Logique classique du premier ordre

Jean Goubault-Larrecq

April 3, 2001

Dans ce chapitre, nous présentons une logique bien plus expressive que la plupart des logiques propositionnelles des chapitres précédents. Ceci est fait en raffinant le langage de la logique, et en remplaçant les variables propositionnelles par des formules plus élaborées décrivant les propriétés de valeurs dans un domaine d'intérêt. On aboutit ainsi à la *logique du premier ordre*, ou *calcul des prédicats*. La plupart des logiques propositionnelles peuvent être étendues au premier ordre (et même à l'ordre supérieur). Nous présentons ici la logique du premier ordre *classique*.

1 Définitions

1.1 Syntaxe

La syntaxe des logiques d'ordre un est à deux niveaux : un pour les termes, qui représentent les objets étudiés, et un pour les formules, autrement dit les propriétés de et entre ces objets.

Soit $\mathcal{V}$ un ensemble infini dénombrable de *variables d'individu* x, y, …, $\mathcal{F}$ un ensemble au plus dénombrable de *symboles de fonction* f, g, …, et $\mathcal{P}$ un ensemble au plus dénombrable, et disjoint du précédent, de *symboles de prédiat* P, Q, … Soit aussi $\mathbf{m} : \mathcal{F} \cup \mathcal{P} \to \mathbb{N}$ une application dite d'*arité*; l'arité de f est définie par $m = \mathbf{m}(f)$.

Définition 1 (Termes, Formules) *Nous appelons* termes *s, t, … les éléments du plus petit ensemble $\mathcal{T}$ tel que :*

- *toute variable est un terme;*
- *pour tout $f \in \mathcal{F}$ d'arité m, pour tous termes t_1, …, t_m dans $\mathcal{T}$, le $(m+1)$-uplet $(f, t_1, \ldots, t_m)$ est aussi dans $\mathcal{T}$.*

*Ce $(m+1)$-uplet, appelé l'*application *de f à t_1, …, t_m, est écrit $f(t_1, \ldots, t_m)$ par commodité.*

Nous appelons formule atomique *ou* atome *A, B, … toute application $P(t_1, \ldots, t_m)$ d'un symbole de prédicat P d'arité m à m termes.*

Les formules *Φ, Ψ, … sont les formules atomiques, les négations $\neg\Phi$ de formules, les conjonctions $\Phi \wedge \Phi'$, les disjonctions $\Phi \vee \Phi'$, les implications $\Phi \Rightarrow \Phi'$, les* quantifications universelles *$\forall x \cdot \Phi$ et les quantifications existentielles $\exists x \cdot \Phi$.*

Nous écrirons aussi $\forall x_1 x_2 \ldots x_n \cdot \Phi$ pour $\forall x_1 \cdot \forall x_2 \cdot \ldots \forall x_n \cdot \Phi$, et $\exists x_1 x_2 \ldots x_n \cdot \Phi$ pour $\exists x_1 \cdot \exists x_2 \cdot \ldots \exists x_n \cdot \Phi$.

Intuitivement, les termes représentent des valeurs dans un domaine donné d'objets. Les constantes sont codées comme des symboles de fonction d'arité 0, et les autres symboles de fonctions représentent des opérations de base sur les valeurs (addition, multiplication, par exemple). Les formules atomiques représentent des propriétés de base des valeurs (par exemple, être inférieur ou égal à, être impair, etc.), et sont combinées entre elles à l'aide des connecteurs propositionnels usuels, ainsi que des quantifications : $\forall x \cdot \Phi$ ("pour tout x, Φ") signifie que, si l'on interprète Φ comme une fonction $x \mapsto f(x)$ envoyant une valeur dénotée x vers la valeur de vérité de Φ, alors $f(v)$ est vraie de toutes les valeurs v; de même, $\exists x \cdot \Phi$ ("il existe x tel que Φ") est vraie dès que $f(v)$ est vraie pour au moins une valeur v.

Définition 2 (Variables libres, liées) *Si t est un terme ou une formule, nous définissons l'ensemble* $\mathrm{fv}(t)$ *de ses* variables libres *et l'ensemble* $\mathrm{bv}(t)$ *de ses* variables liées *par récurrence structurelle :*

- $\mathrm{fv}(x) = \{x\}$, $\mathrm{bv}(x) = \emptyset$ *pour tout* $x \in \mathcal{V}$*;*
- $\mathrm{fv}(f(t_1, \ldots, t_m)) = \mathrm{fv}(t_1) \cup \ldots \cup \mathrm{fv}(t_m)$, $\mathrm{bv}(f(t_1, \ldots, t_m)) = \emptyset$*;*
- $\mathrm{fv}(P(t_1, \ldots, t_m)) = \mathrm{fv}(t_1) \cup \ldots \cup \mathrm{fv}(t_m)$, $\mathrm{bv}(P(t_1, \ldots, t_m)) = \emptyset$*;*
- $\mathrm{fv}(\neg\Phi) = \mathrm{fv}(\Phi)$, $\mathrm{bv}(\neg\Phi) = \mathrm{bv}(\Phi)$*;*
- $\mathrm{fv}(\Phi\vee\Phi') = \mathrm{fv}(\Phi\wedge\Phi') = \mathrm{fv}(\Phi\Rightarrow\Phi') = \mathrm{fv}(\Phi)\cup\mathrm{fv}(\Phi')$, $\mathrm{bv}(\Phi\vee\Phi') = \mathrm{bv}(\Phi\wedge\Phi') = \mathrm{bv}(\Phi\Rightarrow\Phi') = \mathrm{bv}(\Phi) \cup \mathrm{bv}(\Phi')$*;*
- $\mathrm{fv}(\forall x \cdot \Phi) = \mathrm{fv}(\exists x \cdot \Phi) = \mathrm{fv}(\Phi) \setminus \{x\}$, $\mathrm{bv}(\forall x \cdot \Phi) = \mathrm{bv}(\exists x \cdot \Phi) = \mathrm{bv}(\Phi) \cup \{x\}$.

Un terme ou une formule est dit clos *si l'ensemble de ses variables libres est vide. Un* énoncé *est une formule close.*

Une théorie T *est un ensemble d'énoncés.*

Contrairement au cas propositionnel, une variable x peut apparaître dans une formule sans y apparaître libre : par exemple, $\forall x \cdot P(x)$ n'a pas de variable libre, et x n'apparaît que liée (ici, par la quantification universelle).

Comme dans le cas propositionnel, nous pouvons définir des substitutions, à ceci près que nous ne remplaçons pas des atomes mais des variables d'individu. Les définitions sont les mêmes, parce qu'en fait ce sont des définitions sur des éléments d'algèbres libres arbitraires :

Définition 3 (Substitution) *Une* substitution σ *est une application de* $\mathcal{V}$ *vers* $\mathcal{T}$ *telle que l'ensemble* $\mathrm{dom}\ \sigma = \{x \in \mathcal{V} \mid x \neq \sigma(x)\}$*, appelé le* domaine *de* σ*, est fini.*

*L'*image *de* σ *est par définition* $\mathrm{rng}\ \sigma = \{\sigma(x) \mid x \in \mathrm{dom}\ \sigma\}$*, et on pose* $\mathrm{yield}\ \sigma = \bigcup\{\mathrm{fv}(t) \mid t \in \mathrm{rng}\ \sigma\}$.

