Prise de décisions simples

1

Plan

- Théorie de la décision
- Théorie de l'utilité
- Réseaux de décision
- Recherche d'information

2

Théorie de la décision

- Combinaison de la théorie de l'utilité et de la théorie des probabilités pour permettre à un agent de prendre des décisions rationnelles en se basant sur ses croyances et sur ce qu'il veut faire.
- Peut prendre des décisions en présence
  - d'incertain
  - de buts conflictuels
- L'agent a une mesure continue de la qualité d'un état.

3

Combiner les désires et l’incertain

- Les préférences d’un agent sont représentées par une fonction d’utilité qui donne une valeur numérique à tous les états.
- Notation: *U*(*S*): Utilité de l’état *S*.
- Les utilités sont combinées avec les probabilités des états résultant d’une action donnant ainsi une utilité espérée pour chaque action.

4

Combiner les désires et l’incertain

- Une action non déterministe *A* aura les résultats possibles suivants: $Result_i(A)$.
- Avant l’exécution de *A*, l’agent assigne des probabilités à chaque résultat (*i* = indice sur les résultats et *E* = évidence) :

$$P(Result_i(A)|Do(A), E)$$

- À partir de cela, on peut calculer l’utilité espérée d’une action:

$$EU(A|E) = \sum_i P(Result_i(A)|Do(A), E)\, U(Result_i(A))$$

5

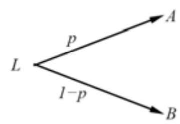
Utilité espérée maximale

- Principe de l’utilité maximale:
  - Un agent rationnel devrait choisir l’action qui maximise l’utilité espérée de l’agent.
- Permet à l’agent de choisir l’action la plus prometteuse dans un environnement incertain.

6

Préférences

- Un agent choisit parmi certains prix (*A*,*B*,etc.) et des loteries, c'est-à-dire des situations avec des prix incertains.
- Loterie *L* = [*p*, *A*; (1-*p*), *B*]



- Relations de préférences:
  - . *A* est préféré à *B*.
  - : Indifférent entre *A* et *B*.
  - : L'agent préfère *A* à *B* ou est indifférent.

7

Contraintes sur les préférences

- Ordonnable: $(A \succ B) \vee (B \succ A) \vee (A \sim B)$
- Transitivité: $(A \succ B) \wedge (B \succ C) \Rightarrow (A \succ C)$
- Continuité: $A \succ B \succ C \Rightarrow \exists p \ [p, A;\ 1-p, C] \sim B$
- Substituable: $A \sim B \Rightarrow [p, A;\ 1-p, C] \sim [p, B; 1-p, C]$
- Monotone: $A \succ B \Rightarrow (p \geq q \Leftrightarrow [p, A;\ 1-p, B] \succsim [q, A;\ 1-q, B])$
- Décomposable:
  $[p, A;\ 1-p, [q, B;\ 1-q, C]] \sim [p, A;\ (1-p)q, B;\ (1-p)(1-q), C]$

8

Théorie de l'utilité

- Principe d'utilité: si les préférences de l'agent obéissent aux contraintes précédentes, alors il existe une fonction à valeur réelle *U*:

$$U(A) > U(B) \Leftrightarrow A \succ B$$
$$U(A) = U(B) \Leftrightarrow A \sim B$$

- Principe de l'utilité espérée maximale: L'utilité d'une loterie est la somme des probabilités de chacun des résultats multipliée par l'utilité de ce résultat:

$$U([p_1, S_1; \ldots ; p_n, S_n]) = \sum_i p_i U(S_i)$$

9

Fonction d'utilité

- Dans un environnement déterministe
  - La fonction d'utilité ne définie qu'un ordre entre les états.
  - Donc, une transformation monotone de la fonction d'utilité ne change pas le comportement de l'agent.

