Quantification et indépendance informationnelle

Manuel Rebuschi

1. Introduction

La sémantique des jeux développée par Hintikka depuis les années 1960, enrichie par les développements de la logique IF à partir de la fin des années 1980, a fini par engendrer un paradigme véritablement nouveau. Aux questions anciennes auxquelles la nouvelle logique prétend apporter des réponses innovantes vient s'ajouter un ensemble de problématiques inédites en philosophie de la logique. Tel est le cas de l'*indépendance informationnelle* dont on essaiera de montrer qu'on ne peut en aucun cas la réduire à un simple artefact du système hintikkien.

La théorie sémantique des jeux (GTS, pour *Game-Theoretical Semantics*), dont on présentera les grandes lignes dans la première section, est fondée sur les «jeux sémantiques» qui sont proposés comme explication de la nature des relations langage-monde. Pour Hintikka, ils constituent une réalisation, en sémantique formelle, des jeux de langage de Wittgenstein. Dès lors, les jeux sémantiques sont censés combiner une approche dénotative relativement standard de la sémantique avec l'idée que les relations sémantiques peuvent être appréhendées au travers d'activités humaines gouvernées par des règles.

Mieux: Hintikka s'appuie sur la sémantique des jeux pour contester l'opposition entre sémantique vérificationniste et sémantique véri-conditionnelle. GTS est censée fusionner les deux approches: la signification d'un énoncé repose sur des activités, précisément sur des «jeux de vérification et de falsification», mais elle réside également dans ses conditions de vérité, puisque l'issue du jeu ne dépend pas des idiosyncrasies des joueurs humains impliqués.

Après une présentation de GTS et de la logique IF, on exposera quelques implications en philosophie des mathématiques, des applications à la sémantique des langues naturelles, puis en logique modale. Enfin, on conclura sur quelques difficultés apparaissant dans l'interprétation des quantificateurs dans le paradigme IF.

2. La Théorie sémantique des jeux (GTS)

L'idée du jeu sémantique est la suivante: à chaque phrase S est associé un jeu, G(S), qui met aux prises le vérificateur initial (noté **V**) et le falsificateur initial (noté **F**). Il faut concevoir **V** comme le promoteur de la phrase, et **F** comme son contradicteur. S'il s'agit d'une phrase complexe (i.e. si elle comporte des connecteurs ou des quantificateurs), elle va être décomposée au fil du jeu jusqu'à atteindre une phrase atomique; dans le cas d'une phrase atomique, **V** gagne si elle est vraie, **F** gagne si elle est fausse. La sémantique des énoncés atomiques est donc établie *avant* la partie.

Il faut distinguer deux niveaux: celui de la partie et celui des stratégies. Le fait qu'un joueur ou l'autre gagne telle ou telle partie n'est d'aucune importance. Ce qui compte véritablement en sémantique des jeux, c'est *l'existence d'une stratégie gagnante* pour le vérificateur: c'est très exactement ce qui donne les conditions de vérité de la phrase.

Considérons l'exemple d'une disjonction:

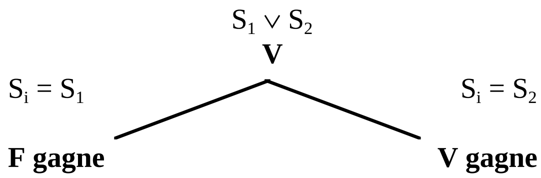
(S) Lausanne est en Angleterre (S_1) **ou** Lausanne est en Suisse (S_2).

Dans G(S), le connecteur principal étant une disjonction, le vérificateur doit choisir l'un des deux termes (S_i), et le jeu se poursuit comme G(S_i). L'idée qui sous-tend le jeu, c'est qu'en affirmant (S), le vérificateur s'engage à assumer l'un ou l'autre des deux énoncés (S_1) ou (S_2), tout en choisissant lequel des deux. On a ici deux cas:

> Cas 1: **V** choisit (S_1), le jeu se poursuit avec (S_1) qui est une phrase atomique; comme (S_1) est fausse, **V** perd (et **F** gagne).
>
> Cas 2: **V** choisit (S_2), et comme (S_2) est vraie, **V** gagne.

V peut donc perdre le jeu G(S) s'il joue mal (cas 1); mais **V** peut aussi gagner systématiquement le jeu G(S), c'est-à-dire qu'il peut gagner indépendamment des choix de **F**. Le vérificateur a donc une stratégie gagnante (choisir S_2): la phrase de départ (S) est donc vraie.

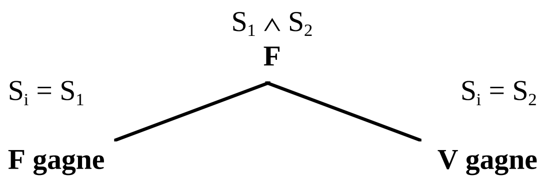
Le jeu en forme extensive permet de visualiser l'ensemble des possibilités: chaque nœud est étiqueté par le joueur dont c'est le tour (ou par l'énoncé de la victoire d'un joueur si c'est un nœud terminal), et les arêtes sont étiquetées par les actions correspondantes:



Il est utile de comparer ce cas avec celui de la conjonction:

> (S') Lausanne est en Angleterre (S_1) **et** Lausanne est en Suisse (S_2).

La règle pour la conjonction est en effet identique à celle pour la disjonction, mais c'est cette fois le falsificateur qui fait le choix. En affirmant (S'), le vérificateur s'engage maintenant à défendre l'un ou l'autre des deux énoncés (S_1) ou (S_2), même si ce n'est pas lui qui choisit lequel. Sur le jeu en forme extensive, la seule différence apparaît au niveau du nœud de départ, puisque c'est au tour du falsificateur de jouer:



Cette fois, **V** peut gagner si **F** joue mal, mais **V** n'a plus de stratégie gagnante (il ne peut plus gagner *indépendamment* des coups joués par **F**). C'est ici le falsificateur qui a une stratégie gagnante, et la phrase S' est fausse.

