Probabilité conditionnelle et probabilité du conditionnel (2)

M. Cozic (DEC, ENS) & P. Égré (CNRS, IJN)

24/04/2006

1 La notion de p-validité

Rappels :
• Langage : $\mathcal{F}(L)_{\Rightarrow} = \mathcal{F} \cup \{(\phi \Rightarrow \psi); \phi, \psi \in \mathcal{F}(L)\}$
• Probabilité et incertitude: $P()$ est une fonction de probabilité pour le langage, et $U(\phi) = 1 - P(\phi)$. L'incertitude de ϕ est la probabilité que ϕ n'ait pas lieu.
• Une inférence $\{\phi_1, ..., \phi_n\} \therefore \phi$, est p-valide ssi elle satisfait pour toute fonction d'incertitude U:

$$U(\phi) \leq U(\phi_1) + ... + U(\phi_n)$$

ie ssi l'incertitude de la conclusion est moindre que la somme de l'incertitude de chaque prémisse, ou encore si la probabilité de la conclusion est plus forte que la somme des probabilités des prémisses.

On notera: $\phi_1, ..., \phi_n \models_p \phi$ pour dire que ϕ est une conséquence probabiliste des prémisses.

• Si ϕ est une formule de $\mathcal{F}(L)_{\Rightarrow}$, alors on note: ϕ^* la formule qui résulte de ϕ lorsqu'on remplace $\Rightarrow$ par $\rightarrow$.

Théorème 1
Si $\Gamma \models_p \phi$, alors $\Gamma^* \models \phi^*$, mais la réciproque est fausse.

Donc, les validités probabilistes peuvent être vues comme un sous-ensemble propre des validités classiques.

• Lemme: toute valuation classique peut être vue comme une fonction de probabilité qui satisfait les axiomes de Kolmogoroff.
• Lemme: si $\phi \models \psi$, alors pour toute fonction d'incertitude U, $U(\psi) \leq U(\phi)$.

2 Les résultats de trivialité de Lewis

Lewis (1976), "Probabilities of Conditionals and Conditional Probabilities", repr. in *Philosophical Papers II*, Oxford UP.

Problème examiné par Lewis: est-ce que la probabilité d'un conditionnel indicatif est égale, en général, à la probabilité conditionnelle ?

"Hypothèse de Stalnaker": $P(A \Rightarrow B) = P(B|A)$
Réponse: non, sous peine de trivialité.

2.1 Premier résultat de trivalité de Lewis

- **Hypothèse de factorisation:** $P(A \Rightarrow B|C) = P(C|AB)$, si $P(AB)$ est positif.
- Soit P une fonction de probabilité, et A et C des énoncés tels que:
  $P(AC) \geq 0, P(A\overline{C}) \geq 0$
  On a :
  (1) $P(A \Rightarrow C) = P(C|A)$ (hyp. de Stalnaker)
  (2) $P(A \Rightarrow C|C) = P(C|AC) = 1$ (factorisation)
  (3) $P(A \Rightarrow C|\overline{C}) = P(C|A\overline{C}) = 0$ (factorisation)
  (4) $P(A \Rightarrow C) = P(A \Rightarrow C|C) \cdot P(C) + P(A \Rightarrow C|\overline{C}) \cdot P(\overline{C})$ (expansion par cas)
  (5) $P(C|A) = 1 \cdot P(C) + 0 \cdot P(\overline{C}) = P(C)$ (par (1), (2), (3) et (4))

Conséquence: si $P(AC)$ et $P(A\overline{C})$ sont positifs, alors A et C sont probabilistiquement indépendants.

Rappel: A et C sont probabilistiquement indépendants ssi $P(AC) = P(A) \cdot P(C)$ ssi $\frac{P(AC)}{P(A)} = P(C)$ ssi $P(C|A) = P(C)$.