Nous écrivons aussi σ *sous la forme* $[\sigma(x_1)/x_1, \ldots, \sigma(x_n)/x_n]$*, où* $x_1, \ldots, x_n$ *contiennent toutes les variables de* $\mathrm{dom}\ \sigma$ *et sont distinctes deux à deux.*

En particulier, $[]$ *est la substitution vide (ou identité).*

Comme dans le cas propositionnel, σ s'étend en un morphisme unique $t \mapsto t\sigma$ de $\mathcal{T}$ vers $\mathcal{T}$ tel que :

- $x\sigma = \sigma(x)$
- $f(t_1, \ldots, t_m)\sigma = f(t_1\sigma, \ldots, t_m\sigma)$

Définition 4 (Instances) *Soit t un terme. Ses instances sont tous les termes de la forme $t\sigma$, où σ est une substitution.*

La composition $\sigma\sigma'$ *de deux substitutions est définie par* $t(\sigma\sigma') = (t\sigma)\sigma'$.

Une substitution σ' *est dite* moins générale que σ*, ce que nous notons* $\sigma' \preceq \sigma$*, si et seulement s'il existe une substitution* σ'' *telle que* $\sigma\sigma'' = \sigma'$.

La composition de substitutions est bien définie (exercice). De plus, elle est associative et a $[]$ comme élément neutre, et $\preceq$ est un préordre. Intuitivement, σ' est moins générale que σ si et seulement si toutes les instances de $t\sigma'$ sont aussi des instances de $t\sigma$, pour tout terme t; en bref, une substitution est plus générale qu'une autre si elle a au moins toutes les instances de l'autre.

Nous étendons ensuite la substitution aux formules par récurrence structurelle :

- $P(t_1, \ldots, t_m)\sigma = P(t_1\sigma, \ldots, t_m\sigma)$
- $(\neg\Phi)\sigma = \neg(\Phi\sigma)$

- $(\Phi \oplus \Phi')\sigma = \Phi\sigma \oplus \Phi'\sigma$, où $\oplus \in \{\vee, \wedge, \Rightarrow\}$
- pour tout $Q \in \{\forall, \exists\}$, $(Qx \cdot \Phi)\sigma = (Qx' \cdot \Phi[x'/x]\sigma)$, où x' est une variable hors de yield $\sigma \cup (\mathrm{fv}(\Phi) \setminus \{x\})$.

Comme la dernière règle n'est pas déterministe, il y a plusieurs instances d'une formule quantifiée par σ. Cependant, toutes ces formules seront équivalentes tant sémantiquement que syntaxiquement. Le fait de transformer $Qx \cdot \Phi$ en $Qx' \cdot \Phi[x'/x]$ avec $x' \notin \mathrm{fv}(\Phi) \setminus \{x\}$ sera bénin, et est appelé *α-renommage*. (Exercice : vérifier que les règles sémantiques de la section 2 et les systèmes de preuve de la section 3 sont effectivement invariants par α-renommage.)

2 Sémantique

Comme dans le cas propositionnel, la sémantique de la logique du premier ordre est décrite par une interprétation. Cependant le langage de la logique du premier ordre est plus riche, et de nouvelles définitions sont nécessaires :

Définition 5 (Interprétation) *Une interprétation I est un ensemble non vide D_I, appelé le* domaine *de l'interprétation, muni d'une application $I(f)$ de D_I^m vers D_I pour chaque symbole de fonction f d'arité m, et d'une application $I(P)$ de D_I^m vers $\mathbb{B}$ pour chaque symbole de prédicat P d'arité m.*

Une affectation *ρ est une application de $\mathcal{V}$ vers D_I. Soit $\mathcal{A} = \mathcal{V} \to D_I$ l'ensemble de toutes les affectations.*

Intuitivement, une interprétation fournit un ensemble de valeurs D_I que nous décrivons à l'aide des termes, un ensemble d'opérations $I(f)$ sur ces valeurs, et un ensemble de propriétés de base $I(P)$ sur des m-uplets de valeurs.

Contrairement au cas propositionnel, les interprétations et les affectations sont des objets différents : une affectation donne une valeur à chaque variable, alors qu'une interprétation décrit le domaine des valeurs et la sémantique des symboles de fonctions et de prédicats.

Les interprétations nous permettent de définir la sémantique des termes du premier ordre (comme des valeurs dans D_I) et des formules (comme des valeurs de vérité dans $\mathbb{B}$) :

Définition 6 (Sémantique) *Pour toute affectation ρ, soit $\rho[v/x]$ l'affectation envoyant chaque variable y autre que x vers $\rho(y)$, et x vers v.*

Dans une interprétation I, et modulo l'affectation ρ, la sémantique des termes et des formules est définie par :

- $[\![x]\!]I\rho = \rho(x)$;
- $[\![f(t_1, \ldots, t_m)]\!]I\rho = I(f)([\![t_1]\!]I\rho, \ldots, [\![t_m]\!]I\rho)$;
- $[\![P(t_1, \ldots, t_m)]\!]I\rho = I(P)([\![t_1]\!]I\rho, \ldots, [\![t_m]\!]I\rho)$;
- $[\![\neg\Phi]\!]I\rho = \overline{\neg}[\![\Phi]\!]I\rho$;
- $[\![\Phi \vee \Phi']\!]I\rho = [\![\Phi]\!]I\rho \overline{\vee} [\![\Phi']\!]I\rho$;
- $[\![\Phi \wedge \Phi']\!]I\rho = [\![\Phi]\!]I\rho \overline{\wedge} [\![\Phi']\!]I\rho$;
- $[\![\Phi \Rightarrow \Phi']\!]I\rho = [\![\Phi]\!]I\rho \overline{\Rightarrow} [\![\Phi']\!]I\rho$;
- $[\![\forall x \cdot \Phi]\!]I\rho = \bigwedge_{v \in D_I} [\![\Phi]\!]I(\rho[v/x])$;
- $[\![\exists x \cdot \Phi]\!]I\rho = \bigvee_{v \in D_I} [\![\Phi]\!]I(\rho[v/x])$;

$$\frac{\Gamma \longrightarrow \Phi[y/x]}{\Gamma \longrightarrow \forall x \cdot \Phi}\ (\forall I) \qquad \frac{\Gamma \longrightarrow \forall x \cdot \Phi}{\Gamma \longrightarrow \Phi[t/x]}\ (\forall E)$$
(y non libre dans Γ)

$$\frac{\Gamma \longrightarrow \Phi[t/x]}{\Gamma \longrightarrow \exists x \cdot \Phi}\ (\exists I) \qquad \frac{\Gamma \longrightarrow \exists x \cdot \Phi \qquad \Gamma, \Phi[y/x] \longrightarrow \Psi}{\Gamma \longrightarrow \Psi}\ (\exists E)$$
(y non libre dans Γ, Ψ)

Figure 1: Déduction naturelle en forme de séquents : règles des quantificateurs

où $\bigwedge$ est la conjonction distribuée et $\bigvee$ et la disjonction distribuée, et $\bar{\neg}$, $\bar{\wedge}$, $\bar{\vee}$ et $\bar{\Rightarrow}$ sont les fonctions booléennes usuelles.

