Université Paul Sabatier
Master 2 DC - **Modèles d'incertitude, de raisonnement et de décision** - 2018-19
Documents autorisés : notes de cours

Exercice 1 - Décision sous incertitude (7 points)

Un producteur de pétrole a prospecté un terrain avec une premiere technique qui a conclu que le terrain contient du pétrole brut ou du gaz avec une probabilité $p = \frac{3}{4}$
(et ne contient rien avec une probabilité $1 - p = \frac{1}{4}$)

1. Modéliser cette information par une fonction de masse m sur l'ensemble d'états $S = \{Brut, Gaz, Rien\}$ (on ne peut pas avoir, dans ce gisement, à la fois du gaz et du brut)

   Réponse :

   $m(\{Gaz, Brut\}) = \frac{3}{4}; m(\{Rien\}) = \frac{1}{4}$

2. Calculer, pour chacun des trois évènements suivants, sa plausibilité (Pl) et son degré de certitude (Bel) : $\{Gaz\}$, $\{Gaz, Brut\}$, $\{Rien\}$

   Réponse :
   $Bel(\{Gaz\}) = 0$, $Bel(\{Gaz, Brut\}) = \frac{3}{4}$, $Bel(\{Rien\}) = \frac{1}{4}$
   $Pl(\{Gaz\}) = \frac{3}{4}$, $Pl(\{Gaz, Brut\}) = \frac{3}{4}$, $Pl(\{Rien\}) = \frac{1}{4}$

3. Le producteur a le choix entre deux décisions :

   – `Ne Pas Forer`, décision qui ne coute rien mais ne rapporte rien
   – `Forer`, avec un résultat incertain, qui coute $f = -30$, et qui rapporte $b = 40$ si il y a du Brut, $g = 100$ si il y a du Gaz, et rien si il n'y a rien

   Pour simplifier le problème, on suppose le décideur neutre par rapport aux gains/pertes, cad que sa fonction d' utilité est de la forme $u(x) = a.x + c$

   3.a Quelle serait, pour un décideur pessimiste, la meilleure des deux décisions ? justifiez votre réponse.

   Réponse :
   On est dans une situation de décision pessimiste à utilité quantitative (argent) et connaissance modélisée par une fonction de croyance. Pour la connaissance on pourrait aussi utiliser la famille de probabilité $F = \{p, p(rien) = \frac{1}{4}$ *et* $p(brut) + p(gaz) = \frac{3}{4}\}$*, ce qui reviendrait au meme.*

Dans le cas pessimiste on utilise une intégrale de Choquet fondée la mesure Bel (ou la probabilité inférieure) ou ce qui revient au même le minimun des utilités esperées sur la famille)

$$\begin{array}{rcl} Ch(Forer) & = & u(f).Bel(\{Rien, Brut, Gaz\}) \\ & + & (u(b+f)-u(f)).Bel(\{Brut, Gaz\}) \\ & + & (u(g+f)-u(b+f)).Bel(\{Gaz\}) \\ & = & u(f).1 \ + \ (u(b+f)-u(f)).\frac{3}{4} \ + \ (u(g+f)-u(b+f)).0 \\ & = & u(f) \ + \ (u(b+f)-u(f)).\frac{3}{4} \\ & = & u(f).\frac{1}{4} \ + \ u(b+f).\frac{3}{4} \end{array}$$

$Ch(NePasForer) = u(0)$

Application numérique :
$Ch(Forer) = a.(-30.\frac{1}{4}+10.\frac{3}{4})+cte = a.0+c$ $\qquad Ch(NePasForer) = u(0) = a.0+c.$
Les deux décisions sont indifférentes

3.b Montrer que dans le contexte de cette premiere expertise la décision est indépendante du prix de vente du gaz (de g)
Réponse :
Ceci découle directement de $Ch(Forer) = u(f) + (u(b+f) - u(f)).\frac{3}{4}$, *qui ne depend pas de* g

3.c A partir de quelle valeur du rapport prix de vente du brut sur coût de forage ($\frac{b}{-f}$) a-t-on interêt à forer ?

Réponse :

On fore des que $Ch(Forer) > Ch(NePasForer)$ *i.e.* $u(f).\frac{1}{4} + u(b+f).\frac{3}{4} > u(0)$ *;*
Avec $u(x) = ax + b$ *cela donne :* $a.f.\frac{1}{4} + a.(b+f).\frac{3}{4} + cte.\frac{1}{4} + c.\frac{3}{4} > a.0 + c$

C'est à dire $\frac{(f+3b+3f)}{4} > 0$ *: on fore des que le prix de vente du brut est plus grand que* $\frac{4}{3}$ *du prix du forage*

4. Une seconde expertise montre maintenant que le terrain contient du gaz avec une probabilité au moins égale à $\frac{1}{2}$.

4.a Décrire la famille de probabilités modélisant la connaissance issue des *deux* expertises.

Réponse :
$F = \{p, p(rien) = \frac{1}{4}, p(gaz) \geq \frac{1}{2}$ *et* $p(brut) + p(gaz) = \frac{3}{4}\}$ *; ici on ne peut pas utiliser les fonctions de croyance*

4.b Quelle serait, au sens de l'intégrale de Choquet pessimiste, la meilleure des deux décisions ?

