Théories formelles de la rationalité
Logique, probabilités & théorie de la décision

D. Bonnay / M. Cozic / H. Galinon

Cogmaster rentrée 2009

Le triptyque du jour

- *Logique*
  théorie des inférences déductives

Si j'accepte un certain nombre de propositions,
quelles autres propositions suis-je *par là même* commis à accepter ?

- *Théorie subjective des probabilités*
  théorie du raisonnement dans l'incertain

Si je crois certaines propositions à certains degrés,
quelles contraintes de cohérences pèsent sur ces croyances,
comment doivent-elles évoluer ?

- *Théorie de la décision*
  théorie du choix

Comment dois-je agir
étant donné ce que je crois et ce que je désire ?

Un tryptique cohérent

Ces trois théories ont en commun

- leur objet
  la rationalité
- leur statut
  elles sont (d'abord) *normatives*
- leur formulation
  elles sont *mathématisées* et *intégrées* entre elles

The penguin curse

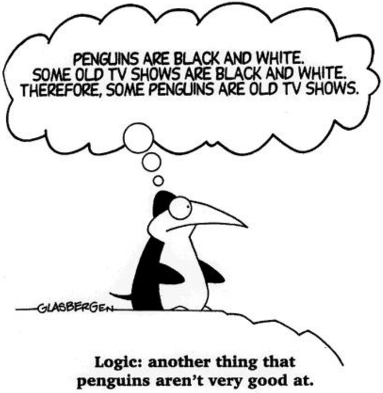
La logique a d'abord affaire à un certain type d'activités mentales : *les inférences déductives*

The penguin curse

La logique a d'abord affaire à un certain type d'activités mentales : *les inférences déductives*

The penguin curse

La logique a d'abord affaire à un certain type d'activités mentales : *les inférences déductives*



Intuitivement, le raisonnement du pingouin n'est pas correct.

Intuitivement, le raisonnement du pingouin n'est pas correct.

... Pourquoi ?

Intuitivement, le raisonnement du pingouin n'est pas correct.

... Pourquoi ?

La logique développe des *langages formels*
dans lesquels peuvent être définis :

- la notion de raisonnement valide
(comme une certaine relation entre prémisses et conclusion)
- des systèmes de preuve
(qu'on peut utiliser pour montrer qu'une conclusion particulière suit bien de certaines prémisses)

La conjonction, $\wedge$

Exemple : 'Pierre est content *et* Marie est triste.'

p : Pierre est content
q : Marie est triste

p	q	$p \wedge q$
1	1	
1	0	
0	1	
0	0	

La conjonction, $\wedge$

Exemple : 'Pierre est content *et* Marie est triste.'

p : Pierre est content
q : Marie est triste

p	q	$p \wedge q$
1	1	1
1	0	
0	1	
0	0	

La conjonction, $\wedge$

Exemple : 'Pierre est content *et* Marie est triste.'

p : Pierre est content
q : Marie est triste

p	q	$p \wedge q$
1	1	1
1	0	0
0	1	
0	0	

La conjonction, $\wedge$

Exemple : 'Pierre est content *et* Marie est triste.'

p : Pierre est content
q : Marie est triste

p	q	$p \wedge q$
1	1	1
1	0	0
0	1	0
0	0	

La conjonction, $\wedge$

Exemple : ‘Pierre est content *et* Marie est triste.’

p : Pierre est content
q : Marie est triste

p	q	$p \wedge q$
1	1	1
1	0	0
0	1	0
0	0	0

L'implication, $\rightarrow$

Exemple : '*Si* Pierre est content *alors* Marie est triste.'

p : Pierre est content
q : Marie est triste

p	q	$p \rightarrow q$
1	1	
1	0	
0	1	
0	0	

L'implication, $\rightarrow$

Exemple : '*Si* Pierre est content *alors* Marie est triste.'

p : Pierre est content
q : Marie est triste

p	q	$p \rightarrow q$
1	1	
1	0	0
0	1	
0	0	

L'implication, $\rightarrow$

Exemple : '*Si* Pierre est content *alors* Marie est triste.'

p : Pierre est content
q : Marie est triste

p	q	$p \rightarrow q$
1	1	1
1	0	0
0	1	
0	0	

L'implication, $\rightarrow$

Exemple : '*Si* Pierre est content *alors* Marie est triste.'

p : Pierre est content
q : Marie est triste

p	q	$p \rightarrow q$
1	1	1
1	0	0
0	1	1
0	0	1

Exercices : find out the truth

En logique, il n'y a pas que $\wedge$ et $\rightarrow$

une petite enquête sur le concept de *vérité*, les paradoxes qu'il suscite et les solutions qu'on peut leur proposer.