Une formule est valide *si elle est vraie dans toute interprétation et toute affectation; sinon, elle est* invalide. *Une formule est* insatisfiable *is elle est fausse dans toute interprétation et toute affectation; sinon, elle est* satisfiable.

Une interprétation dans laquelle une formule Φ est satisfaite est appelée un modèle *de Φ. Un modèle d'une théorie est un modèle de toutes les formules qu'elle contient. Si F est une formule ou une théorie, nous notons $I \models F$ la relation "I est un modèle de F."*

La notion de conséquence sémantique, *notée aussi $\models$, lie une théorie (resp. une formule) F à une autre théorie (resp. formule) F' : $F \models F'$ si tout modèle de F est aussi un modèle de F'.*

3 Systèmes de preuve

Les systèmes de preuve pour la logique du premier ordre sont tous des extensions de ceux pour la logique propositionnelle, où la mention "variable propositionnelle" doit être remplacée par "atome", et où de nouvelles règles sont ajoutées pour tenir compte des quantifications.

Par exemple, la logique classique du premier ordre a des systèmes de Hilbert. Un exemple est $\mathcal{SKC}_1$:

- Axiomes :
  - (K) $\Phi \Rightarrow \Phi' \Rightarrow \Phi$,
  - (S) $(\Phi \Rightarrow \Phi' \Rightarrow \Phi'') \Rightarrow (\Phi \Rightarrow \Phi') \Rightarrow (\Phi \Rightarrow \Phi'')$,
  - (C) $\neg\neg\Phi \Rightarrow \Phi$
  - (E) $(\forall x \cdot \Phi) \Rightarrow \Phi[t/x]$ pour toute formule Φ, toute variable x et tout terme t,
  - (I) $(\forall x \cdot \Phi \Rightarrow \Phi') \Rightarrow \Phi \Rightarrow \forall x \cdot \Phi'$, pour toutes formules Φ, Φ' et pour toute variable x telle que x n'est pas libre dans Φ.
- Règles :
  - (MP) de Φ et $\Phi \Rightarrow \Phi'$, déduire Φ', pour toutes Φ, Φ',
  - (Gen) de Φ, déduire $\forall x \cdot \Phi$, pour toute formule Φ et toute variable x.

où seuls les axiomes (E), (I) et la règle (Gen) sont nouveaux. Comme dans le système $\mathcal{SKC}$, seuls $\forall$, $\Rightarrow$ et $\mathbf{F}$ sont des notations de base dans ce système, de sorte que $\neg\Phi$ est une abréviation de $\Phi \Rightarrow \mathbf{F}$ en particulier. En fait, $\exists$ est aussi une abréviation : $\exists x \cdot \Phi$ est une abréviation de $\neg\forall x \cdot \neg\Phi$.

La règle (Gen) est la *règle de généralisation.* Si nous avons réussi à prouver Φ, alors Φ est valide (si la logique est cohérente), donc elle est vraie dans toute affectation, donc pour toute valeur de x, de sorte que $\forall x \cdot \Phi$ doit être valide aussi, donc prouvable (si nous voulons que le système de preuve soit complet).

$$\frac{}{\Gamma, \Phi \longrightarrow \Delta, \Phi}\ \mathrm{Ax}$$

$$\frac{\Gamma, \Phi, \Phi' \longrightarrow \Delta}{\Gamma, \Phi \wedge \Phi' \longrightarrow \Delta}\ \wedge\mathrm{L} \qquad \frac{\Gamma \longrightarrow \Delta, \Phi \qquad \Gamma \longrightarrow \Delta, \Phi'}{\Gamma \longrightarrow \Delta, \Phi \wedge \Phi'}\ \wedge\mathrm{R}$$

$$\frac{\Gamma, \Phi \longrightarrow \Delta \qquad \Gamma, \Phi' \longrightarrow \Delta}{\Gamma, \Phi \vee \Phi' \longrightarrow \Delta}\ \vee\mathrm{L} \qquad \frac{\Gamma \longrightarrow \Delta, \Phi, \Phi'}{\Gamma \longrightarrow \Delta, \Phi \vee \Phi'}\ \vee\mathrm{R}$$

$$\frac{\Gamma \longrightarrow \Phi, \Delta \qquad \Gamma, \Phi' \longrightarrow \Delta}{\Gamma, \Phi \Rightarrow \Phi' \longrightarrow \Delta}\ \Rightarrow\mathrm{L} \qquad \frac{\Gamma, \Phi \longrightarrow \Delta, \Phi'}{\Gamma \longrightarrow \Delta, \Phi \Rightarrow \Phi'}\ \Rightarrow\mathrm{R}$$

$$\frac{\Gamma \longrightarrow \Delta, \Phi}{\Gamma, \neg\Phi \longrightarrow \Delta}\ \neg\mathrm{L} \qquad \frac{\Gamma, \Phi \longrightarrow \Delta}{\Gamma \longrightarrow \Delta, \neg\Phi}\ \neg\mathrm{R}$$

$$\frac{\Gamma, \Phi[t/x] \longrightarrow \Delta}{\Gamma, \forall x \cdot \Phi \longrightarrow \Delta}\ \forall\mathrm{L} \qquad \frac{\Gamma \longrightarrow \Phi[y/x], \Delta}{\Gamma \longrightarrow \forall x \cdot \Phi, \Delta}\ \forall\mathrm{R}$$

(y non libre dans Γ, Δ)

$$\frac{\Gamma, \Phi[y/x] \longrightarrow \Delta}{\Gamma, \exists x \cdot \Phi \longrightarrow \Delta}\ \exists\mathrm{L} \qquad \frac{\Gamma \longrightarrow \Phi[t/x], \Delta}{\Gamma \longrightarrow \exists x \cdot \Phi, \Delta}\ \exists\mathrm{R}$$

(y non libre dans Γ, Δ)

$$\frac{\Gamma \longrightarrow \Delta, \Phi \qquad \Gamma', \Phi \longrightarrow \Delta'}{\Gamma, \Gamma' \longrightarrow \Delta, \Delta'}\ \mathrm{Cut}$$

Figure 2: Le système **LK** de Gentzen

La logique classique du premier ordre a aussi des systèmes de déduction naturelle. Par exemple, nous pouvons ajouter les règles de la figure 1 à celles du système $\mathcal{ND}$ du chapitre sur la logique propositionnelle classique.

Dans le but d'analyser la structure des preuves, nous serons davantage intéressés par le système **LK** de Gentzen (cf. figure 2). C'est une extension du système $\mathbf{LK}_0$ du chapitre sur la logique propositionnelle classique, auquel nous avons ajouté les règles ∀L, ∀R, ∃L, ∃R sur les quantifications.

Tous ces systèmes sont corrects et complets pour la logique du premier ordre. Nous ne le prouverons que pour **LK**, en section 5.3 après avoir présenté toutes les notions nécessaires.

4 Pouvoir d'expression

Les logiques d'ordre un sont beaucoup plus expressives que les logiques propositionnelles correspondantes, et en particulier la logique classique du premier ordre est beaucoup plus expressive que la logique propositionnelle classique.

