Le raisonnement incertain

dans les systèmes experts

Antoine Cornuéjols
(antoine@lri.fr)

I.I.E.
&
L.R.I., Université d'Orsay

Plan général

Raisonnement non exact

- **Incertain**
  - *Neigera-t-il à Noël ?*
  - *Le fourgon postal passera sans doute par cette route*
- **Vague**
  - *Un fourgon postal avec pas mal de lingots passera aux environs de 10h15*
- **Hypothétique**
  - *Si x est étudiant, alors, jusqu'à preuve du contraire, on peut supposer que x est jeune*

Le raisonnement incertain : sources du problème

- **Théorie du domaine**
  - **Concepts imprécis**
  - **Emploi de règles heuristiques**
- **Les données**
  - **Senseurs insuffisamment précis**
  - **Données manquantes** (par principe, ou trop difficiles à obtenir)

Axiomes du raisonnement incertain

Il existe une fonction monotone croissante Φ telle que :

(A1) $0 \leq \Phi(Q|e) \leq 1$

(A2) $\Phi(A) = \Phi(A \,\&\, A)$

(A3) $\Phi(A \,\&\, B) = \Phi(B \,\&\, A)$

(A4) $\Phi(A \,\&\, (B \,\&\, C)) = \Phi((A \,\&\, B) \,\&\, C)$

(A5) $\Phi(Q|e) + \Phi(\neg Q|e) = 1$

(A6) $\Phi(A) = \Phi(A \,\&\, B) + \Phi(A \,\&\, \neg B)$

(A7) $\Phi(\text{vrai}|e) = 1$

(A8) $\Phi(QR|e) = \Phi(Q|e).\Phi(R|Qe)$

Le raisonnement incertain : approches (1/2)

- **Le raisonnement des SE :**
  - **Guidé par la syntaxe**
  - **Propagation locales des valeurs de vérité**
  - **Prise en compte incrémentale des informations**

➥ ***On voudrait les mêmes facilités pour le raisonnement incertain***

MAIS ... ***c'est impossible*** **!**

Le raisonnement incertain : approches (2/2)

- **Approches extensionnelles**
  - Incertitude ≈ valeur de vérité généralisée
  - Calcul local et incrémental
  - ➥ ***à la MYCIN / certaines logiques***
- **Approches intensionnelles**
  - Basées sur les modèles
  - L'incertitude dépend de la Base de Connaissances entière
  - ➥ ***Calcul Bayésien / Réseaux probabilistes***

R. Incertain : Approches extensionnelles

- **Généralisation des valeurs de vérité**
  - ➥ **Dépendant seulement des valeurs de vérité des sous-formules**
  - Exemple : **CF(A & B) = CF(A),CF(B))**

$\mathbf{A} \xrightarrow{m} \mathbf{B}$: Si A est observé alors on peut modifier la croyance de B d'un montant dépendant de la force *m* de la règle

- **Règle décrivant une association et sa force**

R. Incertain : Approches intensionnelles

- **Mesure de certitude assignée à des *ensembles de mondes***
  - ➥ **Connecteurs = combinaison ensembliste des ensembles de mondes**
  - ❑ Exemple : **p(A & B) = p(A ↔ B)**
  - ❑ **MAIS non déterminable à partir de p(A) et p(B) seulement**

$$\mathbf{A} \xrightarrow{m} \mathbf{B} \quad : \quad p(A|B) = m$$

Approches extensionnelles vs. approches intensionnelles

- **Approches extensionnelles**
  - \+ Modification limitée du raisonnement classique
  - \+ Modularité / explicativité / modifiabilité aisée
  - \- Non sémantiquement correctes / Inexactes
- **Approches intensionnelles**
  - \+ Bien fondées
  - \- Non modularité / explicativité difficile / modifiabilité difficile
  - \- Computationnellement coûteuses