Exemple: $P=$ probabilité associée au lancer d'un dé non-pipé
$A=$"obtenir un nombre pair"
$C=$"obtenir un six"
$P(AC) = \frac{1}{6}$, $P(A\overline{C}) = \frac{1}{3}$
$P(C|A) = \frac{1}{3}$, $P(C) = \frac{1}{6}$: A et C ne sont pas indépendants.

Mais, sous les hypothèses de Lewis: si on suppose que la probabilité de "si c'est un nombre pair, ce sera un six" ($P(A \Rightarrow C)$) est celle de "que ce soit un six, sachant que c'est un nombre pair" ($P(C|A)$), alors on prédit que la probabilité "que ce soit un six" ($P(C)$) est la même que celle de "que ce soit un six, sachant que ce sera un nombre pair" ($P(C|A)$).

2.2 Version "intuitive" de Adams, adaptée de Hajek

On suppose un ensemble fini de mondes possibles W, avec 10^6 mondes. On suppose que C correspond à une proposition qui représente 300000 mondes, idem pour A, et AC 100000 mondes. Donc, $P(AC), P(A\overline{C}) \geq 0$.

Par définition: $P(C|A) = P(AC)/P(A) = 100000/300000 = 1/3$.

Si $(A \Rightarrow C)$ exprime une proposition, est que sa probabilité est égale à $P(C|A)$, alors $(A \Rightarrow C)$ devrait contenir $10^6/3$ mondes (ie $P(A \Rightarrow C)$ est la probabilité d'être dans un monde dans lequel $A \Rightarrow C$ est vrai) : mais c'est un impossible, par définition, une proposition contient un nombre entier de mondes.

"Un rapport entre des aires n'est pas nécessairement une aire", de même "une probabilité conditionnelle n'est pas nécessairement la probabilité du conditionnel".

2.3 Le second résultat de trivialité

Enoncé précis du premier résultat de Lewis: l'interprétation d'un conditionnel $\Rightarrow$ ne peut être un "conditionnel de probabilité universel", ie un conditionnel qui, *pour toute fonction de probabilité* P, satisfait l'équation de Stalnaker.

Problème: ne peut-on pas maintenir l'équation, si on abandonne l'idée du conditionnel comme indépendant de la fonction de probabilité?

Réponse de Lewis: si on suppose que $\Rightarrow$ satisfait l'équation de Stalnaker pour *une classe donnée de fonctions de probabilité*, close par l'opération de conditionnalisation (ie telle que si P est dans la classe, et B est une proposition, $P(\cdot|B)$ est aussi dans la classe), alors on retrouve un résultat de trivialité.

• Soit C, D, E trois énoncés incompatibles deux à deux, de probabilités $P(C)$, $P(D)$, $P(E)$ toutes positives.

Soit $A = C \vee D$, alors $P(AC) = P(C) \geq 0$ et $P(A\overline{C}) = P(D) \geq 0$.

$P(C|A) = P(C|C \vee D) \neq P(C)$ (car $P(C \vee D) \neq 1$, sinon C et D et E ne pourraient être deux à deux incompatibles).

Donc pour tous énoncés C et A, $P(C|A) = P(C)$ seulement si le langage interdit qu'il puisse y avoir trois énoncés de ce genre.

• Si un langage est suffisamment riche pour qu'il y ait trois énoncés A, B et C de ce genre, alors l'équation ne tient pas. Une fonction de probabilité qui n'assigne jamais de probabilité positive à plus de deux propositions incompatibles est *triviale*.

3 Diagnostic du problème

• Pour Lewis, le problème vient de ce qu'on suppose que la probabilité conditionnelle est la probabilité qu'un énoncé conditionnel soit vrai.

• Pour Adams, la probabilité conditionnelle $P(C|A)$ mesure l'*assertabilité* du conditionnel, non pas la probabilité que le conditionnel $A \Rightarrow C$ soit vrai, et Lewis est d'accord avec Adams. La théorie d'Adams revient cependant à nier que les conditionnels aient des conditions de vérité, et donc à nier même qu'ils expriment des propositions (correspondant à des ensembles de mondes possibles). On peut donc comprendre le résultat de trivialité comme signifiant que les conditionnels n'expriment pas de propositions, et n'ont pas de conditions de vérité, mais seulement des conditions d'assertabilité.

