# Explication de recommandations issues d'un modèle additif : de la conceptualisation à l'évaluation

Manuel Amoussou[1],Khaled Belahcene[2],Nicolas Maudet[3],
Vincent Mousseau[1],Wassila Ouerdane[1]

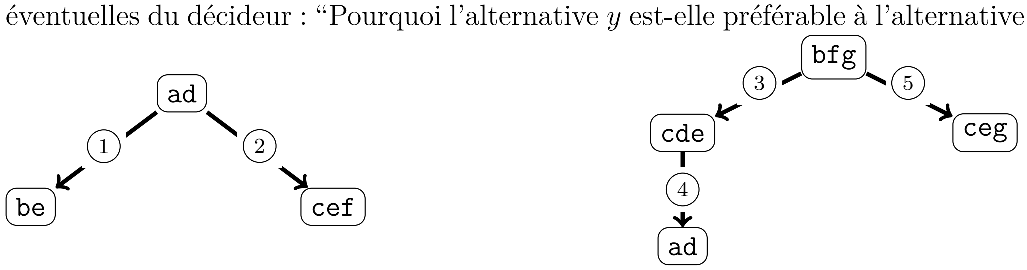
[1] Université Paris-Saclay, CentraleSupélec, MICS, Gif-Sur-Yvette, France
{manuel.amoussou,vincent.mousseau,wassila.ouerdane}@centralesupelec.fr
[2] Université de Technologie de Compiègne, CNRS, UMR 7253 Heudiasyc, Compiègne, France
khaled.belahcene@hds.utc.fr
[3] Sorbonne Université, CNRS, Laboratoire d'Informatique de Paris 6, LIP6, France
nicolas.maudet@lip6.fr

**Mots-clés** : *Aide Multicritère à la Décision ; Explication ; Modèle additif ; PLNE*

## 1 Introduction et Contexte

Nous nous intéressons à la structure et au calcul automatique de l'explication que devra fournir un analyste à un décideur qui l'interroge sur la recommandation qu'il vient de lui faire. Notre proposition s'inscrit dans le cadre de l'Aide MultiCritère à la Décision (AMCD) où le décideur souhaite opérer le choix de la meilleure alternative $x^*$ parmi un ensemble fini d'alternatives $\mathbb{A}$ décrites sur $m$ critères (tous) binaires c'est-à-dire ayant exactement deux niveaux d'évaluation ("fort" et "faible"). Cette hypothèse (binarité des critères) permet de représenter chaque alternative sous la forme du sous-ensemble de critères sur lesquels elle est évaluée au niveau "fort". Par ailleurs, les préférences du décideur étant additives, elles sont donc représentables par une fonction de score $\omega = \langle \omega_i \rangle_{i \in [m]}$ qui associe à chaque critère $i$ une valeur marginale positive $\omega_i$. Le score globale d'une alternative est obtenu en sommant les valeurs marginales des critères qui la composent ; la meilleure alternative $x^*$ étant celle ayant le score global le plus élevé.

Justifier pourquoi le décideur devrait choisir l'alternative $x^*$ revient à exhiber une arborescence dont la racine est occupée par $x^*$ et dont les arcs symbolisent les préférences entre paires d'alternatives de l'ensemble $\mathbb{A}$. De telles arborescences peuvent diverger suivant leur forme caractérisable par leur hauteur (voir figures ci-dessous) et mettent en évidence les comparaisons par paires d'alternatives qui servent à légitimer le choix de $x^*$. Ces comparaisons par paires, compatibles avec $\omega$, constituent les objets de base auxquels s'appliqueront les interrogations éventuelles du décideur : "Pourquoi l'alternative $y$ est-elle préférable à l'alternative $z$ ($y \succsim z$) ?"




$$\omega = \{\texttt{a} : 128,\ \texttt{b} : 126,\ \texttt{c} : 77,\ \texttt{d} : 59,\ \texttt{e} : 52,\ \texttt{f} : 41,\ \texttt{g} : 37\}$$

Pour répondre à cette question, nous nous inspirons des travaux de [2] qui proposent l'explication à l'aide de *preference-swaps* de la comparaison par paire $y \succsim z$ induite par calcul de la relation nécessaire. Cette contribution qui elle-même, tire sa source des *even-swaps* [1], développe une conception de l'explication qui s'attaque aux causes de la complexité de la prise de décision dans un contexte multicritère : la multiplicité et la conflictualité des critères. Elle propose ainsi une explication de la préférence $y \succsim z$ sous la forme d'une séquence de confrontations entre paires de critères (d'où le terme *preference-swaps*), lesquelles confrontations ont de bonnes chances d'être comprises du décideur vu qu'elles expriment la préférence (atomique) d'un (seul) critère par rapport à un (seul) autre.