Raisonner dans l'incertain

- la plupart de nos raisonnements,
  dans la vie ordinaire et dans l'activité scientifique,
  comportent une part d'incertitude

Exemple : qui va terminer en tête de la Ligue 1 cette année ?

Raisonner dans l'incertain

- la plupart de nos raisonnements,
  dans la vie ordinaire et dans l'activité scientifique,
  comportent une part d'incertitude

Exemple : qui va terminer en tête de la Ligue 1 cette année ?

- la logique (déductive) donne peu d'outils pour raisonner dans l'incertitude

Raisonner dans l'incertain

- la plupart de nos raisonnements, dans la vie ordinaire et dans l'activité scientifique, comportent une part d'incertitude

Exemple : qui va terminer en tête de la Ligue 1 cette année ?

- la logique (déductive) donne peu d'outils pour raisonner dans l'incertitude
- la théorie qui prend en charge le raisonnement dans l'incertitude est la théorie des probabilités

Ligue 1 et probabilités (I)

On se donne un espace d'états, correspondant à un ensemble de mondes possibles :
$\{w_{Lyon}, w_{Marseille}, w_{Bordeaux}, w_{PSG}, w_{Lille}, w_{ASSE} \cdots\}$

Ligue 1 et probabilités (I)

On se donne un espace d'états, correspondant à un ensemble de mondes possibles :
$\{w_{Lyon}, w_{Marseille}, w_{Bordeaux}, w_{PSG}, w_{Lille}, w_{ASSE} \cdots\}$

A chaque état w_i est assigné un poids $p(w_i)$:
$p(w_{Lyon}) = 0.2$, $p(w_{Marseille}) = 0.2$,
$p(w_{Bordeaux}) = 0.1$, $p(w_{PSG}) = 0.1$,
$p(w_{Lille}) = 0.05$, $p(w_{ASSE}) = 0.05$, etc.

Ligue 1 et probabilités (I)

On se donne un espace d'états, correspondant à un ensemble de mondes possibles :
$\{w_{Lyon}, w_{Marseille}, w_{Bordeaux}, w_{PSG}, w_{Lille}, w_{ASSE} \dots\}$

A chaque état w_i est assigné un poids $p(w_i)$:
$p(w_{Lyon}) = 0.2$, $p(w_{Marseille}) = 0.2$,
$p(w_{Bordeaux}) = 0.1$, $p(w_{PSG}) = 0.1$,
$p(w_{Lille}) = 0.05$, $p(w_{ASSE}) = 0.05$, etc.

de telle sorte que $\sum_i p(w_i) = 1$

Ligue 1 et probabilités (II)

Un événement correspond à un ensemble d'états :

- Lyon gagne $= \{w_{Lyon}\}$
- Lyon ou Marseille gagnent $= \{w_{Lyon}, w_{Marseille}\}$
- Un club du big Four gagne
  $= \{w_{Lyon}, w_{Marseille}, w_{Bordeaux}, w_{PSG}\}$

Ligue 1 et probabilités (III)

$p(w_{Lyon}) = 0.2$, $p(w_{Marseille}) = 0.2$,
$p(w_{Bordeaux}) = 0.1$, $p(w_{PSG}) = 0.1$,
$p(w_{Lille}) = 0.05$, $p(w_{ASSE}) = 0.05$, etc.

On peut calculer des probabilités complexes à partir des informations données par la distribution de probabilités sur l'espace d'état :

- $p(\text{Lyon gagne}) =$

Ligue 1 et probabilités (III)

$p(w_{Lyon}) = 0.2$, $p(w_{Marseille}) = 0.2$,
$p(w_{Bordeaux}) = 0.1$, $p(w_{PSG}) = 0.1$,
$p(w_{Lille}) = 0.05$, $p(w_{ASSE}) = 0.05$, etc.

On peut calculer des probabilités complexes à partir des informations données par la distribution de probabilités sur l'espace d'état :

- p(Lyon gagne) $=$0.2

Ligue 1 et probabilités (III)

$p(w_{Lyon}) = 0.2$, $p(w_{Marseille}) = 0.2$,
$p(w_{Bordeaux}) = 0.1$, $p(w_{PSG}) = 0.1$,
$p(w_{Lille}) = 0.05$, $p(w_{ASSE}) = 0.05$, etc.