• Pour Lewis, l'abandon de l'hypothèse selon laquelle les conditionnels ont des conditions de vérité est trop forte: "I have no conclusive objection to the hypothesis that indicative conditionals are non-truth-valued sentences, governed by a special rule of assertability that does not involve their nonexistent probabilities of truth. I have an inconclusive objection, however: the hypothesis requires too much of a fresh start." (Lewis 1976, p. 141).

Le problème essentiel de la thèse d'Adams porte sur l'enchâssement des conditionnels dans des énoncés pour lesquels les règles d'évaluation classiques semblent fonctionner.

La solution initiale qu'adopte Lewis est gricéenne: les conditions de vérité du conditionnel indicatif sont celles du conditionnel matériel. Lewis tente d'expliquer pourquoi l'assertabilité du conditionnel correspond à la probabilité du conditionnel, plutôt qu'à la probabilité que le conditionnel matériel soit vrai.

• Problème avec cette analyse (van Rooij 2005, Grice 1989):

Yog et Zog jouent aux échecs. Yog a les blancs 9 fois sur 10.

Sur 100 parties, Yog a gagné 80 fois quand il avait les blancs, et perdu 10 fois quand il avait les noirs.

(1) Si Yog a eu les blancs, il y a une probabilité de 8/9 qu'il ait gagné.

(2) Si Yog n'a pas gagné, il y a une probabilité de 1/2 qu'il n'ait pas eu les blancs.

Problème si on donne portée large à l'opérateur de probabilité, sur le conditionnel matériel:
$P(B \to G) = 8/9$
$P(\neg G \to \neg B) = 1/2$
Les énoncés sont matériellement équivalents, mais la probabilité n'est pas la même. Lewis doit traiter le conséquent et l'antécédent du conditionnel de façon symétrique eu égard à la vérité, mais non symétrique eu égard à l'assertabilité.

• Stalnaker lui-même reconnaît que l'équation entre probabilité du conditionnel et probabilité conditionnelle ne tient pas. En revanche, il n'en tire pas la même conclusion négative qu'Adams sur la non véri-conditionnalité des conditionnels, et il maintient, contre Lewis, une analyse unifiée des conditionnels.

4 Conditionnels indicatifs vs subjonctifs: l'analyse de Stalnaker

Remarque intéressante de Lewis 1976, note 9: "once it is recognized that the Stalnaker conditional is not a probability conditional, the coincidence of logics has a new significance. The hypothesis that assertability of indicative conditionals goes by conditional probabilities, though still sufficiently well-supported by direct evidence, is no longer unrivalled as an explanation of our judgements of validity for inferences with indicative conditional premises or conclusions. The same judgements could be explained instead by the hypothesis that the indicative conditional is the Stalnaker conditional and we judge valid those inferences that preserve truth".

Stalnaker sur le conditionnel indicatif (1975):

(3) Soit le jardinier, soit le majordome a commis le crime. Donc, si le jardinier n'a pas commis le crime, le majordome a commis le crime
$A \vee B \therefore (\neg A \Rightarrow B)$

• Cette inférence n'est pas p-valide (si on interprète "soit...soit" comme un "ou" inclusif).
$P(A) = 0.99$
$P(B) = 0.901$
$P(A \vee B) = 0.991$
$P(B \wedge \neg A) = 0.001$
$P(B|\neg A) = P(B\overline{A})/P(\overline{A}) = 0.001/0.01 = 0.1$

• Remarque sur le "ou" exclusif: si $P(AB) = 0$. $P(B\overline{A}) = P(B)$. Donc,
Exemple: $P(AB) = 0$. $P(A) = 0.5$, $P(B) = 0.1$.
$P(A \vee B) = 0.6$
$P(B|\overline{A}) = P(B)/P(\overline{A}) = 0.1/0.5 = 0.2$
Même dans ce cas, l'inférence n'est pas p-valide. Si l'on considère que l'assertabilité est garantie au-delà d'une probabilité subjective de 1/2, par exemple, alors l'assertabilité de "soit A soit B" reste garantie, mais pas nécessairement celle de "si non-A, alors B".