Les règles du jeu associés aux quantificateurs peuvent être conçus comme des généralisations des règles pour la disjonction et la conjonction.

Les énoncés existentiels sont, comme les disjonctions, joués par le vérificateur initial: dans le jeu associé à la phrase «Il y a un dragon en Suisse», le vérificateur doit pouvoir trouver un objet en Suisse tel que cet objet est un dragon; s'il ne le peut pas (c'est-à-dire s'il n'a pas de stratégie gagnante), la phrase n'est alors pas vraie. Le jeu sémantique associé à la phrase est dans ce cas, comme le dit Hintikka, un «jeu de la recherche et de la découverte».

Pour les quantificateurs universels, c'est au falsificateur de faire le choix. Une phrase universellement quantifiée comme «Tous les corbeaux sont noirs» est vraie si le vérificateur dispose d'une stratégie gagnante, c'est-à-dire si pour tout objet présenté par le falsificateur, le vérificateur peut montrer que cet objet n'est pas un corbeau ou qu'il est noir. La phrase ne sera donc pas «vérifiée» au cours d'une partie – puisque la partie est finie – mais elle sera considérée comme vraie si on dispose d'une stratégie pour déjouer toutes les falsifications (c'est le falsificateur qui a l'initiative).

Si on ajoute à cela que la négation est interprétée comme une rotation des rôles entre vérificateur et falsificateur, on peut présenter la totalité des règles pour les jeux sémantiques. Comme l'indique Hintikka: «On peut apprécier la teneur de ces règles si l'on garde présente à l'esprit l'interprétation intuitive de G(*S*) comme étant une tentative de la part de Moi-Même de vérifier *S* contre les machinations d'une Nature maligne qui cherche à falsifier *S*.»[1] Nous pouvons récapituler les règles comme suit:

[1] Hintikka (1987).

Règles du jeu GTS

Le jeu G(S) associé à une formule S d'un langage L est joué sur un modèle **M** de L, de domaine do(**M**).

(R.∧) G($S_1 \wedge S_2$) commence par le choix que fait le falsificateur de S_1 ou S_2 – disons S_i. La suite du jeu est G(S_i).

(R.∨) G($S_1 \vee S_2$) commence par le choix que fait le vérificateur de S_1 ou S_2 – disons S_i. La suite du jeu est G(S_i).

(R.∃) G($(\exists x)S(x)$) commence par un choix d'un membre du domaine do(M) effectué par le vérificateur. S'il n'y a pas de constante 'c' dans L dont la valeur soit l'individu choisi, on ajoute à L une nouvelle constante (un nouveau nom) 'c' comme nom de l'individu choisi. La suite du jeu est comme dans G(S(c)).

(R.∀) G($(\forall x)S(x)$) procède comme la précédente, si ce n'est que l'individu initial est choisi par le falsificateur.

(R.At) Si A est une formule atomique (ou une identité), alors le vérificateur gagne G(A) si A est vraie dans **M** et le falsificateur perd. Si A est fausse dans **M**, alors le falsificateur gagne et le vérificateur perd.

(R.~) Dans G($\sim S$), les deux joueurs jouent G(S) en permutant les rôles (définis par les règles-G).

Ces règles sont applicables aux énoncés d'un langage du premier ordre, après leur mise en forme normale «négative» (disjonctive ou conjonctive)[2], les énoncés atomiques étant donc préalablement interprétés à la Tarski sur une structure **M** de domaine do(**M**).

Etant données ces règles, on peut définir la vérité et la fausseté selon la sémantique des jeux:

$\mathbf{M} \vDash_{\mathrm{GTS}} S^{+}$ (S est GTS-vraie dans **M**) *ssi* le vérificateur initial a une stratégie gagnante dans G(S) joué sur **M**.

$\mathbf{M} \vDash_{\mathrm{GTS}} S^{-}$ (S est GTS-fausse dans **M**) *ssi* le falsificateur initial a une stratégie gagnante dans G(S) joué sur **M**.

[2] La forme normale négative est une forme normale prénexe (un préfixe de quantificateurs suivi d'une matrice sans quantificateurs), sans subjonctions ni bisubjonctions, les négations précédant immédiatement les atomes.

Cette définition s'avère équivalente aux définitions standard de la vérité. En supposant l'axiome du choix, on a en effet:

$\mathbf{M} \models_{\mathrm{GTS}} S^{+}$ *ssi* $\mathbf{M} \models_{\mathrm{Tarski}} S$.

L'équivalence entre les deux interprétations, GTS et Tarski, se traduit par le fait que pour une formule fermée (énoncé), il y a toujours une stratégie gagnante pour **V** ou pour **F** (i.e. à la bivalence correspond le fait que les jeux sont déterminés).

Expression des conditions de vérité

Observons l'expression des conditions de vérité GTS sur un exemple. Soit la formule:

$(\forall y)(\exists x)\ x > y$

jouée sur $\mathbb{N}$, avec l'interprétation habituelle du prédicat '>'. On va considérer une partie du jeu associé à cette formule:

- Le falsificateur choisit y:= 4; le jeu se poursuit avec: $(\exists x)\ x > 4$.
- Le vérificateur choisit x:=7; le jeu se poursuit avec: $7 > 4$.
- $7 > 4$ est une formule atomique vraie, donc **V** gagne.

Une stratégie gagnante pour **V** n'est pas de choisir x:=7, mais c'est de pouvoir choisir systématiquement, pour tous les choix de valeurs de *y* par **F**, une valeur *n* pour *x* telle que *n* soit supérieure à la valeur de *y*. Une stratégie gagnante pour le vérificateur est donc une fonction de choix *f* telle que: $(\forall y)\ f(y) > y$. Par exemple: $f(y) = y + 3$ est une bonne fonction de choix: c'est une stratégie gagnante pour **V**. Ce qui nous intéresse ce n'est cependant pas *cette* fonction de choix, mais *l'existence* d'une fonction de choix (en fait, d'un ensemble de fonctions de choix). On peut donc exprimer, dans le métalangage, les conditions de vérité de la formule d'origine de la façon suivante:

$(\exists f)(\forall y)\ f(y) > y$.