Approche intensionnelle directe : le calcul des probabilités

- **Probabilité conditionnelle :**

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \qquad P(A|B) = \frac{P(B|A).P(A)}{P(B)}$$

$$P(H|E) = \frac{P(E|H).P(H)}{P(E|H).p(H) + P(E|\neg H).P(\neg H)}$$

$$P(H_i|E) = \frac{P(E|H_i).P(H_i)}{\sum_{k=1}^{m} P(E|H_k).p(H_k)}$$

$$\boxed{P(H_i|E_1E_2...E_n) = \frac{P(E_1E_2...E_n|H_i).P(H_i)}{\sum_{k=1}^{m} P(E_1E_2...E_n|H_k).p(H_k)}}$$

Exemple d'inférence probabiliste

	1	2	3
$p(H_i)$	0.5	0.3	0.2
$p(E_1\|H_i)$	0.4	0.8	0.3
$p(E_2\|H_i)$	0.7	0.9	0.0

- **Après la prise en compte de *E*1 :**

$$\boxed{p(H_i|E_1) = \frac{p(E_1|H_i) \times p(H_i)}{\sum_{k=1}^{3} p(E_1|H_k) \times p(H_k)} \quad , \quad i=1,2,3}$$

$$p(H_1|E_1) = \frac{0.4 \times 0.5}{0.4\times0.5 + 0.8\times0.3 + 0.3\times0.2} = 0.40$$

$$p(H_2|E_1) = \frac{0.8 \times 0.3}{0.4\times0.5 + 0.8\times0.3 + 0.3\times0.2} = 0.48 \longleftarrow$$

$$p(H_3|E_1) = \frac{0.3 \times 0.2}{0.4\times0.5 + 0.8\times0.3 + 0.3\times0.2} = 0.12$$

Problèmes de l'approche probabiliste

- **Il faut fournir beaucoup de nombres**
  - **Impossible de les demander tous à un expert**
- **Ces nombres ne sont pas indépendants (certaines sommes à 1)**
  - **Donc difficile de faire évoluer une base de connaissances**
- **Difficulté pour les experts humains de fournir des estimations fiables de probabilités conditionnelles**

Exemple d'inférence probabiliste

	1	2	3
$p(H_i)$	0.5	0.3	0.2
$p(E_1 \vert H_i)$	0.4	0.8	0.3
$p(E_2 \vert H_i)$	0.7	0.9	0.0

- **Après la prise en compte de $E1$ puis de $E2$:**

$$\boxed{p(H_i|E_1E_2) \; = \; \frac{p(E_1|H_i) \times \; p(E_2|H_i) \times \; p(H_i)}{\sum_{k=1}^{3} p(E_1|H_k) \times \; p(E_2|H_k) \times \; p(H_k)} \quad , \quad i=1,2,3}$$

$$p(H_1|E_1E_2) \; = \; \frac{0.4\times0.7\times0.5}{0.4\times0.7\times0.5 \; + \; 0.8\times0.9\times0.3 \; + \; 0.3\times0.0\times0.2} \; = \; 0.393$$

$$p(H_2|E_1E_2) \; = \; \frac{0.8\times0.9\times0.3}{0.4\times0.7\times0.5 \; + \; 0.8\times0.9\times0.3 \; + \; 0.3\times0.0\times0.2} \; = \; 0.607$$

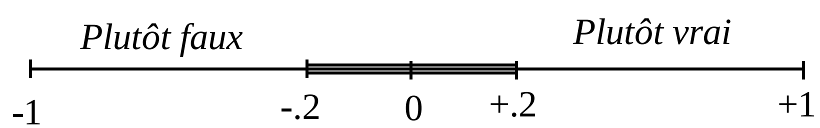
$$p(H_3|E_1E_2) \; = \; \frac{0.3\times0.0\times0.2}{0.4\times0.7\times0.5 \; + \; 0.8\times0.9\times0.3 \; + \; 0.3\times0.0\times0.2} \; = \; \cancel{0.0}$$