D'abord, la logique du premier ordre peut exprimer des propriétés sur des domaines infinis. En logique propositionnelle, pour exprimer des propriétés sur des valeurs, nous ne pouvions que coder ces dernières sur un nombre fini de variables propositionnelles. Ceci ne permet qu'un nombre fini de valeurs, puisqu'il n'y a qu'un nombre fini d'affectations portant sur les variables propositionnelles libres dans une formule propositionnelle. En logique du premier ordre, nous pouvons construire des formules comme :

$$\begin{array}{ll} & \forall x, y, z \cdot x < y \wedge y < z \Rightarrow x < z \\ \wedge & \forall x, y \cdot x < y \Rightarrow \exists z \cdot P(x, z) \wedge \neg P(y, z) \\ \wedge & \forall x \cdot \exists y \cdot x < y \end{array}$$

où P et $<$ sont deux symboles de prédicat binaires, $<$ étant écrit en notation infixe. Cette formule est satisfiable : interprétons $<$ comme la relation "strictement plus petit que" sur les entiers, et P comme la relation "plus petit ou égal à". Mais ses modèles sont tous infinis : $<$ doit dénoter un ordre partiel strict de par les deux premières lignes, et ne peut pas avoir de borne supérieure par la troisième.

Nous pouvons fabriquer des théories plus utiles. Par exemple, prenons un symbole de prédicat (infixe) binaire $<$, et disons qu'il représente un ordre strict, c'est-à-dire qu'il est irréflexif :

$$\forall x \cdot \neg x < x$$

et transitif :

$$\forall x \cdot \forall y \cdot \forall z \cdot x < y \wedge y < z \Rightarrow x < z$$

L'égalité $\doteq$ est un peu plus difficile à axiomatiser, car nous voulons pouvoir remplacer les termes par des termes égaux, autrement dit si $a \doteq b$ est vrai, alors $s[a/x] \doteq t[a/x]$ est vrai aussi pour tous termes s et t; aussi, $\Phi[a/x]$ doit être logiquement équivalent à $\Phi[b/x]$. L'astuce pour axiomatiser l'égalité est de d'abord fixer le langage sur lequel nous travaillerons, autrement dit l'ensemble des symboles de fonctions et de prédicats dont nous aurons besoin, et de produire un axiome par symbole pour exprime que chacun préserve l'égalité. Ceci mène à la théorie suivante :

- Réflexivité : $\forall x \cdot x \doteq x$;
- Symétrie : $\forall x, y \cdot x \doteq y \Rightarrow y \doteq x$;
- Congruence fonctionnelle : pour tout symbole de fonction f, d'arité n :

$$\begin{array}{l} \forall x_1 \cdot \ldots \forall x_n \cdot \forall y_1 \cdot \ldots \forall y_n \cdot \\ \quad x_1 \doteq y_1 \wedge \ldots \wedge x_n \doteq y_n \Rightarrow f(x_1, \ldots, x_n) \doteq f(y_1, \ldots, y_n) \end{array}$$

- Congruence prédicative : pour tout symbole de prédicat P, d'arité n :

$$\begin{array}{l} \forall x_1 \cdot \ldots \forall x_n \cdot \forall y_1 \cdot \ldots \forall y_n \cdot \\ x_1 \doteq y_1 \wedge \ldots \wedge x_n \doteq y_n \wedge P(x_1, \ldots, x_n) \Rightarrow P(y_1, \ldots, y_n) \end{array}$$

Supposons que nous souhaitions prouver une formule. Son langage est fini, donc nous n'avons besoin que d'un nombre fini d'axiomes pour $\doteq$. Si nous devons étendre ce langage, par exemple par l'introduction de nouveaux symboles de fonction (comme des symboles de Skolem, voir Section 5.1), nous devrons alors ajouter les axiomes de congruence correspondants.

Grâce à de telles constructions, nous pouvons traduire les logiques modales, ainsi que la logique intuitionniste en terme de logique classique du premier ordre, en exprimant leur sémantique de Kripke. Ceci est cependant une mauvaise idée en général, car les procédures de recherche de preuve spécialisées pour les logiques modales ou intuitionniste sont en général plus efficaces que les procédures générales pour la logique du premier ordre.

Une grande partie des mathématiques peut être formalisée en logique du premier ordre avec égalité. Par exemple, la théorie des groupes étend celle de l'égalité en ajoutant les axiomes suivants sur le langage 0 (constante), + (fonction binaire) :

$$\begin{array}{ll} \forall x \cdot 0 + x \doteq x \wedge x \doteq x + 0 & \text{(0 est élément neutre)} \\ \forall x \cdot \forall y \cdot \forall z \cdot x + (y + z) \doteq (x + y) + z & \text{(associativité)} \\ \forall x \cdot \exists y \cdot x + y \doteq 0 \wedge y + x \doteq 0 & \text{(inverse)} \end{array}$$

La théorie des anneaux ajoute de nouveaux symboles de fonction 1 (constante), $*$ (fonction binaire) et les axiomes :

$$
\begin{array}{ll}
\forall x \cdot \forall y \cdot x + y \doteq y + x & \text{(commutativité de +)} \\
\forall x \cdot 1 * x \doteq x \wedge x \doteq x * 1 & \text{(1 est neutre pour } *\text{)} \\
\forall x \cdot \forall y \cdot \forall z \cdot x * (y * z) \doteq (x * y) * z & \text{(associativité)} \\
\forall x \cdot \forall y \cdot \forall z \cdot x * (y + z) \doteq x * y + x * z & \\
\forall x \cdot \forall y \cdot \forall z \cdot (y + z) * x \doteq y * x + z * x & \text{(distributivité)}
\end{array}
$$

Les corps ajoutent l'axiome :

$$\forall x \cdot x \not\doteq 0 \Rightarrow \exists y \cdot x * y \doteq 1 \wedge y * x \doteq 1$$

Nous pouvons de même développer des théories des corps commutatifs, des modules, des espaces vectoriels, et ainsi de suite.

Ces théories décrivent les notions mathématiques correspondantes de façon adéquate. Cependant, il existe des notions qui ne peuvent pas être capturées précisément par aucune théorie du premier ordre récursivement énumérable. Nous disons que de telles notions ne sont pas *axiomatisables au premier ordre*. (Trivialement, toute notion mathématique peut être capturée par une théorie du premier ordre, à savoir celle qui liste tous les énoncés valides portant sur la notion; mais cette théorie n'est pas forcément récursivement énumérable.) Par exemple, il n'y a pas de théorie du premier ordre axiomatisable dont les modèles sont tous les groupes finis, et seulement les groupes finis. (Bien que nous puissions décrire les groupes de cardinal p, pour tout entier p fixé.) Nous ne pouvons pas traduire le μ-calcul modal en logique du premier ordre fidèlement non plus, autrement dit nous ne pouvons pas axiomatiser les constructions de la logique en logique du premier ordre. C'est parce que les règles de point fixe du calcul ne peuvent pas être traduites fidèlement en logique du premier ordre. Pour la même raison, il n'y a pas d'axiomatisation au premier ordre dont le seul modèle soit l'ensemble $\mathbb{N}$ des entiers naturels muni des opérations arithmétiques usuelles.

