Problème de détection de connivence

Introduction

Soit un ensembles A d'alternatives, ayant un ensemble de préférences R sur le powerset, chaque préférence pouvant s'écrire de la forme $a \succeq b$ avec $a, b \in 2^A$.

Souvent afin d'effectuer une élicitation incrémentale, on suppose que les préférences sont représentables par une fonction f_θ i.e $\forall x, y \in 2^A$; $a \succeq b \iff f_\theta(a) \geq f_\theta(b)$.

Et donc généralement on suppose que la fonction f_θ est par exemple 1-additive, cependant, si par la suite l'élicitation révèle des synergies entre des pairs objets ça implique que la fonction ne peut plus représenter la préférence et donc il faut passer à une fonction 2-additive dans laquelle θ va contenir des termes qui représentent les synergies entre 2 alternatives, cependant, même si il existe une unique synergie entre 2 éléments la fonction 2-additives contiendra tous les termes qui auraient pu représenter des synergies entre d'autres paires et qui, a ce stade, n'existent pas.

Dans la suite il s'agit de conçevoir un protocole visant a étendre θ afin de l'adapter a la présence de synergies et donc de détecter des sous-ensembles de préférences qui démontrent une interaction non prise en compte par f_θ, on appelera ces sous-ensembles **ensemble connivants**.

Formalisation du probléme

Dans la suite on note:

- A L'ensemble des alternatives $a_1, a_2, \ldots, a_n$
- θ un ensemble de paramètres chacun représentant un sous-ensemble de A dont on sait qu'il y à interaction.
- f_θ une fonction n-additive ou seule les paramètres dans θ sont non nuls, par exemple, si $A = \{1, 2, 3, 4\}$ et $\theta = \{(1,), (4,), (3, 4, 1)(1, 2), (3,)\}$ alors $f_\theta(\{1, 2, 3\}) = u_1 + u_3 + u_{12}$

Et donc le problème ADDITIVE-1 se définit comme:

- **Données**: Un ensemble R de préférences.
- **Question**: Les préférences sont elles représentables par une fonction f_θ qui soit 1-additive c'est a dire ou θ contient uniquement des singletons ?

Et ce problème se généralise a n'importe quel θ comme suit:

- **Données**: Un ensemble R de préférences, θ un ensemble de sous-ensembles.
- **Question**: Les préférences sont elles représentables par une fonction additive f_θ ou les seuls coefficients non nuls sont inclus dans θ?

Donc ADDITIVE-1 est un cas particulier de ADDITIVE-θ avec θ qui contient tous les singletons, et de la même manière ADDITIVE-2 est un cas particulier où θ contient tous les singletons et toutes les paires.

Condition de Fishburn

Fisburn décrit dans son article [?] une condition nécessaire et suffisante à l'existence d'une fonction 1-additive qui représente un ensemble de préférences R. Cette condition peut se formuler comme suit :

Si il existe deux ensembles de sous-ensembles $(A_1, A_2, \ldots, A_j), (B_1, B_2, \ldots, B_j)$ avec $j \geq 2$, tel que $A = (A_1, A_2, \ldots, A_j) \approx_1 B = (B_1, B_2, \ldots, B_j)$ avec $\approx_1$ une relation qui signifie que chaque singleton est présent autant de fois dans A que B alors si $\exists i; A_i \succeq B_i$ alors $\exists j; B_j \succeq A_j$

Formulé autrement on en déduit que si on trouve deux sous-ensembles de sous-ensembles $A_1, A_2, \ldots, A_n$ et $B_1, B_2, \ldots, B_n$ qui contiennent le même nombre de chaque singleton et que $\forall i; A_i \succeq B_i$ alors l'ensemble de préférences n'est pas représentable une fonction 1-additive.

On peut généraliser ce résultat a un θ quelconque mais d'abord définissons pour $\omega \in \theta$ et $S \subset 2^A$:

$$in(\omega, S) = \begin{cases} 1 & : \omega \subseteq S \\ 0 & : Sinon \end{cases}$$

Ensuite définissons pour un sous ensemble de sous ensembles $K = (A_0, A_1, \ldots, A_j) \subset 2^{2^A}$

$$n(\omega, K) = \sum_i in(w, A_i)$$

Alors on peut définir $\approx_\theta$ par:

$$A = (A_0, \ldots, A_j) \approx_\theta B = (B_0, \ldots, B_j) \iff \forall \omega \in \theta; n(\omega, A) = n(\omega, B)$$

En d'autres termes chaque sous-ensemble ayant un terme d'interaction dans θ est présent en même nombre en somme dans A et dans B.