10

Utilité de l'argent

- L'argent ne fonctionne pas comme une fonction d'utilité.
  - Par exemple, si on donne le choix entre un millions de dollars et une chance sur deux de gagner 3 millions, la majorité du monde vont choisir le million.
  - Pourtant la valeur monétaire espérée est de:
    - ½ * 0$ + ½ * 3 000 000$ = 1 500 000$

11

Utilité de l'argent

- Pour prendre une décision rationnelle, il faut regarder l'amélioration du niveau de vie.
  - Supposons k0 vos possessions actuelles, k1 vos possessions plus un million et k2 vos possessions plus trois millions.
  - $EU(Accepte) = ½ * U(S_{k0}) + ½ * U(S_{k2})$
  - $EU(Refuse) = U(S_{k1})$
- Si $U(S_{k0}) = 5$, $U(S_{k2}) = 10$ et $U(S_{k1}) = 8$
- Alors la décision rationnelle est de garder le million.

12

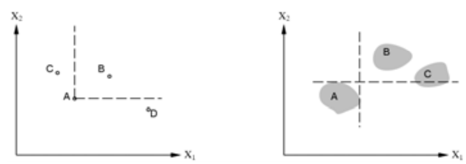
Fonction d'utilité à plusieurs attributs

- Pour certains problèmes, on a plusieurs attributs pour caractériser une situation.
  - Ex: Pour la construction d'un aéroport, il faut tenir compte du coût du terrain, de la distance par rapport à la ville, du bruit des avions, etc.
- Les attributs sont représentés par:
  - $\mathbf{X} = X_1, \ldots, X_n$
- On suppose que plus la valeur d'un attribut augmente, plus l'utilité est élevée.

13

Dominance stricte

- Une option *A* domine une autre option *B* si toutes les valeurs des attributs de *A* sont plus grandes que celles pour *B*.



À gauche, l'option A est strictement dominée par B, mais pas par C et D. À droite (incertain), A est strictement dominée par B mais pas par C.

14

Dominance stochastique

- Survient plus souvent que la dominance stricte.
- L'action A_1, domine de manière stochastique A_2, avec des distributions de probabilité de $p_1(x)$ et $p_2(x)$ sur l'attribut X si:

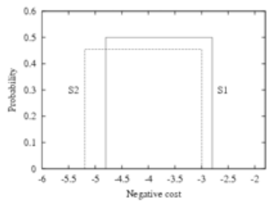
$$\forall x \quad \int_{-\infty}^{x} p_1(x')\,dx' \leq \int_{-\infty}^{x} p_2(x')\,dx'$$

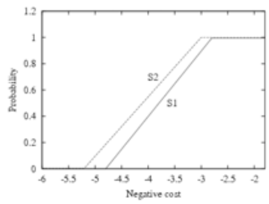
Distribution cumulative

15

Exemple

- La distribution cumulative de S_1 est toujours plus petite que celle de S_2, alors S_1 domine de manière stochastique S_2.





Dominance

- Si une action est stochastiquement dominée par une autre action sur tous les attributs, elle peut être rejetée.

Structure des préférences

- Si l'on a *n* attributs avec *d* valeurs distinctes, alors pour définir l'utilité $U(x_1,\ldots,x_n)$ on aura besoin de d^n valeurs en pire cas.
- Le pire cas c'est lorsqu'il n'y a aucune structure dans les préférences.
- Normalement, les préférences d'un agent sont structurées, donc on peut avoir:

$$U(x_1, \ldots, x_n) = f[f_1(x_1), \ldots, f_n(x_n)]$$

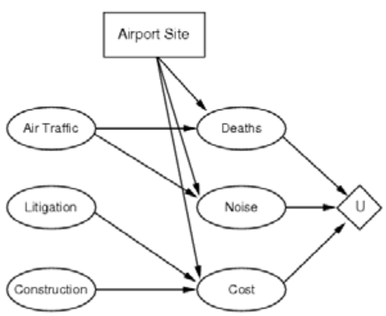
- Où *f* est une fonction simple, ex. l'addition.

Réseaux de décisions

- Ajoute des nœuds pour les actions et les utilités au réseaux bayésiens.
- Un réseau de décisions représente:
  - L'état courant de l'agent;
  - Les actions possibles;
  - L'état résultant de son action;
  - L'utilité de l'état résultant.
- Permet donc d'implémenter un agent basé sur l'utilité comme vu au chapitre 2.