• Remarque de Adams sur le rejet de cette inférence: "if it were always rational to infer "if not A then B" from "either A or B", then this, coupled with the fact that the inverse inference is always rational,..., would imply that "either A or B" and "if not A then B" were logically equivalent. And by the same token this would imply that "if A then B" should be equivalent to "either not A, or B", which is itself equivalent to a material conditional" (Adams 1998, p. 124).

• Stalnaker (1975) fait remarquer que le schéma du syllogisme disjonctif (the "direct argument") est intuitivement valide. Mais comme Adams, il reconnaît qu'on ne peut accepter sa validité en toute généralité, sauf à retrouver les paradoxes du conditionnel matériel. Par exemple : $A \therefore A \vee B$, donc on aurait: $A \therefore \neg A \Rightarrow B$: "le jardinier a commis le crime ; donc s'il n'a pas commis le crime, le majordome a commis le crime" : non valide intuitivement.

• De même, on peut remarquer que l'inférence $(A \vee B) \therefore (\neg A > B)$ n'est pas valide au sens de la sémantique de Stalnaker. Soit w qui satisfait A et donc $A \vee B$. Il se peut que $f(\neg A, w)$ ne satisfasse pas B.

Thèse de Stalnaker : l'argument direct n'est pas sémantiquement valide, mais il est néanmoins *raisonnable*. Le but de Stalnaker est de rendre compte de cette notion d'inférence raisonnable.

Pour cela, Stalnaker introduit une contrainte pragmatique sur l'analyse des conditionnels. Stalnaker définit la notion de *context set* comme désignant l'ensemble des mondes compatibles avec ce que les croyances du locuteur sur ce que savent les interlocuteurs. On appelle C cet ensemble de mondes.

Contrainte de sélection: "if the conditional is being evaluated at a world in the context set, then the world selected must, if possible, be within the context set as well"
"the idea is that when a speaker says "if A", then everything he is presupposing to hold in the actual situation is presupposed to hold in hypothetical situation in which A is true"

$$\text{si } w \in C \text{ alors } f(A, w) \in C, \text{ autant que possible}$$

Subjonctif vs indicatif: l'usage du subjonctif dans un conditionnel serait une manière d'indiquer que la fonction de sélection sélectionne en dehors du context set. Par exemple, dans "si Kennedy n'avait pas été assassiné,...", on se place dans un monde contrefactuel qui n'est pas celui des connaissances partagées des interlocteurs.

Généralisation sur les disjonctions: un énoncé disjonctif n'est approprié que dans un contexte qui autorise chaque disjoint à être vrai sans l'autre (le locuteur doit envisager les deux possibilités).

Inférence raisonnable: tout contexte dans lequel la prémisse peut être assertée ou supposée de façon appropriée, est un contexte qui entraîne la proposition exprimée par la conclusion.

Soit la prémisse $(A \vee B)$: elle est appropriée si $B \wedge \neg A$ est compatible avec le context set. Dans ce cas, l'antécédent $\neg A$ de $\neg A \Rightarrow B$ est compatible aussi, ie $f(\neg A, w) \in C$. Par ailleurs, tous les $\neg A$ mondes de C sont forcément des B-mondes. Donc $f(\neg A, w) \models B$.

Conséquence: Stalnaker propose de maintenir une analyse unifiée des conditionnels indicatifs et contrefactuels. Pour Stalnaker, les conditionnels ont bien des conditions de vérité. Comme Lewis, Stalnaker rend compte des conditionnels indicatifs selon une inspiration gricéenne. En revanche, Stalnaker nie que le conditionnel indicatif ait les conditions de vérité du conditionnel matériel.