Approches extensionnelles : MYCIN *(1/5)*

- **Motivation historique**
  - Pas de base de données suffisante pour fournir des valeurs statistiques
  - Expertise et stratégies de contrôle à représenter explicitement
  - Capacité d'expliquer la ligne de raisonnement
  - Exigence de modifiabilité

Approches extensionnelles : MYCIN *(2/5)*

- **Les coefficients de certitude** (*certainty factor* : CF)
  - **Des faits ou hypothèses**



  - **Des règles**

    Exprimant la certitude en la conclusion, les prémisses étant supposées vraies

    Mais quelle vraie signification ?

$$CF(h,e) \; = \; \begin{cases} \dfrac{p(h|e) - p(h)}{1 - p(h)} & si \;\; p(h|e) \; \geq \; p(h) \\ \dfrac{p(h|e) - p(h)}{p(h)} & si \;\; p(h|e) \; \leq \; p(h) \end{cases}$$

Approches extensionnelles : MYCIN *(3/5)*

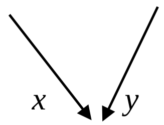
- **Propagation des coefficients de certitude**

  $R_i :\ a_1 \ \&\ a_2 \ \&\ ...\&\ a_n \quad \varnothing \quad conséquents\ \ (\mathrm{CF(R)})$

  ❶ **CF(a_1 & a_2 & ...& a_n) = min (CF(a_1), CF(a_2), ..., CF(a_n))**

  ❷ **CF(*conséquents*) = CF(a_1 & a_2 & ...& a_n) . CF($\mathrm{R_i}$)**

$$CF(H) = \begin{cases} x+y-x.y & si\, x \,\&\, y \ > \ 0 \\ x+y+x.y & si\, x \,\&\, y \ < \ 0 \\ \dfrac{x+y}{1-\min(|x|,|y|)} & \sin on \end{cases}$$



CF(H) = ?

Approches extensionnelles : MYCIN *(4/5)*

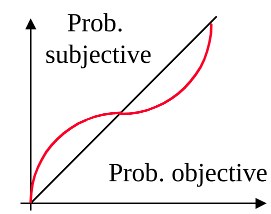
- **Le contrôle**
  - ❑ **Ne pas appliquer une règle plus d'une fois**
  - ❑ **Recherche exhaustive,**
  - ❑ **... sauf ...**
    - – Si fait certainement vrai ou certainement faux
    - – Si CF(fait) < .2

Approches extensionnelles : MYCIN *(5/5)*

- **Les défauts sémantiques**
  - ❑ **Traitement des inférences bidirectionnelles**
    - – (**Inférences déductives et abductives** : "il n'y a pas de fumée sans feu")
  - ❑ **"Explaining away"**
    - – "Si A alors C" & "Si B alors C"
  - ❑ **Limites de la modularité**
    - – "Si le sol est humide alors (.9) il a plu"
    - – *MT :* "Le système d'arrosage a fonctionné toute la nuit" (exception ou suppresseur)
    - – De même : "Si le sol est humide alors il a plu"
      "Si le système d'arrosage a fonctionné alors le sol est humide"
  - ❑ **Problème des règles différentes utilisant des sources identiques**
  - ❑ Les experts confondent CF et probabilités conditionnelles
  - ✚ Défauts d'utilisation ...

Biais dans le jugement humain

- **Précision limitée (échelle 7 ± 2 échelons)** (jugement absolu sur un stimulus unidimensionnel)
- **Calibration**



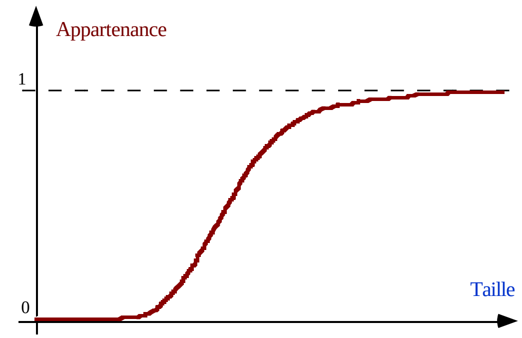
- **Trop conservatif**
  - ❑ Moins si petit échantillon, intéressement, fortes prob. a priori
- **Mais ne tiennent pas compte des probabilités a priori**
- **Corrélations positives OK, sauf si contre préjugés**