On peut calculer des probabilités complexes à partir des informations données par la distribution de probabilités sur l'espace d'état :

- p(Lyon gagne) $=$0.2
- p(Lyon ou Marseille gagnent) $=$

Ligue 1 et probabilités (III)

$p(w_{Lyon}) = 0.2$, $p(w_{Marseille}) = 0.2$,
$p(w_{Bordeaux}) = 0.1$, $p(w_{PSG}) = 0.1$,
$p(w_{Lille}) = 0.05$, $p(w_{ASSE}) = 0.05$, etc.

On peut calculer des probabilités complexes à partir des informations données par la distribution de probabilités sur l'espace d'état :

- p(Lyon gagne) $=$0.2
- p(Lyon ou Marseille gagnent) $=$0.2 $+$ 0.2 $=$ 0.4

Ligue 1 et probabilités (III)

$p(w_{Lyon}) = 0.2$, $p(w_{Marseille}) = 0.2$,
$p(w_{Bordeaux}) = 0.1$, $p(w_{PSG}) = 0.1$,
$p(w_{Lille}) = 0.05$, $p(w_{ASSE}) = 0.05$, etc.

On peut calculer des probabilités complexes à partir des informations données par la distribution de probabilités sur l'espace d'état :

- p(Lyon gagne) $=$0.2
- p(Lyon ou Marseille gagnent) $=0.2 + 0.2 = 0.4$
- p(Un club du big Four gagne) $=$

Ligue 1 et probabilités (III)

$p(w_{Lyon}) = 0.2$, $p(w_{Marseille}) = 0.2$,
$p(w_{Bordeaux}) = 0.1$, $p(w_{PSG}) = 0.1$,
$p(w_{Lille}) = 0.05$, $p(w_{ASSE}) = 0.05$, etc.

On peut calculer des probabilités complexes à partir des informations données par la distribution de probabilités sur l'espace d'état :

- p(Lyon gagne) $=$0.2
- p(Lyon ou Marseille gagnent) $=0.2 + 0.2 = 0.4$
- p(Un club du big Four gagne) $=0.2 + 0.2 + 0.1 + 0.1 = 0.6$

Ligue 1 et probabilités (III)

$p(w_{Lyon}) = 0.2$, $p(w_{Marseille}) = 0.2$,
$p(w_{Bordeaux}) = 0.1$, $p(w_{PSG}) = 0.1$,
$p(w_{Lille}) = 0.05$, $p(w_{ASSE}) = 0.05$, etc.

On peut calculer des probabilités complexes à partir des informations données par la distribution de probabilités sur l'espace d'état :

- p(Lyon gagne) $=$0.2
- p(Lyon ou Marseille gagnent) $=$$0.2 + 0.2 = 0.4$
- p(Un club du big Four gagne) $=$$0.2 + 0.2 + 0.1 + 0.1 = 0.6$
- p(Lyon gagne si un club du big Four gagne) $=$

Ligue 1 et probabilités (III)

$p(w_{Lyon}) = 0.2$, $p(w_{Marseille}) = 0.2$,
$p(w_{Bordeaux}) = 0.1$, $p(w_{PSG}) = 0.1$,
$p(w_{Lille}) = 0.05$, $p(w_{ASSE}) = 0.05$, etc.

On peut calculer des probabilités complexes à partir des informations données par la distribution de probabilités sur l'espace d'état :

- p(Lyon gagne) $=0.2$
- p(Lyon ou Marseille gagnent) $=0.2 + 0.2 = 0.4$
- p(Un club du big Four gagne) $=0.2 + 0.2 + 0.1 + 0.1 = 0.6$
- p(Lyon gagne si un club du big Four gagne) $=\frac{0.2}{0.6} = \frac{1}{3}$

Exercices : be smart !

Raisonnement dans l'incertain :
comment calculer les probabilités pertinentes dans une situation donnée.

la théorie de la décision

- théorie de la décision : théorie formelle de l'action
- explication naïve de l'action :

Pierre a choisi CO8 parce que Pierre désire une excellente introduction à la théorie de la décision et croit que CO8 fournit une excellente introduction à la théorie de la décision