De façon générale, on peut supprimer tous les quantificateurs d'un énoncé du premier ordre en simulant les quantificateurs existentiels à l'aide de fonctions de choix (ou «fonctions de Skolem»): c'est la mise en forme normale de Skolem, pratiquée dans des contextes étrangers à GTS[3]; en restaurant les quantificateurs (sur les fonctions), on obtient une formulation des conditions de vérité GTS de la formule de départ.

La relation entre GTS et skolémisation peut alors être lue «dans les deux sens»:

1. La sémantique des jeux fournit une interprétation «naturelle» des fonctions de Skolem (ce sont des fonctions de choix, i.e. les composantes d'une stratégie pour le vérificateur); on complète la symétrie entre fonctions de Skolem et stratégies en étendant la notion de «fonction de Skolem» à la prise en charge des choix du vérificateur face à une disjonction. Ainsi, pour une formule de la forme:

$$(\forall y_1)(\forall y_2)\ldots (S_1 \vee S_2)$$

on considère la fonction «de Skolem» h telle que:

$$(S_1 \wedge h(y_1, y_2\ldots) = 0) \vee (S_2 \wedge h(y_1, y_2\ldots) \neq 0)$$

D'où la skolémisation:

$$(\exists h)(\forall y_1)(\forall y_2)\ldots ((S_1 \wedge h(y_1, y_2\ldots) = 0) \vee (S_2 \wedge h(y_1, y_2\ldots) \neq 0))$$ [4]

2. La skolémisation d'une formule fournit ses conditions de vérité GTS. C'est une formule du métalangage, qui est un fragment du second ordre[5]. Si on présuppose l'axiome du choix, on obtient l'équivalence entre une formule et sa skolémisation (i.e. la donnée de ses conditions de vérité). L'implication entre la formule et sa skolémisation est d'ailleurs une formulation possible de l'axiome du choix:

3 La forme normale de Skolem n'est pas équivalente à la formule d'origine, mais préserve la satisfaisabilité.

4 Un cas «dégénéré» est donné par la simple disjonction: $(S_1 \vee S_2)$; h est alors une fonction constante, et la skolémisation a la forme suivante: $(\exists h)\ ((S_1 \wedge h = 0) \vee (S_2 \wedge h \neq 0))$.

5 Il s'agit du fragment Σ^1_1; une formule de ce fragment est composée d'une quantification existentielle (du second ordre) sur des variables de fonctions, suivie d'une formule du premier ordre.

$(\forall x)(\exists y)\, S[x, y] \supset (\exists f)(\forall x)\, S[x, f(x)]$

(La fonction f sélectionne un élément dans chaque classe $\{y : S[x, y]\}$[6], cf. Hintikka 2001: 8).

La présentation de GTS conduit à articuler trois plans:

- Le langage ordinaire du premier ordre, dans lequel sont exprimées les formules habituelles;
- Le métalangage Σ^1_1, où s'expriment les conditions de vérité à partir des formes de Skolem;
- L'interprétation sémantique, à la Tarski ou GTS, ce qui est équivalent si l'on admet l'axiome du choix.

3. La Logique IF

La logique IF (pour *Independence Friendly*, littéralement «respectueuse de l'indépendance») est une extension de la logique standard du premier ordre. Développée par Hintikka et Sandu depuis la fin des années 1980, elle est le lieu où émergent les quantificateurs informationnellement indépendants.

L'un des angles d'attaque pour introduire la logique IF réside dans la notion de *portée* d'un quantificateur. Cette notion est en effet duale. On y trouve simultanément l'idée de portée «géographique», qui correspond à la zone où la variable est liée (i.e. la zone qui suit immédiatement le quantificateur), et celle de portée «hiérarchique», qui marque la dépendance logique entre quantificateurs. Selon Hintikka, il faut distinguer ces deux notions. En témoignent les quantificateurs ramifiés, également connus sous le nom de quantificateurs de Henkin.

Un exemple souvent débattu a la forme suivante:

(1) «Un proche (y) de chaque villageois (x) et un proche (u) de chaque citadin (z) se détestent mutuellement»

[6] On a: $(\forall x)(\exists y)\, y \in Cx \supset (\exists f)(\forall x)\, f(x) \in Cx$, avec $Cx = \{y : S[x, y]\}$.

A première vue, deux formalisations en logique du premier ordre semblent possibles:

(2) $(\forall x)(\exists y)(\forall z)(\exists u)\ S[x, y, z, u]$

(3) $(\forall z)(\exists u)(\forall x)(\exists y)\ S[x, y, z, u]$

Mais chacune de ces formalisations pose problème. En effet, dans (2), le choix du proche du citadin dépend du choix du villageois, et dans (3) le choix du proche du villageois dépend du choix du citadin. On peut souhaiter «libéraliser» l'usage des parenthèses de telle sorte que dans (2), $(\exists u)$ puisse être dans la portée «géographique» de $(\forall x)$ sans être dans sa portée hiérarchique»[7].

Considérons alors le quantificateur de Henkin:

(4) $$\left.\begin{array}{l} (\forall z)(\exists u) \\ (\forall x)(\exists y) \end{array}\right\rangle S[x, y, z, u]$$

Cette formalisation est visiblement adéquate, mais impossible à mettre sous forme linéaire: elle se situe en fait au-delà du premier ordre. Cela signifie que la logique standard du premier ordre impose une contrainte non justifiée de dépendance mutuelle systématique des quantificateurs: si un quantificateur existentiel est sous la portée géographique d'un quantificateur universel, il en dépend logiquement (ou «hiérarchiquement»).