La logique floue *(fuzzy logic)*

- **Concerne l'imprécis, le vague**
  - *"Un jour assez pluvieux"*
  - **≠ Incertitude**
- **Historique :**
  - Travaux précurseurs de **Lofti Zadeh** (1965, ...) (Théorie des ensembles flous)
  - **Gros développements**
    - **Théoriques : (En France : Dubois & Prade)**
    - **Pratiques : contrôleurs flous (Japon), ...**
- **Références :**
  - B. Bouchon-Meunier : "**La logique floue**". PUF "Que sais-je?".
  - L. Gacogne : "**Elements de logique floue**". Hermès 1997.
  - M. Stefik : "**Introduction to Knowledge Systems**". Morgan Kaufmann, 1995.

Logique floue : concepts de base

- **Ensembles flous**
  - **Degré d'appartenance υ_A**
  - **Fonction caractéristique**
  - **Variables linguistiques**



Logique floue : *Modus ponens généralisé*

- **Modus ponens généralisé**

***Règle floue* :** **si X est A alors Y est B**

Fonctions d'appartenance **:** υ_A υ_B

***Fait observé* :** **X est A'**

Fonction d'appartenance : $\upsilon_{A'}$

***Conclusion* :** **Y est B'**

Fonction d'appartenance : $\upsilon_{B'}$

- **La fonction d'appartenance de B' est calculée comme une combinaison de υ_R et de $\upsilon_{A'}$:**

$$\upsilon_{B'}(y) = \sup_{x \text{ P } supp(A')} (\hat{}\ (\upsilon_{A'}(x), \upsilon_A(x) \varnothing \upsilon_B(y))$$

Raisonnement hypothétique : *approches*

- **Raisonnement par défaut : TMS**
  - **On veut faire des déductions en l'absence de certaines informations, sachant qu'elles peuvent arriver ultérieurement en cours de déduction.**

    Si x est étudiant, alors, jusqu'à preuve du contraire, on peut supposer que x est jeune.

    Albert est étudiant, il est donc supposé jeune ... puis on apprend qu'il a 70 ans
    - **Remise en cause des faits déduits : Albert est jeune (et de ceux qui en découlent)**
    - **Algorithme pour gérer ce type de remise en cause : TMS [Doyle,79]**
- **Raisonnement par hypothèse : ATMS**
  - **On établit tous les mondes possibles avec leurs justifications (hypothèses qui permettraient de les établir).**
  - **Très utile (e.g. diagnostic de pannes)**
  - **ATMS [De Kleer,86]**

Raisonnement hypothétique : TMS *(1/7)*

- **Les faits ne sont plus vrais ou faux mais ont un *statut***
  - **IN : Il existe une raison (justification) d'y croire**
  - **OUT : Rien ne le justifie (jusqu'à présent)**

- **Le *statut* peut évoluer de IN à OUT ou de OUT à IN**

R1 : <u>si</u> oiseau(X) & OUT(autruche(X)) <u>alors</u> vole(X)

R2 : <u>si</u> autruche(X) <u>alors</u> OUT(vole(X))

IN(oiseau(titi)) ∅ IN(vole(titi))

peut être contredit ultérieurement par l'ajout de IN(autruche(titi))

Raisonnement hypothétique : TMS *(2/7)*

- **Idées fondamentales**
  - **Conserver pour chaque fait déduit ses justifications**

    vole(titi) [[IN (oiseau(titi)), OUT (autruche(titi))]]
  - **En cas de contradiction, rétablir la cohérence en changeant éventuellement les statuts de certains faits**
  - **Séparer le module déductif (moteur d'inférence) de TMS**
    - Le module déductif transmet à TMS des déductions sous la forme ***(fait, justification)*** ou ***(contradiction, justification)***
    - TMS calcule pour chaque fait un statut de façon à obtenir un état globalement cohérent

