Proportion analogique et raisonnement à partir de cas

Henri Prade
IRIT, Toulouse
France

Vue d'ensemble

- Proportion analogique:

  « A est à B ce que C est à D »

  A, B, C, D vecteurs (Booléens)

- Expression logique vraie pour chaque composante

  $\Leftrightarrow$ la proportion analogique est satisfaite

- Extensions pour des composantes $\in [0,1]$

- Applications au **raisonnement à partir de cas**

Propriétés requises pour les proportions analogiques

- **(a : b :: c : d) = (c : d :: a : b)** symétrie
- **(a : b :: c : d) = (a : c :: b : d)** permutation centrale
- **(a : a :: b : x) ⇒ x = b** « identité », déterminisme

cependant, insuffisant pour une caractérisation complète

- Conséquences immédiates :

  (a : b :: c : d) = (b : a :: d : c)

  (a : b :: c : d) = (d : b :: c : a)

  (a : b :: c : d) = (d : c :: b : a)

  (a : b :: c : d) = (b : d :: a : c)

  (a : b :: c : d) = (c : a :: d : b)

  (a : b :: a : x) ⇒ x = b

	a	b	c	d	Propor. Anal. **(a : b :: c: d)**	
1	**1**	**1**	**1**	**1**	OUI	**1**
2	**1**	**1**	**0**	**0**	OUI	**1**
3	**1**	**0**	**1**	**0**	OUI	**1**
4	**1**	**0**	**0**	**1**	NON	0
5	**0**	**1**	**1**	**0**	NON	0
6	**0**	**1**	**0**	**1**	OUI	**1**
7	**0**	**0**	**1**	**1**	OUI	**1**
8	**0**	**0**	**0**	**0**	OUI	**1**
9-16 autres 4-uplets					NON	0

a = b iff c = d

a > b iff c > d

(a : b :: c: d) ≠ (a ≡ b) ≡ (c ≡ d) (S. Klein, 1982)

3 formes équivalentes de la bonne définition

- $(a : b :: c : d) = ((a \equiv b) \equiv (c \equiv d)) \wedge ((a \,\Delta\, b) \rightarrow (a \equiv c))$
- $(a : b :: c : d) = ((a \equiv b) \wedge (c \equiv d)) \vee ((a \equiv c) \wedge (b \equiv d))$
- $(a : b :: c : d) =$

  $$((a \rightarrow b) \equiv (c \rightarrow d)) \wedge ((b \rightarrow a) \equiv (d \rightarrow c))$$

Rappel la définition numérique basée sur la différence:

$$(a - b) = (c - d)$$

couvre 2 cas: $a \geq b$ et $b \geq a$

Propositions

- Si $\mathbf{a} \rightarrow \mathbf{b} = \mathbf{1}$ et $(\mathbf{a} : \mathbf{b} :: \mathbf{c} : \mathbf{d}) = \mathbf{1}$ alors $\mathbf{c} \rightarrow \mathbf{d} = \mathbf{1}$

- $(\mathbf{a} : \mathbf{b} :: \mathbf{c} : \mathbf{d}) = 1$ et $(\mathbf{c} : \mathbf{d} :: \mathbf{e} : \mathbf{f}) = 1$
  impliquent $(\mathbf{a} : \mathbf{b} :: \mathbf{e} : \mathbf{f}) = 1$

- Un triplet **(a b c)** *peut être complété* par **d**

  de façon à ce que $(a : b :: c : d) = 1$

  si et seulement si $(\mathbf{a} \equiv \mathbf{b}) \vee (\mathbf{a} \equiv \mathbf{c}) = \mathbf{1}$

$$(a : b :: c : d) = ((a \equiv b) \equiv (c \equiv d)) \wedge ((a \equiv b) \vee (a \equiv c))$$

Equation de la proportion analogique

Quand elle existe, l'*unique* solution de l'équation

$$(a : b :: c : x) = 1$$

est $\mathbf{x = (a \equiv (b \equiv c))}$

Cas graduel

« fuzzification » de

$(a : b :: c : d) = ((a \equiv b) \wedge (c \equiv d)) \vee ((a \equiv c) \wedge (b \equiv d))$ $\quad (A)$

$(a : b :: c : d) = ((a \rightarrow b) \equiv (c \rightarrow d)) \wedge ((b \rightarrow a) \equiv (d \rightarrow c))$ $\quad (B)$

En utilisant un des trois choix:

- i) $a \wedge b = \min(a, b)$; $a \vee b = \max(a, b)$; $a \rightarrow b = \max(1 - a, b)$;
  $a \equiv b = \min(a \rightarrow b, b \rightarrow a)$;
- ii) $a \wedge b = \min(a, b)$; $a \vee b = \max(a, b)$; $a \rightarrow b = \max(1, b/a)$;
  $a \equiv b = \min(b/a, a/b)$;
- iii) $a \wedge b = \min(a, b)$; $a \vee b = \max(a, b)$; $a \rightarrow b = \min(1, 1 - a + b)$;
  $a \equiv b = 1 - | a - b |$