Nous pouvons quand même construire une forme d'arithmétique appelée *arithmétique de Peano du premier ordre* $\mathbf{PA}_1$, mais il nous faudra un nombre infini d'axiomes, et elle ne décrira pas complètement l'arithmétique, bien qu'elle en décrive une bonne partie. Le langage de $\mathbf{PA}_1$ consiste en la constante 0, la fonction unaire s (successeur), les fonctions binaires $+$ et $*$ (addition et multiplication, que nous écrivons en infixe), les relations binaires $\doteq$ et $\dot{\leq}$ (égalité et relation “inférieur ou égal à”, écrites en infixe; nous utilisons des points pour les distinguer des affirmations d'égalité et d'inégalité $=$ et $\leq$). Les axiomes de $\mathbf{PA}_1$ — les éléments de la théorie $\mathbf{PA}_1$ — sont :

- Axiomes d'égalité :
  - Réflexivité : $\forall x \cdot x \doteq x$;
  - Symétrie : $\forall x, y \cdot x \doteq y \Rightarrow y \doteq x$;
  - Transitivité : $\forall x, y, z \cdot x \doteq y \wedge y \doteq z \Rightarrow x \doteq z$;
  - Congruence:

$$
\begin{array}{l}
\forall x, y \cdot x \doteq y \Rightarrow s(x) \doteq s(y) \\
\forall x, y, z, t \cdot x \doteq z \wedge y \doteq t \Rightarrow x + y \doteq z + t \\
\forall x, y, z, t \cdot x \doteq z \wedge y \doteq t \Rightarrow x * y \doteq z * t \\
\forall x, y, z, t \cdot x \doteq z \wedge y \doteq t \wedge x \dot{\leq} y \Rightarrow z \dot{\leq} t
\end{array}
$$

- Axiomes de Peano de base :
  - $\forall x \cdot x \doteq 0 \vee \exists y \cdot x \doteq s(y)$ (tout entier est 0 ou un successeur);
  - $\forall x \cdot s(x) \not\doteq 0$ (0 n'est le successeur d'aucun entier);
  - $\forall x, y \cdot s(x) \doteq s(y) \Rightarrow x \doteq y$ (s est injective; en particulier, tout entier non nul a un prédécesseur unique);
  - $\forall x \cdot x + 0 \doteq x$, $\forall x, y \cdot x + s(y) \doteq s(x + y)$ (définition de l'addition);

– $\forall x \cdot x * 0 \doteq 0$, $\forall x, y \cdot x * s(y) \doteq (x * y) + x$ (définition de la multiplication);
– $\forall x, y \cdot x \dot{\leq} y \Leftrightarrow \exists z \cdot x + z \doteq y$ (définition de $\dot{\leq}$);

- Schéma de récurrence : pour toute formule Φ du premier ordre dans le langage de $\mathbf{PA}_1$, et pour toute variable x :

$$\Phi[0/x] \wedge (\forall x \cdot \Phi \Rightarrow \Phi[s(x)/x]) \Rightarrow \forall x \cdot \Phi$$

Le schéma de récurrence est une collection infinie d'axiomes. Nous voudrions écrire à la place :

$$\forall P \cdot P(0) \wedge (\forall x \cdot P(x) \Rightarrow P(s(x))) \Rightarrow \forall x \cdot \Phi$$

où P est une variable de prédicat. (Ceci est l'*axiome de récurrence du second ordre*.) Nous ne pouvons pas le faire en logique du premier ordre, puisque nous ne pouvons quantifier que sur les variables, et les variables dénotent des valeurs, pas des prédicats sur les valeurs. Mais les dommages sont limités : bien que $\mathbf{PA}_1$ ait une infinité d'axiomes, l'ensemble de ses axiomes est récursif, ce qui signifie que nous pouvons décider par algorithme si une formule donnée est un axiome de $\mathbf{PA}_1$ ou non.

Un modèle de $\mathbf{PA}_1$ est le modèle souhaité : l'ensemble $\mathbb{N}$ des entiers naturels, où la constante 0 est interprétée comme zéro, s est interprétée comme l'application envoyant n vers $n+1$, $+$ est interprétée comme l'addition, $*$ comme la multiplication, $\doteq$ comme l'égalité et $\dot{\leq}$ comme l'ordre naturel $\leq$ vérifie tous les axiomes. Il s'agit du *modèle standard* de $\mathbf{PA}_1$.

Mais $\mathbf{PA}_1$ a d'autres modèles, les *modèles non standard* de l'arithmétique [Rob66], comme on s'y attendait au su du fait que l'arithmétique de Peano n'est pas axiomatisable en logique du premier ordre. Ceci est une limitation inhérente à la logique du premier ordre, et peut être expliqué dans ce cas par le théorème de Löwenheim-Skolem, que nous présenterons en section 5.2.

La logique du premier ordre peut non seulement exprimer cette forme faible d'arithmétique, mais elle est encore assez expressive pour formaliser la *théorie des ensembles*, par exemple au moyen des axiomes de Zermelo et Frænkel, sur laquelle toutes les mathématiques peuvent être fondées (à l'exception de quelques points n'intéressant que les logiciens). Voir [Joh92] pour une discussion de la théorie des ensembles de Zermelo-Frænkel, de ses concepts et résultats fondamentaux.

5 Propriétés méta-mathématiques

La logique du premier a de nombreuses propriétés logiques intéressantes, et la plupart peuvent être approchées soit du point de vue sémantique, au moyen d'interprétations (la *théorie des modèles*), soit du point de vue syntaxique, au moyen d'arguments portant sur les preuves, disons en **LK** (la *théorie de la preuve*).

5.1 Formes prénexes et de Skolem

Notre but en démonstration automatique est de découvrir si des formules sont valides (resp. prouvables). Cependant, avant de chercher des preuves, il est commode de simplifier les formules à prouver. Étant donnée une formule du premier ordre Φ, nous voulons calculer une formule Φ' plus simple (en un sens à préciser) dont la validité (resp. la prouvabilité) soit équivalente à celle de Φ.

Définition 7 (Prénexe) *Une formule est dite* prénexe *si et seulement si elle est de la forme $Q_1 x_1 \cdot \ldots Q_n x_n \cdot \Phi$, où $Q_1, \ldots, Q_n$ sont des quantificateurs et Φ est sans quantificateur.*

Que Φ soit sans quantificateur signifie qu'elle a été construite sans utiliser de quantifications. Autrement dit, Φ est en fait une *formule propositionnelle* où les atomes jouent le rôle de variables propositionnelles.

Théorème 8 *Pour toute formule du premier ordre Φ, il existe une formule prénexe Φ', calculable à partir de Φ, telle que Φ et Φ' sont (sémantiquement, resp. prouvablement) équivalentes.*

Preuve : Exercice : l'idée est de pousser les quantificateurs vers l'extérieur par des règles comme $\neg\forall x \cdot \Psi \longrightarrow \exists x \cdot \neg\Psi$ ou $\Psi \wedge (\forall x \cdot \Psi') \longrightarrow \forall x' \cdot \Psi \wedge \Psi'[x'/x]$, où $x' \notin \mathrm{fv}(\Psi)$. La terminaison du processus est obtenu par exemple par l'utilisation d'un ordre récursif sur les chemins (rpo). Si vous avez bien conçu vos règles de transformation, alors l'équivalence sémantique devrait être évidente, alors que l'équivalence pour la prouvabilité devrait être aussi simple, quoique pénible à démontrer formellement. □

Le théorème 8 montre qu'au lieu de travailler avec une formule arbitraire, nous pouvons nous restreindre aux formules prénexes, qui ont une structure plus précise. Nous pouvons encore simplifier, grâce à l'idée suivante.

