0%	0.0%	0.0%		
x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7	[a, b] : ✗	
86.6%	13.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%		
x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7	[c, e] : ✓ [e, a] : ✗	
53.9%	46.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%		

Illustration : le résumé des interactions

Évolution de l'éllicitation								$\mathbb{PI}^0$ & Réactions du décideur	
x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7	$\mathbb{PI}^0 = \{r^{43} \succ r^{24} \succ x^{50} \succ x^7 \succ r^{17}\}$	
11.2%	9.9%	13.4%	10.5%	9.6%	12.5%	17.7%	15.2%		
x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7	$[a, d] : \checkmark$ $[b, f] : \checkmark$	
39.2%	34.6%	6.4%	17.4%	2.4%	0.0%	0.0%	0.0%		
x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7	$[a, f] : \checkmark$ $[b, e] : \checkmark$ $[c, d] : \checkmark$	
47.6%	29.5%	2.8%	20.1%	0.0%	0.0%	0.0%	0.0%		
x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7	$[a, b] : \times$	
86.6%	13.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%		
x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7	$[c, e] : \checkmark$ $[e, a] : \times$	
53.9%	46.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%		
x^{56}	x^{26}	x^{25}	x^{42}	x^{11}	x^{50}	x^{37}	x^7	Recommander x^{56} !!!	
100%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%		

Ébauche de protocole interactif

Squelette d'un protocole interactif

Entrées. A , A^R et $\mathbb{PI}^0$.

$t = 0$

Tant que $|\mathbb{N}_{\mathbb{PI}^t}^A| \neq 0$ et **ConditionArret** non vérifiée

Choisir $(x, y) \in \mathbb{N}_{\mathbb{PI}^t}^A$

 Tenter d'expliquer (x, y)

 Si l'explication existe :

 Collecter S^t (Les réactions du décideur)

 Mises à jour :

 ♦ $\mathbb{PI}^{t+1} = \mathbb{PI}^t \cup S^t$

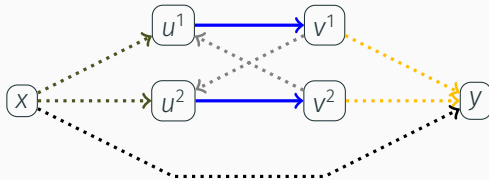
 ♦ $t \leftarrow t + 1$

 Sinon retirer (x, y) de $\mathbb{N}_{\mathbb{PI}^t}^A$

- Ressource critique : les déductions de la relation nécessaire
- **ConditionArret** sur la durée de l'interaction ou l'évolution du nombre de recommandations admissibles (seuil, convergence) ...
- **Choisir** : stratégie ou déduction désignée explicitement par le décideur
- Choix de l'explication de $(x, y) \in \mathbb{N}_{\mathbb{PI}^t}^A$ quand il en existe plusieurs.

Formalisation et Modélisation

Formalisation du problème



$$\begin{array}{ccccc} E_1 : \text{dotted green} \rightarrow & E_4 : \text{solid blue} \rightarrow & E_2 : \text{dotted orange} \rightarrow \\ E_3 : \text{dotted black} \rightarrow & E_5 : \text{dotted black} \rightarrow & \end{array}$$

Rechercher un chemin dans le Graphe d'explication $\mathcal{G}^{\text{PI}}_{(x,y)}$
(une instantiation possible : le chemin(-explication) le plus court)

Modélisation à l'aide de la PLNE

Variables

- w_i continues (scores marginaux des critères)
- $\text{arc}_{(u,v)}$ binaires indiquant les arcs(-comparaisons) retenus sur le chemin-explication.
- $s_{(u,v)}^{[i,j]}$ binaires valant 1 si le swap $[i, j]$ participe à la décomposition de la comparaison (u, v)

Contraintes

- Fonction de score normalisée et positivité de ses composantes :

$$\sum_{i \in [m]} w_i = 1 \quad (1)$$

$$w_i \geq \epsilon \quad \forall i \in [m] \quad (2)$$

- Prise en compte de la $\mathbb{PI}$

$$\sum_{i \in \text{pro}(u,v)} w_i \geq \sum_{j \in \text{con}(u,v)} w_j \quad \forall (u,v) \in \mathbb{PI} \quad (3)$$

- Contraintes de "chemin"

$$\sum_{(x,u) \in E} \text{arc}_{(x,u)} = 1 \quad (4)$$

$$\sum_{(u,z) \in E} \text{arc}_{(u,z)} = \sum_{(z,v) \in E} \text{arc}_{(z,v)} \quad \forall z \in V \setminus \{x, y\} \quad (5)$$

- Contraintes liant les variables $s_{(u,v)}^{[i,j]}$ et $\text{arc}_{(u,v)}$

$$\sum_{i \in \text{pro}(u,v)} s_{(u,v)}^{[i,j]} = \text{arc}_{(u,v)} \quad \forall j \in \text{con}(u,v) \quad (6)$$

$$\sum_{i \in \text{con}(u,v)} s_{(u,v)}^{[i,j]} \leq \text{arc}_{(u,v)} \quad \forall i \in \text{pro}(u,v) \quad (7)$$

- Chemin-explication questionnable

$$\sum_{(u,v) \in E \setminus E_4} \sum_{[i,j] \in \tilde{\mathbb{PI}}} s_{(u,v)}^{[i,j]} \geq 1 \quad (8)$$

$\tilde{\mathbb{PI}}$: ensemble des swaps non déduits.

- Contraintes liant les variables $s_{(u,v)}^{[i,j]}$ à w_i et w_j

$$w_i - w_j \geq (1 + \epsilon) \times s_{(u,v)}^{[i,j]} - 1 \quad (9)$$

avec ϵ un réel positif arbitrairement petit.

Fonction objectif

$$\sum_{(u,v) \in E} \text{arc}_{(u,v)} \times g((u,v))$$

avec $g(\cdot)$ la fonction de pondération des arcs(-comparaisons).

- Instanciation précise du protocole en fonction du contexte
 - Absence de déduction (provocation, suggestion)
 - Stratégie de choix de déductions et/ou d'explications
 - Changer de mécanismes de support
- Évaluation empirique avec un décideur humain

Merci!