La proposition de Hintikka et Sandu consiste simplement à ajouter un symbole, la barre oblique '/', pour marquer l'indépendance entre quantificateurs. A partir de l'exemple ci-dessus, on peut ainsi obtenir les trois formalisations suivantes:

(5) $(\forall x)(\exists y)(\forall z)(\exists u/\forall x)\ S[x, y, z, u]$

(6) $(\forall z)(\exists u)(\forall x)(\exists y/\forall z)\ S[x, y, z, u]$

(7) $(\forall x)(\forall z)(\exists y/\forall z)(\exists u/\forall x)\ S[x, y, z, u]$

[7] La pertinence de cette interprétation de la phrase de départ est évidemment discutable. Mais ce qui m'intéresse ici, c'est précisément la formalisation de *cette* interprétation (la formalisation de l'interprétation avec des quantificateurs dépendants ne posant évidemment aucun problème).

(5), (6), (7) étant des formules équivalentes à (4), mais cette fois sous une forme linéaire. En ajoutant la barre oblique, on est en fait passé de la logique ordinaire du premier ordre à la logique IF du premier ordre.

Expression des conditions de vérité

On a vu que les conditions de vérité (GTS) des formules du premier ordre ordinaire pouvaient s'exprimer à l'aide de leurs formes de Skolem. Les skolémisations présentent le même type d'intérêt pour les formules IF. On peut ainsi considérer les skolémisations respectives des formules (2), (3) et (5):

(2') $(\exists f)(\exists g)(\forall x)(\forall z)\ S[x, f(x), z, g(x, z)]$

(3') $(\exists f)(\exists g)(\forall x)(\forall z)\ S[x, f(x, z), z, g(z)]$

(5') $(\exists f)(\exists g)(\forall x)(\forall z)\ S[x, f(x), z, g(z)]$

La dépendance mutuelle entre quantificateurs (dans la logique ordinaire du premier ordre) se traduit donc par l'inclusion mutuelle des ensembles de variables arguments des fonctions de Skolem associées. Si on passe à la logique IF, les fonctions de Skolem peuvent être définies pour des ensembles de variables non inclus les uns dans les autres (et réciproquement).

La logique IF n'envisage pas l'indépendance uniquement pour les quantificateurs, mais aussi pour la disjonction. Cela se traduit semblablement au niveau des fonctions de Skolem. Tandis que la «skolémisation» d'une disjonction ordinaire après deux quantificateurs universels donne une fonction à deux variables:

$(\forall x)(\forall y)\ (S_1 \vee S_2)$

$(\exists h)(\forall x)(\forall y)\ ((S_1 \wedge h(x, y) = 0) \vee (S_2 \wedge h(x, y) \neq 0))$

celle de la même disjonction, rendue indépendante de l'un des deux quantificateurs, fournit une fonction à une seule variable:

$(\forall x)(\forall y)\ (S_1\ \mathbf{(\vee/\forall y)}\ S_2)$

$(\exists h)(\forall x)(\forall y)\ ((S_1 \wedge h(x) = 0) \vee (S_2 \wedge h(x) \neq 0))$

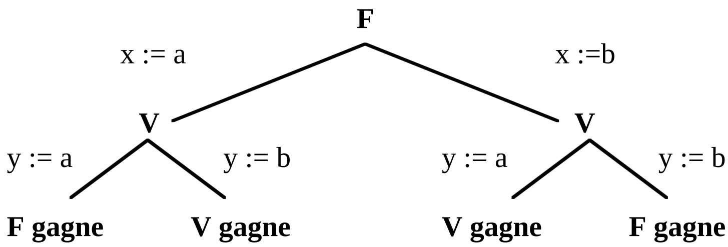
Traduction sur le plan sémantique (GTS)

La sémantique des jeux s'étend tout naturellement aux formules IF. La dépendance logique (ou «hiérarchique») des quantificateurs se traduit par une dépendance informationnelle des choix de valeurs pour les variables. Ainsi, la formule (2) sera lue: «Pour tout x on peut trouver un y, tel que pour tout z on peut trouver un u, tel que…», le choix d'une valeur pour la variable u étant implicitement considéré comme dépendant des choix de valeurs pour les variables x et z. Par contraste, on lira la formule (4) en marquant l'indépendance du choix d'une valeur pour u par rapport au choix pour x: «Pour tout x on peut trouver un y, tel que pour tout z on peut trouver un u *indépendamment de x*, tel que…».

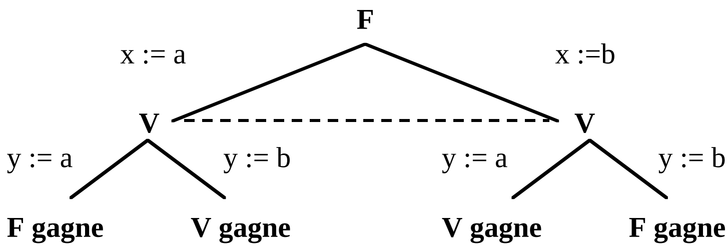
C'est donc bien au niveau de l'interprétation sémantique qu'apparaît l'indépendance informationnelle annoncée dans le titre de cette contribution. L'idée est qu'en jouant un quantificateur existentiel Q_1 indépendant d'un quantificateur universel Q_2, le vérificateur doit choisir une valeur pour la variable liée par Q_1 sans avoir d'information au sujet de la valeur antérieurement choisie par le falsificateur pour la variable liée par Q_2.

Comparons les jeux sémantiques associés à deux énoncés simples, l'un en logique ordinaire du premier ordre, l'autre en logique IF, interprétés sur un modèle **M** à deux éléments, a et b.

(G1) (∀x)(∃y) x ≠ y



(G2) (∀x)(∃y/∀x) x ≠ y



Entre G1 et G2, on passe d'un jeu à information parfaite à un jeu à information imparfaite (ou d'un jeu «transparent» à un jeu «opaque»): dans G2, le vérificateur ne peut pas discerner sa position dans le jeu entre les deux nœuds reliés en pointillé[8]. Dans le jeu à information imparfaite, le vérificateur n'a plus de stratégie gagnante. Ou plus précisément: **V** a toujours une stratégie gagnante – la même que dans G1 – mais il ne peut plus y accéder; il n'a pas de *stratégie gagnante uniforme*, ce qui est en fait exigé pour établir la vérité de la formule.