Soit une formule de la forme, par exemple, $\forall x \cdot \Phi$: elle est valide si et seulement si pour tout x, Φ est vraie. Remplaçons x par une nouvelle constante c dans Φ. Alors $\forall x \cdot \Phi$ est valide si et seulement si, pour toute interprétation de c, $\Phi[c/x]$ est vraie, autrement dit, si et seulement si $\Phi[c/x]$ est valide. En somme, nous pouvons coder la quantification universelle sur x par l'utilisation d'une constante c et la quantification universelle implicite sur toutes les interprétations dans la définition de la validité.

Maintenant considérons une formule de la forme $\exists x \cdot \forall y \cdot \Phi$: elle est valide si et seulement s'il existe une valeur pour x telle que pour tout y, Φ soit vraie. Un contre-exemple de la validité serait, pour chaque valeur donnée de x, une valeur de y telle que Φ ne serait pas vraie; en d'autres termes, un contre-exemple pour y est une application $f(x)$, où f est un nouveau symbole de fonction, et pas seulement une constante c. Sans avoir à recourir à la notion de contre-exemple, nous pouvons dire que $\exists x \cdot \forall y \cdot \Phi$ est valide si et seulement si $\exists x \cdot \Phi[f(x)/y]$ est valide, où la quantification universelle sur y est implicite, cachée dans la définition de la validité.

Formellement, ceci nous mène à :

Définition 9 *Une* formule existentielle *est une formule de la forme $\exists x_1 \cdot \ldots \exists x_n \cdot \Phi$, où Φ est sans quantificateur.*

Une formule universelle *est une formule de la forme $\forall x_1 \cdot \ldots \forall x_n \cdot \Phi$, où Φ est sans quantificateur.*

Ce qui suit est dû pour l'essentiel à Thoralf Skolem pour les formules universelles, mais Jacques Herbrand l'utilisait abondamment pour les formules existentielles :

Théorème 10 (Herbrand-Skolem) *Soit Φ une formule du premier ordre.*

Il existe une formule existentielle Φ', calculable à partir de Φ, telle que Φ' est valide si et seulement si Φ est valide. Nous disons que Φ' est obtenue en herbrandisant *Φ, et que Φ' est une* forme de Herbrand *de Φ.*

De façon duale, il existe une formule universelle Φ'', calculable à partir de Φ, telle que Φ'' est insatisfiable si et seulement si Φ est insatisfiable. Nous disons que Φ'' est obtenue en skolémisant *Φ, et que Φ'' est une* forme de Skolem *de Φ.*

Preuve : Nous pouvons supposer que Φ est en forme prénexe, par le théorème 8. Nous traitons de la skolémisation, l'herbrandisation s'en déduisant en niant la forme de Skolem de la négation. (Et réciproquement.)

Soit donc Φ de la forme $Q_1 x_1 \cdot \ldots Q_n x_n \cdot \Psi$, où Ψ est sans quantificateur. Pour chaque $1 \leq i \leq n$ tel que Q_i est le quantificateur existentiel $\exists$, soit $x_{i_1}, \ldots, x_{i_{m_i}}$ la liste des variables universellement quantifiées d'indices inférieurs à i, soit f_i un nouveau symbole de fonction d'arité m_i et t_i le terme $f_i(x_{i_1}, \ldots, x_{i_{m_i}})$. Alors, soit Ψ' la formule Ψ où toute variable existentiellement quantifiée x_i est remplacée par t_i, et Φ' la formule $\forall x_{i_1} \cdot \ldots \forall x_{i_m} \cdot \Psi'$, où $x_{i_1}, \ldots, x_{i_m}$ sont les variables universellement quantifiées. (Autrement dit, nous remplaçons les variables existentielles par des fonctions nouvelles de toutes les variables universelles qui les précèdent. Dans le cas de l'herbrandisation, nous remplaçons les variables universelles par des fonctions nouvelles de toutes les variables existentielles les précédant.)

Nous affirmons que Φ est satisfiable si et seulement si Φ' l'est. Pour le prouver, nous montrons que tout énoncé $\forall y_1 \cdot \ldots \forall y_k \cdot \exists z \cdot \Phi_1$ est satisfiable si et seulement si $\forall y_1 \cdot \ldots \forall y_k \cdot \Phi_1[f(y_1, \ldots, y_k)/z]$ est satisfiable, où f est un symbole de fonction qui n'apparaît pas dans Φ_1. L'affirmation s'en déduit par récurrence sur le nombre de quantificateurs existentiels dans le préfixe de Φ.

Si $\forall y_1 \cdot \ldots \forall y_k \cdot \Phi_1[f(y_1, \ldots, y_k)/z]$ est satisfiable, alors il y a un modèle I, et une interprétation de f dans I, tels que $\Phi_1[f(y_1, \ldots, y_k)/z]$ est vraie pout toutes les affectations de valeurs v_1, ..., v_k aux variables y_1, ..., y_k respectivement. Alors $I(f)(v_1, \ldots, v_k)$ est une valeur pour z qui rend Φ_1 vraie dans I, pour toute affectation des y_i, $1 \leq i \leq k$. Donc I satisfait $\forall y_1 \cdot \ldots \forall y_k \cdot \exists z \cdot \Phi_1$.

Réciproquement, supposons qu'il existe une interprétation I satisfaisant $\forall y_1 \cdot \ldots \forall y_k \cdot \exists z \cdot \Phi_1$. Pour toute affectation de valeurs v_1 à y_1, ..., v_k à y_k, soit $G(v_1, \ldots, v_k)$ l'ensemble des valeurs de z qui rendent Φ_1 vraie. Par hypothèse, $G(v_1, \ldots, v_k)$ est non vide pour tout k-uplet de valeurs $(v_1, \ldots, v_k)$. Nous utilisons maintenant l'axiome du choix pour en déduire qu'il existe une fonction g envoyant chaque $(v_1, \ldots, v_k)$ vers un élément de $G(v_1, \ldots, v_k)$. Étendons l'interprétation I en une nouvelle interprétation I' telle que $I'(s) = I(s)$ pour tous les symboles dans Φ_1, et telle que $I'(f) = g$, nous voyons alors que I' satisfait $\forall y_1 \cdot \ldots \forall y_k \cdot \Phi_1[f(y_1, \ldots, y_k)/z]$. $\square$

Note : Alors que la conversion en forme prénexe produit une formule logiquement équivalente, l'herbrandisation et la skolémisation ne le garantit pas. L'herbrandisation ne préserve que le statut de validité, et la skolémisation ne préserve que le statut de satisfiabilité.