Réponse :

$$\begin{array}{lcl}
Ch(Forer) & = & u(f).P_*(\{Rien, Brut, Gaz\}) + (u(b+f) - u(f)).P_*(\{Brut, Gaz\}) \\
& & +(u(g+f) - u(b+f)).P_*(\{Gaz\}) \\
& = & u(f).\frac{3}{4} + (u(b+f) - u(f)).\frac{3}{4} + (u(g+f) - u(b+f)).\frac{1}{2} \\
& = & u(f).\frac{1}{4} + u(b+f).\frac{1}{4} + u(g+f).\frac{1}{2} \\
Ch(NePasForer) & = & u(0)
\end{array}$$

Application numérique :

$$\begin{array}{lcl}
Ch(Forer) & = & a. - 30.\frac{1}{4} + 10.\frac{1}{4} + 70.\frac{1}{2}) + cte \\
& = & a.30 + cte \\
Ch(NePasForer) & = & a.0 + cte
\end{array}$$

On fore!!!

En fait on fore des que $\frac{f}{4} + \frac{f+b}{4} + \frac{f+g}{2} > 0$ *c'est à dire des que le prix du forage* $(-f)$ *est inférieur à* $\frac{b+2g}{4}$.

Exercice 2 - Complexité (6 points)

Rappel : une instance de SAT ou de UNSAT est une formule en FNC (forme normale conjonctive) $f = C_1 \wedge \ldots \wedge C_m$ où chaque clause C_i est la conjonction de littéraux. SAT est le problème de décider si f est satisfiable, tandis que UNSAT est le problème de décider si f est insatisfiable. SAT est NP-complet et UNSAT est co-NP-complet. Un problème Π est NP-difficile si pour chaque problème $\Pi_0 \in$ NP, il existe une réduction polynomiale de Π_0 vers Π.

On se propose d'étudier la complexité des problèmes suivants.

MULTI-SAT

Instance : Une formule f en FNC et un entier k.
Question : Est-ce que f possède un nombre de solutions superieur ou égal à k ?

NOPT-SAT

Instance : Une formule $f = C_1 \wedge \ldots \wedge C_m$ en FNC et un entier $k \leq m$.
Question : Est-il vrai que pour chaque sous-ensemble $S = \{C_{i_1}, \ldots, C_{i_k}\}$ de k clauses $(1 \leq i_1 < i_2 < \ldots < i_k \leq m)$, il **n**'existe **pas** d'affectation qui satisfait la formule $C_{i_1} \wedge \ldots \wedge C_{i_k}$?

Par exemple, si $f = (x_1 \vee \overline{x_2}) \wedge (\overline{x_1} \vee x_2) \wedge (x_1 \vee x_2) \wedge (\overline{x_1})$ et $k = 3$, alors la réponse pour le problème MULTI-SAT est "non" car la formule f est insatisfiable. La réponse pour le problème NOPT-SAT est également "non" car l'affectation $x_1 = x_2 = vrai$ satisfait les trois clauses $(x_1 \vee \overline{x_2})$, $(\overline{x_1} \vee x_2)$, $(x_1 \vee x_2)$.

Pour chacune des affirmations suivantes, précisez si elle est vraie ou fausse. **Justifiez chacune de vos réponses.**

1. MULTI-SAT $\in$ PSPACE

   Réponse :

   VRAIE car un algorithme naïf de recherche exhaustive n'utilise qu'un espace mémoire *polynomial.*

2. MULTI-SAT est NP-difficile.
   Réponse :
   VRAIE car il y a une réduction polynomiale de SAT vers MULTI-SAT (plus précisement vers le cas particulier de MULTI-SAT avec $k = 1$*).*

3. NOPT-SAT $\in$ coNP
   Réponse :
   VRAIE. coNP est la classe de problèmes pour lesquels il existe un vérificateur de contre-exemples de complexité polynomiale. Un contre-exemple de NOPT-SAT correspond à une affectation qui satisfait k *clauses. On peut vérifier un tel contre-exemple en temps polynomial.*

4. NOPT-SAT est coNP-complet.
   Réponse :
   VRAIE. On sait que NOPT-SAT $\in$ *coNP. Donc, il suffit de démontrer une réduction polynomiale de UNSAT vers NOPT-SAT. Mais UNSAT est équivalent au cas particulier* $k = m$ *de NOPT-SAT.*

Exercice 3 - Langages de représentation (7 points)

Partie 1

On dispose des connaissances suivantes, plus ou moins certaines, concernant l'élection au poste de directeur du département, pour laquelle X est candidat :

A1 X est membre sortant du conseil de département.
A2 Il est certain que X ne sait pas animer une réunion.
A3 Il est presque certain (0.7) que si X est présenté par la direction sortante, il sera élu.
A4 Il est relativement certain (0.6) qu'un candidat ne sachant pas animer une réunion ne sera pas élu.
A5 Il est assez certain (0.5) qu'un membre sortant du conseil sera élu.
A6 Selon une rumeur, X serait présenté par la direction sortante (certitude $x > 0$).