GTS fournit donc une bonne sémantique pour la logique IF – sans changement des règles associées aux constantes logiques, le seul changement intervenant au niveau des stratégies. GTS fournit même une interprétation naturelle pour la logique IF: on peut comprendre le passage de la logique ordinaire du premier ordre à la logique IF du

[8] Un interprétation «concrète» de GTS peut être que chaque joueur abstrait, **V** et **F**, est en fait une équipe de joueurs. Voir Pietarinen, Sandu (2000).

premier ordre comme un passage des jeux à information parfaite aux jeux à information imparfaite – c'est-à-dire comme le passage d'un cas particulier au cas général.

Les jeux à information imparfaite ne sont pas toujours déterminés – p.ex. G2 n'est pas déterminés: **V** n'a pas de stratégie gagnante, mais **F** non plus! Cela signifie qu'en passant du premier ordre ordinaire au premier ordre IF, on perd la bivalence: la formule IF jouée en G2 n'est ni vraie, ni fausse sur **M**.

Terminons cette rapide présentation de la logique IF par quelques arguments de ses auteurs en sa faveur, et particulièrement de Hintikka selon qui la logique IF est notre vraie logique naturelle:

- Dès qu'on comprend la notion habituelle de quantificateur, on comprend la notion de quantificateur indépendant – l'indépendance informationnelle est une notion «mafieuse» (on ne peut pas y échapper);
- En effet: le vrai sujet de la logique du premier ordre, ce ne sont pas les quantificateurs considérés isolément, mais les dépendances et indépendances mutuelles entre quantificateurs;
- Il n'y a aucune raison qui justifie la contrainte imposée par Frege aux dépendances mutuelles entre quantificateurs (ou, de manière équivalente, à l'usage des parenthèses);
- L'indépendance informationnelle est un trait des langues naturelles (voir Section 5).

4. Implications pour les fondements

Fin du *tertium non datur*

Comme on vient de le voir à propos de l'interprétation sémantique de la logique IF, une phrase qui n'est pas vraie n'est pas nécessairement fausse, car le fait qu'il n'y ait pas de stratégie gagnante pour le vérificateur n'implique pas qu'il y en ait une pour le falsificateur.

Ceci est lié au caractère particulier de la négation GTS, la négation *duale* (~), qui correspond à une inversion des rôles entre les deux joueurs[9]. On pourrait ajouter à la logique IF la négation (faible) contradictoire ($\neg$), c'est-à-dire une négation telle que les modèles où $\neg S$ est vraie seraient les complémentaires des modèles où S est vraie. Pour la négation $\neg S$ d'une formule IF, elle-même formule de IF étendue ($\mathrm{IF}^{\mathrm{ext}}$), et pour un modèle $\mathbf{M}$, on peut définir la vérité et la fausseté de la manière suivante:

$$\mathbf{M} \models_{\mathrm{GTS}} (\neg S)^{+} \text{ ssi } \mathbf{M} \not\models_{\mathrm{GTS}} S^{+}$$

$$\mathbf{M} \models_{\mathrm{GTS}} (\neg S)^{-} \text{ ssi } \mathbf{M} \not\models_{\mathrm{GTS}} S^{-}$$

La négation contradictoire affirme quelque chose à propos d'un jeu (elle ne peut pas être jouée, i.e. interprétée dans GTS). Une conséquence de son introduction est la suivante: si S est une phrase IF sans valeur de vérité, $(\neg S)$ sera à la fois vraie *et* fausse[10].

Il est utile ici de noter la version du *Théorème de séparation* valable pour la logique IF: si S est une phrase d'un langage IF du premier ordre, et si $(\neg S)$ est représentable dans ce langage, alors S est représentable par une phrase du premier ordre *ordinaire*. Cela signifie en effet que les seules phrases d'un langage IF (non étendu) dont la négation contradictoire s'exprime dans le même langage sont les phrases du premier ordre ordinaire.

Ce rejet du tiers-exclu ne fait pas de la logique IF une sœur de la logique intuitionniste: loin de restreindre le premier ordre, IF est une *extension* de la logique classique. A partir de GTS on peut cependant considérer des restrictions *constructivistes* (non intuitionnistes) de la logique ordinaire du premier ordre: par exemple la restriction aux fonctions de Skolem récursives – une restriction qui confère un minimum de plausibilité cognitive aux jeux sémantiques. On obtient alors moins de phrases vraies et moins de phrases fausses que dans les

[9] Cf. Hintikka, Sandu (1997: 373-375).

[10] Il n'y a pas de formule simultanément vraie et fausse en logique IF. C'est uniquement avec $\mathrm{IF}^{\mathrm{ext}}$ qu'on obtient une logique à quatre valeurs.

jeux sans limitations sur les stratégies, d'où un abandon du tiers-exclu[11].

Entre premier et second ordre

Comme on l'a signalé en introduisant les quantificateurs indépendants, de la même manière que pour les formules ordinaires du premier ordre on peut skolémiser les formules IF du premier ordre. On a pour la logique IF le même théorème que pour la logique ordinaire, à savoir qu'en supposant l'axiome du choix, chaque formule IF du premier ordre est équivalente à sa skolémisation, c'est-à-dire à une formule du fragment Σ_1^1 du second ordre.

Mais il y a mieux: toute formule Σ_1^1 est équivalente à une formule du premier ordre préfixée d'un quantificateur de Henkin, et les quantificateurs de Henkin sont définissables dans la logique IF du premier ordre. Cela signifie que toute formule Σ_1^1 est équivalente à une formule de la logique IF du premier ordre[12]. Donc la logique IF du premier ordre est équivalente au fragment Σ_1^1 – ce qui a fait dire à Hintikka que la logique IF était d'ordre «1 + epsilon».

Cette expressivité accrue de la logique du premier ordre permet donc d'exprimer le métalangage (les conditions de vérité des phrases IF) dans le langage (comme des phrases IF). La sémantique cesse d'être «ineffable».