En fait, l'herbrandisation préserve aussi le statut de prouvabilité, et la skolémisation le statut de cohérence de la formule, ce qui signifie que ce processus est justifié non seulement sémantiquement, mais aussi en ce qui concerne les preuves. Nous le montrerons en section 5.3.

On peut mieux herbrandiser/skolémiser si on évite de mettre la formule en forme prénexe. Définissons deux fonctions, une d'herbrandisation h et une de skolémisation s par [And86] :

- si Φ est atomique, alors $h(\Phi) = s(\Phi) = \Phi$;
- $h(\Phi' \wedge \Phi'') = h(\Phi') \wedge h(\Phi'')$, $s(\Phi' \wedge \Phi'') = s(\Phi') \wedge s(\Phi'')$;
- $h(\Phi' \vee \Phi'') = h(\Phi') \vee h(\Phi'')$, $s(\Phi' \vee \Phi'') = s(\Phi') \vee s(\Phi'')$;
- $h(\neg\Phi') = \neg s(\Phi')$, $s(\neg\Phi') = \neg h(\Phi')$,
- $h(\Phi' \Rightarrow \Phi'') = s(\Phi') \Rightarrow h(\Phi'')$, $s(\Phi' \Rightarrow \Phi'') = h(\Phi') \Rightarrow s(\Phi'')$;
- $h(\exists x \cdot \Phi') = h(\Phi')$, $s(\exists x \cdot \Phi') = s(\Phi')[f(x_1, \ldots, x_n)/x]$, où x_1, ..., x_n sont les variables libres de $\exists x \cdot \Phi'$ et f est un nouveau symbole de fonction n-aire;
- $h(\forall x \cdot \Phi') = h(\Phi')[f(x_1, \ldots, x_n)/x]$, où x_1, ..., x_n sont les variables libres de $\exists x \cdot \Phi'$ et f est un nouveau symbole de fonction n-aire, et $s(\forall x \cdot \Phi') = s(\Phi')$.

(Exercice : montrer que toute formule Φ, $h(\Phi)$ est valide si et seulement si Φ est valide, et que $s(\Phi)$ est satisfiable si et seulement si Φ est satisfiable. On peut même choisir le même symbole de Skolem f pour deux occurrences de formules quantifiées logiquement équivalentes.)

5.2 Théorie sémantique de Herbrand

Maintenant que nous avons restreint la classe des formules à considérer (pour la validité) aux formules existentielles, nous présentons quelques notions fondamentales, découvertes par Jacques Herbrand dans les années 30 :

Définition 11 (Univers de Herbrand) *Soit B un ensemble dit de* constantes de base*, tel que $B \neq \emptyset$ s'il n'y a pas de symbole de fonction* 0*-aire. Soit B_0 le plus petit de ces B, autrement dit $\emptyset$ s'il y a des symboles de fonction* 0*-aires, sinon $\{a\}$ où a est une nouvelle constante.*

*L'ensemble des termes clos formés dans le langage étendu par l'ensemble B de constantes est appelé l'*univers de Herbrand sur B*. L'*univers de Herbrand *est $H(B_0)$.*

Une interprétation de Herbrand *I est un ensemble d'atomes clos dans le langage étendu par l'ensemble de constantes B. Elle définit une interprétation sur H par :*

- *$I(f)(t_1, \ldots, t_n)$, où les t_i sont des termes clos, est le terme clos $f(t_1, \ldots, t_n)$;*
- *$I(P)(t_1, \ldots, t_n)$, où les t_i sont des termes clos, est $\top$ si et seulement si l'atome clos $P(t_1, \ldots, t_n)$ est dans l'ensemble I, et $\perp$ sinon.*

Les univers de Herbrand sont les interprétations les plus simples que nous puissions imaginer, étant donnée la syntaxe de la logique du premier ordre. En effet, ils consistent à interpréter les termes comme des termes clos, comme le lemme suivant le montre. La condition étrange sur la vacuité de B ou de B_0 assure que $H(B)$ est bien le domaine d'une interprétation, autrement dit, que $H(B)$ est non vide.

Lemme 12 *Soit ρ une affectation sur $H(B)$. Alors ρ est une substitution close, et pour toute interprétation de Herbrand I, pour tout terme t, $[\![t]\!]I\rho = t\rho$.*

Preuve : Remarquer que $t\rho$ sur le côté droit signifie t auquel la substitution ρ a été appliquée. La preuve est par une récurrence structurelle immédiate sur t. $\square$

En somme, sur les univers de Herbrand, les substitutions closes *sont* les affectations.

L'intérêt principal des univers et des interprétations de Herbrand tient aux théorèmes suivants. Le lemme suivant dit que les interprétations de Herbrand ne sont pas moins générales que les interprétations générales. C'est-à-dire que si la définition 11 montre que les interprétations de Herbrand définissent des interprétations sur $H(B)$, le lemme suivant montre la réciproque :

Lemme 13 *Soit I une interprétation sur un univers de Herbrand $H(B)$. Alors il existe une interprétation de Herbrand I' sur $H(B)$ telle que pour toute formule Φ et pour toute affectation ρ, $[\![\Phi]\!]I\rho = [\![\Phi]\!]I'\rho$.*

Preuve : Soit I' l'ensemble des atomes clos A tels que $[\![A]\!]I\rho = \top$. Par définition, pour tout atome A — pas nécessairement clos — $[\![A]\!]I'\rho = [\![A\rho]\!]I'[]$ par le lemme 12, autrement dit $[\![A]\!]I'\rho = [\![A\rho]\!]I[]$ (par définition de I') $= [\![A]\!]I\rho$ (par une récurrence facile sur les sous-termes de A). On en déduit le lemme par récurrence structurelle sur Φ. Nous venons d'en prouver le cas de base; les autres cas sont immédiats. $\square$

À partir de maintenant, nous confondrons donc les deux notions d'interprétation sur $H(B)$ et d'interprétation de Herbrand.

Ce qui suit montre que rien de plus que la syntaxe (les univers et interprétations de Herbrand) n'est nécessaire pour décider de la validité :

Théorème 14 *Soit Φ une formule existentielle (une forme de Herbrand).*
Les assertions suivantes sont équivalentes :

(i) *Φ est valide;*

(ii) *pour tout univers de Herbrand $H(B)$, Φ est vraie dans toute interprétation de Herbrand sur B;*

(iii) *Φ est vraie dans toute interprétation de Herbrand sur B_0.*

Preuve : (i) implique (ii), qui implique (iii). Il ne reste qu'à montrer que (iii) implique (i). Comme Φ est existentielle, elle a la forme $\exists x_1 \cdot \ldots \exists x_n \cdot \Psi$, où Ψ est sans quantificateur. Par (iii), quelle que soit l'interprétation de Herbrand I, il existe des termes clos t_1, ..., t_n tels que $\Psi[t_1/x_1, \ldots, t_n/x_n]$ soit vraie dans I, autrement dit il existe une instance close $\Psi\sigma$ de Ψ qui soit vraie dans I.