B avec iii) et ii) réminiscent de la vue de la proportion analogique comme *égalité de différences* $(a - b = c - d)$, ou de *rapports* $(b/a = d/c)$

$(a : a :: c : c + s) = 1 - s$ $\qquad$ **$(a : a + s :: c : c - t) = 1 - \min(s, t)$**

Raisonnement à partir de cas (1)

• un répertoire R de *cas*
sous la forme de *paires* (problème, solution)

$$R = \{(P_t, S_t) \mid t = 1, m\}.$$

• un nouveau problème P_0

pour lequel la solution $X = S_0$ est *inconnue*

✓ chaque problème P_t : un k-uplet $P_t = (p_t1, \ldots, p_tk)$
où p_tj exprime à quel degré
« la propriété j est vraie pour le problème P_t »

✓ de même: $P_0 = (p_01, \ldots, p_0k)$

Raisonnement à partir de cas (2)

sur la base de descriptions multi-attributs de situations ou de problèmes, P et P_0, on calcule leur similarité

$$sim_{\boldsymbol{P}}(P, P_0) = \min_j sim_j(pj, p_0j)$$

"plus P et P0 sont similaires, plus il est *possible* que S et S_0 soient similaires" (au sens d'une relation de similarité $sim_{\boldsymbol{S}}$)

$$E(S_0) = \max_{(P, S) \in R} \min(sim_{\boldsymbol{P}}(P, P_0), sim_{\boldsymbol{S}}(S, S_0))$$

E(S0) : degré de possibilité d'avoir S_0 comme solution de P_0

une disjonction pondérée des ensembles flous
de valeurs proches de solutions S des problèmes similaires

Raisonnement à partir de cas et proportion analogique (1)

"objet properties-interpretation"

répertoire de locations (taille, distance au centre-ville; prix)

nouvelle maison: prix inconnu

$$((\mathbf{grande}(h_i), \mathbf{loin}(h_i)) : \mathbf{chère}(h_i)\ ::$$

$$((\mathbf{grande}(h), \mathbf{loin}(h)) : \mathbf{x}) = 1$$

1er terme de l'analogie: $a = \min(\text{grande}(h_i), \text{loin}(h_i))$

(3ème évalué de manière similaire)

trouver solution(s) $x = \text{chère}(h)$

Raisonnement à partir de cas et proportion analogique (2)

Exemple

comment varie le prix d'une maison selon qu'elle a ou non un équipement particulier

3 cas connus:

- 2 maisons similaires,

une ayant l'équipement et l'autre pas

(villa, equip, expen) = (1 1 0) (villa, not-equip, cheap) = (1 0 1)

- 1 maison, assez différente, ayant l'équipement equipment

(apart, equip, expen) = (0 1 0)

Extrapoler le prix d'une 4ème maison

(apart, not-equip, ?) = (0 0 x)

similaire à la 3ème sauf qu'elle possède pas équipement

$x = 1$? = cheap

Détails

(villa, equip, expensive) = (1 1 0)
(villa, not-equip, cheap) = (1 0 1)
(apart, equip, expensive) = (0 1 0)
(apart, not-equip, ?) = (0 0 x)

(1: 1 :: 0 : 0) = 1
(1: 0 :: 1 : 0) = 1
Conduit à postuler

(0: 1 :: 0 : x) = 1

S’étend à des degrés de vérité

De manière graduelle

Supposons
A = (1, 1, .2, .9),
B = (1, 0, .8, .8),
C = (1, 1, .3, .6),
où les degrés expriment respectivement à quel point le prix est *pas_cher* et la taxe *élevée*

Si on applique une approche basée sur la différence, implémentée à l'aide de l'implication de Lukasiewicz,

on obtient D = (0, 0, .9, .5),
c'est-à-dire que

le prix est plutôt pas_cher (.9)
et la taxe pas trop élevée (.5).

i) agréger *disjonctivement* les valeurs extrapolées entre plusieurs proportions analogiques

ii) *pondérer* les niveaux de possibilité de ces valeurs selon le degré auquel tient la proportion analogique

$$E(d') = \max_{(a, a')\in R, (b, b') \in R, (c, c')\in R} \min(PA(a,b,c,d), PA(a',b',c',d'))$$

où PA(a,b,c,d) indique à quel degré la proportion analogique tient entre les sous-vecteurs a, b, c, d.

On a comme pour la similarité $PA(a,b,c,d) = \min_j(PA(a_j,b_j,c_j,d_j)$

il est d'autant plus *possible* qu'une valeur soit en proportion analogique avec trois autres,
que la proportion analogique est satisfaite à un haut degré pour chacune des composantes caractérisant le problème

Conclusion

- Première formalisation logique de la proportion analogique

- Extension de l'approche à des degrés de vérité

Quelles extensions avec des connecteurs flous font-elles sens?

- Comparison avec les approches numériques

- Applications: raisonnement à partir de cas
  raisonnement approché
  apprentissage