Définition d'un prédicat de vérité

On peut définir un prédicat de vérité pour un langage IF dans ce même langage. On peut donc échapper à «la malédiction de Tarski». L'idée est la suivante: si on autorise l'indépendance entre quantificateurs, on pourra projeter le métalangage dans le langage *via* l'encodage de Gödel, de telle sorte que les nombres de Gödel se trouvent

[11] Cette restriction aux fonctions de Skolem récursives peut évidemment s'appliquer à la logique IF.

[12] Cf. Hintikka, Sandu (1997: 368).

uniquement sous la portée des quantificateurs qui les concernent, et jamais sous la portée des quantificateurs qui visent les autres nombres, à savoir ceux qui ne sont pas des nombres de Gödel. Autrement dit, les quantificateurs indépendants permettent de préserver la distinction entre usage et mention.

Pour un langage comportant l'arithmétique élémentaire, Hintikka (1998) montre que l'on peut construire un prédicat «x est vrai» directement en fonction du nombre de Gödel de la phrase à laquelle il s'applique – ce n'est donc pas une définition récursive de la vérité à partir de la notion de satisfaction, comme la définition de Tarski[13]. L'application du prédicat à une phrase est équivalente à une formule de Σ_1^1, non exprimable dans le premier ordre ordinaire (c'est le maintien de l'interdit de Tarski), mais bien entendu exprimable dans la logique IF du premier ordre.

On échappe aux paradoxes du menteur du fait qu'il n'y a pas de tiers exclu dans IF: les énoncés qui affirment d'eux-mêmes qu'ils sont faux n'ont simplement pas de valeur de vérité. Enfin, d'après Hintikka, contrairement à ce que pensait Tarski, ce résultat s'applique aux langues naturelles puisque l'indépendance informationnelle fait partie des universaux linguistiques.

Incomplétude de la logique IF du premier ordre

La logique IF hérite des propriétés du fragment Σ_1^1 du second ordre. En particulier, la logique IF n'est pas sémantiquement complète. Pour Hintikka, cela présente un net avantage pour la réalisation de son projet néo-logiciste.

En effet, grâce à l'incomplétude sémantique de la nouvelle logique, Hintikka pense pouvoir contourner l'incomplétude «descriptive» des axiomes propres de l'arithmétique. L'idée est la suivante: le théorème d'incomplétude de Gödel affirme que l'arithmétique élémentaire

[13] La seule condition exigée pour pouvoir définir un prédicat de vérité dans un langage est que l'on puisse représenter dans ce langage une fonction qui envoie chaque individu du domaine sur l'individu qui le représente dans la numération de Gödel. Cf. Hintikka (1998: 314)

formalisée dans le premier ordre n'est pas déductivement complète. Comme la logique ordinaire du premier ordre est sémantiquement complète (Gödel 1930), les axiomes arithmétiques sont, classiquement, considérés comme descriptivement incomplets. Mais avec IF, on a une logique sémantiquement incomplète: il devient alors possible d'envisager des axiomes descriptivement complets. Autrement dit, les modèles non-standards (au sens de Henkin) pourraient être purement et simplement éliminés.

Autre conséquence de l'incomplétude de IF: la logique cesse d'être purement computationnelle ou mécanique. A l'opposé de l'image colportée par les plus polémiques des antilogicistes au début du vingtième siècle, la créativité n'est pas réservée aux seuls mathématiciens. En effet, en logique IF du premier ordre, certaines règles d'inférences «valides» ne sont pas déductibles des axiomes; il faut recourir à des méthodes sémantiques pour les trouver.

Une illustration en est offerte par le recours systématique à l'axiome du choix par GTS. Cela indique, selon Hintikka, que cet axiome est un principe *logique*. S'il est indécidable en théorie des ensembles – formalisée dans le premier ordre ordinaire –, c'est que la théorie des ensembles est un mauvais cadre de fondement. L'équivalence entre une formule et sa skolémisation est tellement naturelle dans le contexte de la sémantique des jeux qu'il serait artificiel de s'en priver (cf. Hintikka 1998).

5. Applications et extensions

La sémantique des langues naturelles

Présentée par ses auteurs comme une possible révolution[14], l'invention de la logique IF permet d'envisager une réunification de la logique après un siècle de dispersion. En sémantique des langues naturelles, notamment dans la lignée des travaux de Montague, les formalismes

[14] Cf. Hintikka, Sandu (1996).

ont fortement évolué depuis la logique du premier ordre des origines. Le fait que le nouveau paradigme constitué par GTS et la logique IF offre un nouveau socle qui puisse être commun à la sémantique des langues naturelles *et* aux fondements témoigne à lui seul de l'importance de ces travaux.

La sémantique des jeux est une sémantique non compositionnelle: elle procède de l'extérieur vers l'intérieur. Pourtant la compositionnalité est souvent mise en avant comme une contrainte raisonnable imposée aux théories sémantiques si nous voulons qu'elles reflètent un tant soit peu l'appréhension cognitive des significations. GTS serait-elle hors-jeu?

Elle n'est en fait pas plus mal positionnée que les approches plus «standard». Car les modèles (impliqués dans les interprétations à la Tarski ou à la Montague) censés modéliser ce que l'on doit connaître pour saisir la signification d'une expression ont généralement le défaut d'être infinis. A cela, GTS oppose des jeux *finis*. En sémantique des jeux, pour comprendre une expression, il faut *maîtriser* le jeu sémantique associé. Maîtriser le jeu, ce n'est ni jouer une partie (le locuteur n'est pas le vérificateur initial), ni savoir comment le gagner (i.e. connaître la stratégie gagnante). Pour comprendre la phrase: «Tous les corbeaux sont noirs» (S), on n'a pas besoin de posséder les moyens de connaître la vérité ou la fausseté de S.

On distingue alors deux niveaux de signification: la signification abstraite et la signification stratégique. La première est définie par les règles du jeu, c'est ce que l'on connaît quand on comprend la phrase – ce qui correspond à la maîtrise du jeu: on sait que la phrase est vraie si et seulement si le vérificateur a une stratégie gagnante. La seconde correspond à l'idée que l'on se fait des stratégies gagnantes pour le vérificateur: la connaissance de la signification stratégique, c'est la connaissance des fonctions de Skolem qui constituent la stratégie gagnante de **V**.