- explication de l'action en termes d'états mentaux : des croyances et des désirs
- la théorie de la décision représente
  (i) les actions,
  (ii) les états mentaux,
  (iii) la façon dont les états mentaux se combinent pour engendrer l'action

la théorie bayésienne

- aujourd'hui (et dans une bonne partie de C08/MS2), la théorie de référence = théorie bayésienne de la décision (TB)
- la théorie bayésienne reçoit des interprétations normatives (comment les gens doivent agir) et descriptives (comment ils agissent)

la théorie bayésienne

- aujourd'hui (et dans une bonne partie de C08/MS2), la théorie de référence = théorie bayésienne de la décision (TB)
- la théorie bayésienne reçoit des interprétations normatives (comment les gens doivent agir) et descriptives (comment ils agissent)
- 2 idées centrales dans la TB :
  1. les croyances et les désirs viennent par degrés
  2. les choix résultent d'une combinaison des degrés de croyances et de désirs, exprimée par le critère d'espérance d'utilité

une situation de décision

- analyse d'une situation de décision : actions, états de la nature (ou événements) et conséquences
- vin rouge ou vin blanc ?

	poulet	viande
vin rouge	vin rouge avec poulet	vin rouge avec viande
vin blanc	vin blanc avec poulet	vin blanc avec viande

- le choix de Pierre dépend de son jugement sur la probabilité de poulet vs. viande (états de la nature) et de la manière dont il désire les 4 conséquences possibles
- les croyances portent sur les états de la nature (ou événements), les désirs sur les conséquences

l'espérance d'utilité

- les croyances sont représentées par une distribution de probabilité, les désirs par une fonction d'utilité

	poulet 1/2	viande 1/2
vin blanc	1	- 1
vin rouge	0	1

l'espérance d'utilité

- les croyances sont représentées par une distribution de probabilité, les désirs par une fonction d'utilité

	poulet 1/2	viande 1/2
vin blanc	1	- 1
vin rouge	0	1

- l'espérance d'utilité :

vin blanc	1.(1/2) + - 1.(1/2) = 0
vin rouge	0.(1/2) + 1.(1/2) = 1/2

- Pierre préfère le vin rouge au vin blanc et choisit le vin rouge plutôt que le vin blanc

Exercices : be smart again !

- Représenter une situation de décision par des matrices de probabilité et d’utilité.
- Utiliser le critère d’espérance d’utilité.

Logique et Sciences Cognitives

Hypothèse centrale des sciences cognitives

La pensée peut être comprise en termes

- de **représentations** mentales,
- et de **procédures de calcul** qui opèrent sur ces représentations.

Logique et Sciences Cognitives

Hypothèse centrale des sciences cognitives

La pensée peut être comprise en termes

- de **représentations** mentales,
- et de **procédures de calcul** qui opèrent sur ces représentations.

Des outils logiques pour les sciences cognitives

La logique fournit :

- des **langages formels** pour modéliser les représentations mentales,
- et une **théorie du calcul** pour modéliser les procédures de calcul sur ces représentations

For the sake of it

Ces théories sont, en tant que théories normatives, toujours en cours de développement :

- *Extensions de l'analyse logique*

logique épistémique, logique dynamique, logique dynamique probabiliste

- *Extensions des probabilités à l'épistémologie*

épistémologie bayésienne

- *Généralisation de la théorie bayésienne*

modèles à probabilités multiples

- *Réflexion sur les fondements*

développement de systèmes non standard reposant sur des conceptions alternatives de la rationalité

Théories normatives et descriptives

Il y a un fossé entre les normes du raisonnement et la manière dont les agents raisonnent...

- Psychologie du raisonnement :
  - Comment les humains raisonnent-ils ?
  - Comment mesurer la difficulté d'un raisonnement ?

  La logique et la théorie des probabilités fournissent des outils descriptifs pour la psycho du raisonnement

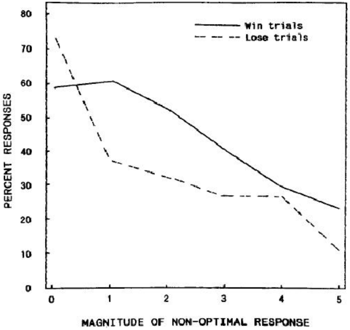
- Economie expérimentale
  - Comment les humains décident-ils ?
  - Quels sont les facteurs qui affectent la prise décision ?

  La théorie de la décision fournit des outils descriptifs / benchmarks pour l'économie comportementale

Nous sommes *vraiment* irrationnels

(Epstein & al., 1992,
Irrational reactions to negative evidence outcomes)

- Deux urnes transparentes pleines de billes
- Urne 1 : 10 billes dont 1 rouge,
- Urne 2 : 100 billes dont entre 9 et 5 rouges
- Les sujets gagnent (ou perdent selon les conditions) de l'argent si une bille rouge est tirée
- A chaque essai, l'expérimentateur donne les chances de gagner



, "I picked the ones with the more red jelly beans because it looked like there were more ways to get a winner, even though I knew there were also more whites, and that the percents were against me." Other

(Epstein & al., 1992)