On considère les symboles **propositionnels** suivants : *nul* (X ne sait pas animer une réunion), *elu* (X sera élu), *sor* (X est membre sortant du conseil de département), *dir* (X est présenté par la direction sortante).

1. Traduire les connaissances sous la forme d'une base de clauses possibiliste **BP1**
   Réponse :
   Les clauses sont $C_1 : (sor, 1)$, $C_2 : (nul, 1)$; $C_3 : (\neg dir \vee elu, 0.7)$, $C_4 : (\neg nul \vee \neg elu, 0.6)$, $C_5 : (\neg sor \vee elu, 0.5)$, $C_6 : (dir, x)$.

2. On suppose $0 < x \leq 0.5$. Calculer Incons(BP1). X sera-t-il élu et avec quelle certitude ?
   Réponse :
   Incons(BP1) = 0.5 : Il y a un modèle de l'ensemble de clauses $\{C_1, C_2, C_3, C_4\}$ *et une réfutation si on ajoute* C_5.
   Avec C_2 *et* C_4 *on produit par résolution possibiliste la clause* $(\neg elu, 0.6)$ *et* $0.6 > 0.5$ *donc* $N(\neg elu) \geq 0.6$.

3. On obtient confirmation de la rumeur avec $x \geq 0.7$. Soit **BP2** la nouvelle base possibiliste. Calculer Incons(BP2). X sera-t-il élu et avec quelle certitude ?
   Réponse :
   Incons(BP2) = 0.6 : Il y a un modèle de l'ensemble de clauses $\{C_1, C_2, C_3, C_6\}$ *et une réfutation si on ajoute* C_4.
   Avec C_3 *et* C_6 *on produit par résolution possibiliste la clause* $(elu, 0.7)$ *et* $0.7 > 0.6$ *donc* $N(elu) \geq 0.7$.

Partie 2

On considère le schéma relationnel *Logement (Ville, Type, Situation)*. Un utilisateur recherche un logement à Paris ou à Lyon, en résidence universitaire ou en studio, avec les assertions de préférence suivantes :

Je préfère habiter à Paris (s1) ; si j'habite à Paris, je préfère une chambre en résidence universitaire (s2), tandis qu'à Lyon, je préfère un studio (s3) ; le centre ville me convient mieux à Paris (s4), alors qu'à Lyon, je préfère habiter en banlieue (s5).

1. Représenter la requête à préférences de l'utilisateur par un CP-net.
   Réponse :
   3 variables V, T, S. Un arc de V vers T et un arc de V vers S.
   La table CPT(V) contient p ¿ l. La table CPT(T) contient p : u ¿ s, l : s ¿ u. La table CPT(S) contient p : c ¿ b, l : b ¿ c

2. On considère l'instance suivante I du schéma *Logement* :

Tuple	Ville	Type	Situation
t1	Paris (p)	ResUniv(u)	centre (c)
t2	Paris	Studio (s)	centre
t3	Paris	ResUniv	banlieue (b)
t4	Paris	Studio	banlieue
t5	Lyon (l)	Studio (s)	banlieue
t6	Lyon	ResUniv	banlieue
t7	Lyon	Studio	centre
t8	Lyon	ResUniv	centre

   Donner les comparaisons induites par le CP-net sur les tuples de l'instance I.
   Réponse :
   On a puc ¿ luc, psc ¿ lsc, pub ¿ lup, psb ¿ lsb par CPT(V).
   On a puc ¿ psc, pub ¿ psb, lsc ¿ luc, lsb ¿ lub par CPT(T).

On a puc ¿ pub, psc¿ psb, lub ¿ luc, lsb ¿ lsc par CPT(S).
En termes de tuples, on a : $t_1 > t_2$, $t_1 > t_3$, $t_2 > t_4$, $t_3 > t_4$, $t_4 > t_5$, $t_5 > t_6$, $t_5 > t_7$, $t_6 > t_8$, $t_7 > t_8$

3. On considère une utilité décomposable U(V, T, S) = U1(V, T) + U2(V, S) telle que :
   U1(p, u) = 0.75 U1(p, s) = 0.5 U1(l, u) = 0 U1(l, s) = 0.25
   U2(p, c) = 0.75 U2(p, b) = 0.5 U2(l, c) = 0 U2(l, b) = 0.25
   (a) Construire le GAI-net associé
   (b) Quel est l'ordre représenté sur les tuples de I ?

   Réponse :
   L'ordre sur les tuples est le suivant : $t_1 > t_2$, t_2 *et* t_3 *équivalents,* $t_3 > t_4$, $t_4 > t_5$, $t_5 > t_6$, t_6 *et* t_7 *équivalents,* $t_7 > t_8$.

4. On souhaite ajouter une représentation de l'assertion suivante : si j'habite à Paris, je préfère un studio au centre ville plutôt qu'une chambre en résidence universitaire en banlieue.
   Comment modifier les tables du GAI-net précédent ? *Réponse :*
   Il suffit de modifier U1 ou U2. Par exemple, U1(p,s) = 0.6.