Et donc extraire un le sous ensemble inclus dans R de la forme $A_i \succeq B_i$ avec $A \approx_\theta B$ est un certificat du fait que les préférences de R ne sont pas représentables avec une fonction additive qui aurait pour seuls paramètres non nuls ceux dans θ, on désignera ce sous-ensemble : sous-ensemble connivant dans R pour θ.

Partant de la on définit le problème suivant CONNIVENCE-θ

- **Données**: Ensemble de relations R, ensemble de paramètres θ.
- **Question**: Existe-il un sous-ensemble connivant dans R pour θ?

Etude polyédrale

Avant tout, essayons de représenter l'impact de la sélection de chaque contrainte sur le rapport entre $n(\omega, A)$ et $n(\omega, B)$ pour les différents $\omega \in \theta$.

Pour se faire on défini les coefficients $\forall r \in R; r = (A \succeq B); \forall \omega \in \theta$; c^r_ω par $c^r_\omega = n(\omega, A) - n(\omega, B)$

Et donc $\forall r \in R$, $\omega \in \theta$, $r = (A \succeq B)$

$$c^r_\omega = \begin{cases} 1 & : \omega \in A \land \neg\omega \in B \\ 0 & : (\neg\omega \in A \land \neg\omega \in B) \lor (\omega \in A \land \omega \in B) \\ -1 & : \neg\omega \in A \land \omega \in B \end{cases}$$

Par exemple la contrainte $\{1, 2\} \succeq \{2, 3\}$ avec $\theta = \{1, 2, 3\}$ est représentée par le vecteur $[1, 0, -1]$.

On défini par la suite le vecteur $c^r = [c^r_{w_1}, c^r_{w_2}, \ldots, c^r_{w_k}]$ qui represent l'impact du choix de la contrainte r et ou tous les coefficients valent $0, 1$ où -1.

De manière analogue on définit: $c^{w_i} = [c^{r_1}_{w_i}, c^{r_2}_{w_i}, \ldots, c^{r_j}_{w_i}]$, et ainsi un ensemble connivant est un sous ensemble de lignes tel que $\sum_r c^{w_i}_r = 0 \, \forall w_i \in \theta$.

Preuve de NP-Completude

Dans cette partie on va prouver que le probléme de calcul d'un sous-ensemble connivant est NP-Complet.

D'abord montrons que CONNIVENCE-$\theta \in NP$.

Etant donné un ensemble R de contraintes, un ensemble A d'alternatives et un ensemble θ tel que $|R| = n$, $|A| = k$ et $|\theta| = m$; pour vérifier qu'un sous-ensemble de contrainte R' est connivant il suffit de calculer les coefficients c^r_ω, pour se faire on itére sur chaque contrainte de la forme $a_i \succeq b_i$ en vérifiant la présence de chaque paramètre ce qui se fait en $O(mk)$ et donc le faire sur tout R prend au plus $O(mnk)$.
Une fois les coefficients crées, il suffit de faire une somme sur les colonnes en $O(m)$ et ensuite de tester la nullité du vecteur obtenu en $O(n)$ la complexité globale est donc de $O(mnk)$ ce qui en fait un probléme NP.

Maintenant réduisons le probléme CONNIVENCE-θ vers le probléme SUBSET SUM défini par

- **Données**: Un ensemble d'entier relatifs E
- **Décision**: Existe il un sous-ensemble $E' \subset E$ tel que $\sum_{a \in E'} a = 0$?

Pour se faire il suffit de trouver un moyen de transformer une instance E de SUBSET SUM en une matrice de coefficients c^r_w de façon a ce que trouver un ensemble connivant revienne a trouver un sous-ensemble $E' \subset E$ qui somme a 0.

Notons $\forall e \in \mathbf{Z}$, $b(e)$ la representation binaire de l'entier $|e|$ et $b^-(e)$ cette même représentation mais en remplaçant les 1 par des -1.

A présent nous allons construire une matrice de dimensions $(|E|, log(max(E)))$ qui va contenir pour chaque entier une ligne représentant $b(e)$ si $e > 0$ et $b^-(e)$ sinon.

A présent il est clair que trouver un sous-ensemble de lignes qui somme a 0 reviens a résoudre le problème d'ensembles connivants.