19

Exemple



20

Type de noeuds

- Nœuds de chance (Ovales):
  - représente des variables aléatoires, comme dans les réseaux bayésiens.
  - Chaque nœud a une distribution conditionnelle à ses parents (nœuds de chance ou décision).
- Nœuds de décision (rectangles):
  - Représente les points où il y a un choix d'actions à faire.
- Nœuds d'utilité (losanges):
  - Représente la fonction d'utilité de l'agent.

21

Utiliser les réseaux de décisions

1. Poser les variables d'évidence (ou d'observation) pour l'état courant.
2. Pour chaque valeur possible du nœud de décision:
   a) Donner au nœud de décision cette valeur
   b) Calculer les probabilité a posteriori des nœuds parents du nœud d'utilité en utilisant un algorithme d'inférence probabiliste standard.
   c) Calculer l'utilité de l'action.
3. Retourner l'action avec la plus haute utilité.

22

Recherche d'information

- En général, l'agent n'a pas toutes les informations ou évidences dont il a besoin pour prendre une décision.
- Une partie importante de la prise de décision est de savoir quelles questions poser.
  – Par exemple, le médecin ne connaît pas tous les résultats de tests possibles et toutes les réponses aux questions quand le patient arrive.
- La théorie de la valeur de l'information permet à l'agent de choisir quelle information acquérir.

23

Exemple

- Une compagnie pétrolière veut acheter une des n zones de droit de forage.
- Une seule zone contient du pétrole d'une valeur de C dollars et chaque zone coûte C/n dollars.
  – Donc, le profit espéré est de: $C/n - C/n = 0$
- Supposons que la compagnie a la possibilité d'acheter une étude sur la zone 3, combien devrait-elle être prête à payer ?

24

Exemple (suite)

- Avec une probabilité de 1/*n*, l'étude va indiquer qu'il y a du pétrole dans la zone 3.
  - Profit de: $C - C/n = ((n - 1) * C) / n$ dollars.
- Avec une probabilité de (n-1)/n, l'étude va indiquer qu'il n'y a pas de pétrole. Le prochain bloc fera passer la pb de 1/n à 1((n-1).
  - Profit de: $C/(n - 1) - C/n = C / ((n - 1)n)$ dollars.
- Le profit espéré sachant l'information est:

$$\frac{1}{n} \times \frac{(n-1)C}{n} + \frac{n-1}{n} \times \frac{C}{n(n-1)} = C/n$$

25

Valeur de l'information

- La valeur d'une information est la différence entre la valeur espérée après l'information, moins la valeur espérée avant l'information.
- Dans notre exemple, l'information vaut:
  - $C/n - 0 = C/n$
  - La compagnie devrait donc être prête à payer jusqu'à *C/n* dollars pour l'étude.

26

Valeur d'une information parfaite

- La valeur de la meilleure action α est:

$$EU(\alpha|E) = \max_A \sum_i U(Result_i(A))\ P(Result_i(A)|Do(A), E)$$

- Après la nouvelle évidence E_j:

$$EU(\alpha_{E_j}|E, E_j) = \max_A \sum_i U(Result_i(A))\ P(Result_i(A)|Do(A), E, E_j)$$

- On ne connaît pas la valeur de E_j, donc on doit faire la moyenne sur toutes les valeurs possibles, ce qui donne finalement (avec VPI = valeur de l'information parfaite) :

$$VPI_E(E_j) = \left(\sum_k P(E_j = e_{jk}|E) EU(\alpha_{e_{jk}}|E, E_j = e_{jk})\right) - EU(\alpha|E)$$

27

Agent de recherche d'information

function INFORMATION-GATHERING-AGENT(*percept*) **returns** an *action*
static: D, a decision network

integrate *percept* into D
$j \leftarrow$ the value that maximizes $VPI(E_j) - Cost(E_j)$
if $VPI(E_j) > Cost(E_j)$
then return REQUEST(E_j)
else return the best action from D

28

Analyse de décision

- L'analyse de décision est utilisée dans plusieurs domaines où les décisions ont une grande importance:
  - Monde des affaires
  - Gouvernement
  - Loi
  - Stratégie militaire
  - Diagnostique médical
  - Santé publique
  - Conception en ingénierie
  - Gestion des ressources

29