Raisonnement hypothétique : TMS *(3/7)*

- **Base de faits.** Ensemble d'éléments, appelés noeuds dans la terminologie TMS, de la forme (fait, (liste de justifications), statut)
- **Contexte**. Ensemble des noeuds IN à un moment donné.
- **Justification**. Ensemble de noeuds (faits, règles ou contraintes) IN (qui doivent avoir le statut IN) supportant la dérivation du noeud considéré, et de noeuds OUT qui doivent avoir le statut OUT pour que le noeud puisse être établi (statut IN). On parle souvent de *IN-justifieurs* et de *OUT-justifieurs*.
- **Justification valide**. Une justification est valide ssi l'un au moins de ses ensembles de IN-justifieurs ont tous un statut IN, et si tous ses OUT-justifieurs ont un statut OUT.
- **Prémisses**. Lorsqu'une justification est réduite a Ø, elle est toujours valide, et on parle de prémisse (au sens de prémisses d'un argument et non des prémisses, c'est-à-dire des antécédents d'une règle).
- **Hypothèse.** Noeud dont toutes les justifications valides contiennent des OUT-justifieurs. (Ce qui signifie que ce noeud peut être remis en cause si l'un des OUT-justifieurs prend le statut IN).

Raisonnement hypothétique : TMS *(4/7)*

- **La tâche de TMS :**
  - **La tâche de TMS est de calculer (mettre à jour) une assignation A (associant un statut à chaque fait) consistante et bien fondée.**
  - **Consistante :**
    - **Pour tout fait f,**

      A(f) = IN ssi il existe une justification valide pour f.

      A(f) = OUT sinon
    - **Pour toute contradiction, A(contradiction) = OUT**
  - **Bien-fondée : 2 faits ne se co-justifient pas.**

    Il existe une chaîne de justifications valides aboutissant à des prémisses ou à des hypothèses.

Raisonnement hypothétique : TMS *(5/7)*

- **Exemple :**

R1 : Si homme(X) alors personne(X)

R2 : Si personne(X) alors humain(X)

R3 : Si humain(X) alors personne(X)

Supposons	**f1 : (homme(fred), {Ø}, IN)**
Alors :	**f2 : (personne(fred), {[IN(f1)], [IN(f3)]}, IN)**
et :	**f3 : (humain(fred), {[IN(f2)]}, IN)**

Raisonnement hypothétique : TMS *(6/7)*

Pierre est un étudiant

Typiquement un étudiant est un adulte

Typiquement un adulte est engagé dans la vie professionnelle

Typiquement un étudiant n'est pas engagé dans la vie professionnelle

R1 : Si étudiant(X) et OUT(defaut-1(X)) alors adulte(X)

R2 : Si adulte(X) et OUT(defaut-2(X)) alors engagé(X)

R3 : Si étudiant(X) et OUT(defaut-3(X)) et engagé(X) alors F

Alors : **(adulte(pierre), {[IN (étudiant(pierre), OUT (défaut-1(pierre))]}, IN)**

d'après R2 : **(engagé(pierre), {[IN(adulte(pierre)), OUT(défaut-2(pierre))], IN)**

et R3 : **(F, {[IN(étudiant(pierre),engagé(pierre)), OUT(défaut-3(pierre))]}, IN)**

D'où : **{étudiant(pierre), adulte(pierre), défaut-2(pierre)}**

ou bien : **{étudiant(pierre), adulte(pierre), ¬engagé(pierre)}**

Raisonnement hypothétique : TMS *(7/7)*

- ***Bilan***

  - **TMS ne propose qu'une seule extension dépendant des heuristiques de choix utilisées ainsi que de l'ordre dans lequel lui sont transmises les justifications**

  - **TMS a été utilisé dans de nombreuses implémentations**

  - **TMS est utilisé pour implémenter la logique des défauts**