La signification stratégique intervient dans la résolution d'anaphores[15], dont le traitement se trouve au cœur de la sémantique des langues naturelles depuis le *dynamic turn*. Ici, GTS peut tenir tête aux autres théories présentes sur le marché, notamment à la DRT (*Discourse Representation Theory*) de Kamp[*], et à la DPL (*Dynamic Predicate Logic*) de Groenendijk et Stokhof. La non-compositionnalité de GTS s'avère alors avantageuse en ce qu'elle autorise la dépendance vis-à-vis du contexte, un trait évidemment recherché en sémantique dynamique. Autre aspect naturellement dynamique de la sémantique des jeux: la possibilité de composer les jeux en super-jeux – ou de décomposer des jeux complexes en sous-jeux simples: le passage du niveau de la phrase à celui du discours est on ne peut plus naturelle. La conjonction dynamique est simplement traitée comme le déroulement successif des jeux correspondants.

Les pronoms anaphoriques sont traités par GTS comme des descriptions cachées à la Russell qui prennent leur valeur dans un ensemble de choix contextuellement défini au cours du jeu (voir Sandu 1997). Par rapport à DRT (qui manipule des «représentants» d'objets), la sémantique des jeux présentait l'inconvénient de traiter l'anaphore comme de la co-référence: l'ensemble de choix étant composé d'objets directement extraits du domaine de l'interprétation, on rencontrait un blocage pour la résolution d'anaphores sur des objets fictifs ou inexistants. Mais la situation a été récemment modifiée par un découplage complet entre traitement sémantique de l'anaphore et interprétation (référentielle) du discours (voir Janasik, Sandu, 2003). Il en ressort que ce qui prime dans le traitement GTS de l'anaphore, ce n'est pas l'interprétation modèle-théorétique ultime des termes singuliers, mais bel et bien le *flux informationnel* qui se diffuse au travers du jeu sémantique[16].

[15] Grossièrement, la résolution d'anaphore consiste à trouver le référent du pronom «il» dans un discours comme celui-ci: «Un homme marche dans le parc. *Il* siffle».

[*] Sur la DRT, cf. ici même l'article de A. Ter Meulen (nde).

[16] L'idée est que sur un exemple comme: «Un homme marche dans le parc. *Il* siffle», le fait que l'anaphore soit résolue (tandis qu'elle échouerait si l'ordre des phrases étaient inversé) est

Les logiques modales IF

Avec la sémantique des mondes possibles, on peut aisément envisager des extensions IF de la logique modale. En effet, rien n'interdit de rendre informationnellement indépendants des quantificateurs portant sur des mondes possibles. Il suffit pour cela de doter la logique modale d'une sémantique des jeux, et d'introduire l'indépendance informationnelle au sein des jeux sémantiques – exactement comme pour la logique du premier ordre.

En logique modale propositionnelle, on associera à un énoncé de la forme $\Box\varphi$, à un modèle **M** et un monde possible w, une règle du type: «le falsificateur choisit, s'il y en a un, un monde alterne w', et le jeu continue avec φ en w'; s'il n'y a pas de monde accessible, le vérificateur a gagné». On associera une règle équivalente aux énoncés de la forme $\Diamond\varphi$, en inversant les rôles.

Des opérateurs modaux informationnellement indépendants peuvent alors être proposés, comme dans la formule suivante:

$$\Box_1 \Box_2 (\Diamond/\Box_2)\, \varphi$$

qui est équivalente à la formule modale ordinaire:

$$\Box_1 \Diamond \Box_2\, \varphi$$

Mais on peut également considérer des formules qui n'ont pas d'équivalent en logique modale ordinaire:

$$\Box_1 \Box_2 (\Diamond/\Box_1)\, \varphi$$

Tulenheimo (2002) établit que la logique modale IF résultante est strictement plus expressive que la logique modale standard, au sens où deux modèles indiscernables par un énoncé modal ordinaire peuvent être distingués par des énoncés modaux IF.

---

totalement indifférent au fait qu'il y ait ou non un homme dans le parc. Seule l'information que «un homme» a été introduit au cours du jeu correspondant au traitement de la première phrase est ici pertinente, pour assurer une stratégie gagnante au vérificateur dans le jeu correspondant à la seconde phrase.

Un cas particulier intéressant de logique modale IF est fourni par la logique épistémique. La distinction traditionnelle entre connaissance *de dicto* et connaissance *de re* peut être avantageusement reformulée. En effet, si la logique modale quantifiée ordinaire suffit à rendre compte du contraste entre, par exemple, les deux phrases suivantes:

«On sait qu'il y a un espion dans le voisinage» $K (\exists x) S[x]$

«On connaît un espion dans le voisinage» $(\exists x) K S[x]$

elle ne permet pas de formaliser la connaissance *de re* d'un individu dépendant d'un quantificateur universel, comme dans: «On sait quelle personne [*de re*] chacun déteste le plus». La connaissance *de dicto* correspondante aurait la forme suivante: «On sait que pour toute personne, il y a une personne qu'elle déteste le plus», aisément formalisable:

$$K (\forall x)(\exists y) S[x, y]$$

tandis que la forme *de re* exige des quantificateurs indépendants:

$$K (\forall x)(\exists y/K) S[x, y]$$

Sur ce modèle, la logique épistémique IF propose une reclassification des énoncés: le clivage *de re* / *de dicto* doit être remplacé par celui entre *knowing-that* (savoir-que) et *knowing-wh* (*who*, *what*, *whether*…) (savoir qui, quoi, si…). Si le *knowing-that* correspond à la connaissance *de dicto*, formalisée en logique épistémique standard, le *knowing-wh* s'avère plus étendu que la connaissance *de re*. A celle-ci (formalisée typiquement en: $K (\exists x/K) S[x]$) viennent s'ajouter les phrases en *savoir-si*: *savoir si A ou B* n'est plus représenté comme en logique épistémique classique par: $KA \vee KB$, mais à l'aide d'une disjonction indépendante:

$$K (A (\vee/K) B)$$

On retrouve avec le *knowing-wh* également des exemples qui ne peuvent pas être formalisés en logique épistémique ordinaire. Ainsi le *savoir-qui* est-il représenté à l'aide d'un quantificateur indépendant: $K (\exists x/K) [x = h]$, formule qui peut être considérée comme équivalente à l'énoncé sans barre oblique: $\exists x\ K\ [x = h]$. Mais si je veux affirmer que ma grand-mère sait (*de dicto*) que je sais qui est Hintikka, sans qu'elle-même sache qui est Hintikka, on obtient une formule de logique épistémique IF sans équivalent en logique épistémique standard:

$$K_{GM}\ K_{MR}\ (\exists x\ /\ K_{GM})\ [x = h]$$

6. Conclusion

Selon Hintikka, la signification véritable des quantificateurs – telle que nous la donne GTS – ne fait que recouper l'usage ordinaire qu'en font les mathématiciens. Une formule telle que: $(\forall m)(\exists n)\ n > m$, sera ordinairement lue de la manière suivante: «Pour tout m, *on peut trouver* un nombre n qui lui est strictement supérieur». Doit-on y voir la marque du caractère naturel des «jeux de la recherche et de la découverte»? Les quantificateurs indépendants seraient-ils du coup des quantificateurs comme les autres? Rien n'est moins sûr. Certains aspects de la logique IF du premier ordre sont problématiques et incitent à la prudence.

Première bizarrerie: en logique IF, objectualité et compositionnalité sont incompatibles. On a plus haut mentionné la non-compositionnalité de GTS. Cette non-compositionnalité n'est pas particulièrement problématique pour les langages du premier ordre ordinaire, sachant qu'une sémantique compositionnelle (par exemple à la Tarski) demeure toujours disponible. Les choses se compliquent pour les langages IF. On peut encore construire une sémantique compositionnelle. C'est le cas de la *trump semantics* de Hodges, qui est une extension compositionnelle de GTS. Mais cette sémantique présente un aspect contre-intuitif: les phrases ne sont plus satisfaites par des ensembles de suites d'objets, mais par des ensembles d'ensembles de suites d'objets.

Seconde bizarrerie: selon que l'on considère, ou non, que les choix du vérificateur sont tacitement indépendants de ses propres choix antérieurs, c'est-à-dire en fonction des pouvoirs épistémiques attribués aux joueurs, la formule suivante (due à Hodges):

$$(\forall x)(\exists z)(\exists y/\forall x)\ x \neq y$$

sera vraie ou fausse dans un modèle à plus de deux éléments. Mais alors, de quoi parle-t-on quand on asserte la formule de Hodges? Des objets (du modèle)? Des stratégies gagnantes? Le problème est plus général: il est lié à l'équivalence entre logique IF du premier ordre et un fragment du second ordre[17].

La logique IF est une logique en plein développement: tout n'est pas fixé, et beaucoup de questions sont encore «négociables». Il reste que le phénomène d'indépendance informationnelle qu'elle met au jour, indépendamment des difficultés rencontrées par l'interprétation des quantificateurs IF, paraît bien être un trait central et jusque-là ignoré de la quantification.

Références bibliographiques

HINTIKKA J. (1985). Is Truth Ineffable? In Coll. *Les Formes actuelles du vrai (Entretiens de Palerme 1985)*. Palerme: Enchiidion. 1989. 89-120. [Trad. fr. Schmitz F. dans Hintikka (1994). 9-47].

HINTIKKA J. (1987). Game Theoretical Semantics as a Synthesis of Verificationist and Truth-Conditional Meaning Theories. Dans le Pore E. (ed). *New Directions in Semantics*. Academic Press. 235-258. [Trad. fr. Lavand N. dans Hintikka (1994). 136-167].

HINTIKKA J. (1994). *La Vérité est-elle ineffable? et autres essais.* Combas: L'Eclat.

HINTIKKA J. (1996). *The Principles of Mathematics Revisited*. Cambridge University Press.

HINTIKKA J. (1998). Truth Definitions, Skolem Functions and Axiomatic Set Theory. *The Bulletin of Symbolic Logic* 4, 303-337.

HINTIKKA J. (2002). Hyperclassical Logic (A.K.A. IF Logic) and its Implications for Logical Theory. *The Bulletin of Symbolic Logic* 8, 404-423.

[17] Voir Rebuschi (2003) pour une discussion plus détaillée.

HINTIKKA J., SANDU G. (1996). A Revolution in Logic? *Nordic Journal of Philosophical Logic* 1/2. 169-183.

HINTIKKA J., SANDU G. (1997). Game-Theoretical Semantics. Dans van Benthem J., Ter Meulen A. (eds). *Handbook of logic and language*. Amsterdam: Elsvier / MIT Press. 361-410.

HODGES W. (1997). Compositional Semantics for a Language of Imperfect Information. *Journal of the IGPL* 5, 539-563.

JANASIK T., SANDU G. (2003). Dynamic Game Semantics. Dans Peregrin J. (ed). *Meaning: The Dynamic Turn*. Amsterdam: Elsevier.

PIETARINEN A. (2000). Knowledge *De Re*, Identity, and Games of Imperfect Information. Preprint.

PIETARINEN A., SANDU G. (2000). Games in Philosophical Logic. *Nordic Journal of Philosophical Logic* 4, 143-173.

REBUSCHI M. (2003). About Games and Substitution. Dans Peregrin J. (ed). *Meaning: The Dynamic Turn*. Amsterdam: Elsevier.

SANDU G. (1997). On the Theory of Anaphora: Dynamic Predicate Logic vs Game-Theoretical Semantics. *Linguistics and Philosophy* 20, 147-174.

TULENHEIMO T. (2002). On IF Modal Logic and its Expressive Power. Preprint.