La présente contribution généralise cette idée en proposant des explications par décomposition de la comparaison par paire $y \succsim z$ en sous-confrontations entre des sous-ensembles de critères qui, pour des besoins d'intelligibilité, sont contraintes syntaxiquement. Quatre langages de décomposition ont ainsi été définis : $\mathcal{L}_0$ ou $\Delta(1,\, 1)$, $\mathcal{L}_1$ ou $\Delta(1,\, m)$, $\mathcal{L}_2$ ou $\Delta(m,\, 1)$, $\mathcal{L}_3$ ou $\mathcal{L}_1 \cup \mathcal{L}_2$ ; l'ensemble $\Delta(p,\, q)$ regroupant les préférences d'un sous-ensemble d'au plus $p$ critères sur un autre de cardinalité au plus $q$ critères. On identifiera parmi ces quatre langages, celui ($\mathcal{L}_0$) des *preference-swaps* décrits précédemment et remarquera que les paires de sous-ensembles de critères admises par ces langages sont non décomposables car composées d'au moins un sous-ensemble de type singleton (voir figures ci-dessous : les cas 1 à 4). Concrètement, sont développées dans la contribution, la modélisation du problème de calcul de l'explication de $y \succsim z$ étant donné $\omega$ ainsi que l'étude de sa complexité algorithmique, celle du calcul de l'arborescence justifiant la recommandation et dont tous les arcs sont décomposables et l'ébauche de pistes de solutions lorsqu'une telle arborescence n'existe pas.

| 1 | 2 | 4 | 3 | 5 |
|---|---|---|---|---|
| $\mathtt{a} \succsim \mathtt{b}$ <br> $\mathtt{d} \succsim \mathtt{e}$ | $\mathtt{a} \succsim \mathtt{cf}$ <br> $\mathtt{d} \succsim \mathtt{e}$ | $\mathtt{ce} \succsim \mathtt{a}$ | $\mathtt{b} \succsim \mathtt{de}$ <br> $\mathtt{fg} \succsim \mathtt{c}$ | $\mathtt{bf} \succsim \mathtt{ce}$ |
| $\mathcal{L}_0$−décomposable | $\mathcal{L}_1$−décomposable | $\mathcal{L}_2$−décomposable | $\mathcal{L}_3$−décomposable | non $\mathcal{L}_3$−décomposable |

## 2 Modélisation et Complexité algorithmique

L'existence d'une explication de la préférence $y \succsim z$ dans chacun des quatre langages considérés peut être formulée algorithmiquement sous la forme d'un problème de décision dont l'instance est décrite par la donnée des critères composant les alternatives de la comparaison par paire et de la fonction de score $\omega$. Les questions associées reviennent en fonction du langage $\mathcal{L}_u$, $u \in \{0,\, 1,\, 2,\, 3\}$, à rechercher l'existence d'une fonction de l'ensemble $2^{\text{PRO}}$ (resp. $2^{\text{CON}}$) vers l'ensemble $2^{\text{CON}}$ (resp. $2^{\text{PRO}}$) où $\text{PRO} = y \backslash z$ et $\text{CON} = z \backslash y$ de sorte à garantir les trois propriétés suivantes : *(i)* **Couverture** des critères des ensembles PRO et CON, *(ii)* **Disjonction** des sous-confrontations composant l'explication (pas de superposition), et *(iii)* **Compatibilité** avec $\omega$ des préférences exprimées par les sous-confrontations explicatives. On démontre que le problème de l'existence d'une explication de la comparaison $y \succsim z$ est NP-difficile pour tous les langages sauf pour $\mathcal{L}_0$ où le problème appartient à la classe P. Le calcul de l'explication s'effectue par la résolution d'un Programme Linéaire en Nombres Entiers (PLNE) de taille polynomiale et contenant des variables binaires qui permettent de récupérer, lorsqu'il admet une solution, les sous-confrontations explicatives. Quant au calcul de l'arborescence explicative de la recommandation, il peut se faire séquentiellement en utilisant le PLNE calculant l'explication d'une comparaison ou plutôt à l'aide d'un PLNE contenant par ailleurs d'autres variables binaires et des contraintes destinées à garantir la structure arborescente recherchée.

## Références


[1] J. S. Hammond, R. L Keeney, and H. Raiffa. Even Swaps : A Rational Method for Making Trade-offs. *Harvard business review*, 76 :137–8, 1998.

[2] K. Belahcène, Ch. Labreuche, N. Maudet, V. Mousseau and W. Ouerdane. Explaining robust additive utility models by sequences of preference swaps. *Theory and Decision*, 10.1007/s11238-016-9560-1, 2017.