Maintenant, soit I' une interprétation arbitraire (pas nécessairement sur $H(B_0)$) sur un domaine non vide arbitraire D. Cette interprétation induit une interprétation de Herbrand I sous forme de l'ensemble des atomes clos A tels que A soit vrai dans I'. Alors par (iii), il existe une instance close $\Psi\sigma$ de Ψ qui soit vraie dans I, donc dans I'. (L'argument est similaire à celui du lemme 13.) Ainsi l'affectation qui envoie chaque x_i vers l'interprétation du terme clos $x_i\sigma$ dans

D satisfait Ψ. Donc I' satisfait $\exists x_1 \cdot \ldots \exists x_n \cdot \Psi$, c'est-à-dire Φ. Comme I' est arbitraire, (i) est prouvée. $\square$

Ce théorème a comme conséquence le résultat important qui suit :

Théorème 15 (Herbrand, version sémantique) *Soit Φ une formule existentielle $\exists x_1 \cdot \ldots \exists x_n \cdot \Psi$, où Ψ est sans quantificateur. Supposons de plus qu'il y a au moins une constante dans le langage.*

Alors Φ est valide si et seulement s'il existe un entier k, et k instances closes $\Psi\sigma_1, \ldots, \Psi\sigma_k$ de Ψ telles que $\Psi\sigma_1 \vee \ldots \vee \Psi\sigma_k$ soit propositionnellement valide.

Preuve : Si $\Psi\sigma_1 \vee \ldots \vee \Psi\sigma_k$ est propositionnellement valide, alors Φ est clairement valide, puisque dans toute interprétation, une des formules $\Psi\sigma_i$, $1 \leq i \leq k$ doit être valide.

Réciproquement, si Φ est valide, par le théorème 14, pour toute interprétation de Herbrand I sur $H(B_0)$, il y a une instance close $\Psi\sigma_I$ satisfaite par I. (Remarquer que puisqu'il y a au moins une constante dans le langage, B_0 est vide, et $H(B_0)$ est exactement l'ensemble des termes clos.) En particulier, l'ensemble de tous les $\neg\Psi\sigma_I$ est propositionnellement insatisfiable, et nous affirmons qu'il existe alors un sous-ensemble fini insatisfiable $\neg\Psi\sigma_1, \ldots, \neg\Psi\sigma_k$ de l'ensemble des $\neg\Psi\sigma_I$: on en conclut que $\Psi\sigma_1 \vee \ldots \vee \Psi\sigma_k$ est propositionnellement valide.

Nous prouvons l'affirmation en montrant que, de tout ensemble Γ (propositionnellement) insatisfiable d'atomes clos, on peut extraire un sous-ensemble fini Δ qui est aussi insatisfiable. Construisons un arbre de décision binaire comme dans le cas propositionnel, à ceci près que ses nœuds sont étiquetés par les atomes clos, et comme il peut y en avoir un nombre infini, ses branches peuvent être infinies. Les branches d'un tel *arbre de Herbrand* sont juste les interprétations de Herbrand. Maintenant, sur chaque branche I, il existe un atome clos A_I dans Γ (autrement dit, un nœud sur I) qui est rendu faux par I, et nous pouvons définir A_I comme le plus haut de ces nœuds sur la branche. Élaguons chaque branche I juste en-dessous de A_I. Nous obtenons un arbre, appelé l'*arbre clos T*, dont les branches sont toutes finies. Par le lemme de König (une conséquence de l'axiome du choix, qui énonce que tout arbre dont tout nœud n'a qu'un nombre fini de fils, et dont toutes les branches sont finies, est fini), T est fini. Comme les A_I sont des nœuds dans T, l'ensemble Δ de tous les A_I est fini. De plus, il est insatisfiable par construction. (Le résultat est en fait juste même quand Γ est n'importe quel ensemble de formules propositionnelles, et pas seulement des atomes : c'est le théorème de *compacité* de la logique propositionnelle classique. Exercice : prouvez-le, en remarquant que toute formule propositionnelle ne contient qu'un nombre fini d'atomes.) $\square$

Le thèorème de Herbrand est important parce qu'il fournit une procédure pour tester si une formule donnée est valide : herbrandiser la formule en une formule existentielle Φ, puis énumérer toutes les suites d'instances closes $\Psi\sigma_1, \ldots, \Psi\sigma_k$ et tester si leur disjonction est propositionnellement valide. Ceci peut ne pas terminer, car nous ne connaissons pas de borne a priori sur k, mais au moins si la formule est valide, nous finirons par le découvrir par ce moyen. (Nous disons que ce processus est *complet*. Il est clairement correct, autrement dit il ne peut déclarer valide aucune formule invalide.) Cette méthode a été explorée jusqu'au début des années 1960, où des façons plus efficaces d'utiliser le théorème de Herbrand ont été conçues. Ces méthodes formeront le sujet des chapitres à venir.

Le théorème de Herbrand est en général présenté sous sa forme duale :

Théorème 16 (Herbrand, dual) *Soit Φ une formule universelle $\forall x_1 \ldots \forall x_n \cdot \Psi$, où Ψ est sans quantificateur. Supposons de plus qu'il y ait au moins une constante dans le langage.*

Alors Φ est insatisfiable si et seulement s'il existe un entier k, et k instances closes $\Psi\sigma_1, \ldots, \Psi\sigma_k$ de Ψ telles que $\Psi\sigma_1 \wedge \ldots \wedge \Psi\sigma_k$ soit propositionnellement insatisfiable.

Une composante mystérieuse de ce théorème est qu'il a besoin d'un entier k, alors qu'on imaginerait volontiers $k = 1$. En d'autres termes, nous pourrions nous attendre à ce que $\exists x \cdot \Psi$ soit valide si et seulement s'il existait un terme t tel que $\Psi[t/x]$ soit valide. Un tel terme t s'il existe est appelé un *témoin* de la formule existentielle $\exists x \cdot \Psi$, et est une représentation sous forme de terme d'une valeur v telle que $\Psi[v/x]$ est vraie. Ceci est le cas en logique intuitionniste, mais en logique classique toutes les valeurs ne sont pas représentables par des termes. Par exemple, $\exists x \cdot P(a) \wedge P(b) \Rightarrow P(x)$ est valide en classique, et une valeur v pour x serait a si $P(a)$ est vraie, et b sinon, autrement dit nous pourrions choisir "si $P(a)$ alors a sinon b" comme témoin ... mais ce n'est pas un terme du premier ordre ! Nous venons de comprendre un peu mieux ce que dit le théorème de Herbrand : $\exists x \cdot \Psi$ est valide si et seulement si nous pouvons trouver des témoins pour x sous la forme de programmes si-alors-sinon finis qui retournent des termes.

Le théorème de Herbrand entraîne immédiatement :

Théorème 17 (Énumérabilité récursive) *L'ensemble des énoncés du premier ordre valides est récursivement énumérable. Autrement dit, il existe une procédure qui énumère tous les énoncés valides du premier ordre, et eux seulement.*

Toutes les propriétés récursives P de formules (autrement dit, les propriétés décidables, celles qui peuvent être vérifiées par un algorithme qui termine) sont récursivement énumérables : énumérer toutes les formules, et vérifier si P est vraie pour chaque. Cependant, il existe des propriétés récursivement énumérables qui ne sont pas récursives, et malheureusement la validité au premier ordre n'est pas décidable :

Théorème 18 (Indécidabilité) *Il n'existe pas d'algorithme terminant qui, étant donnée une formule du premier ordre Φ en entrée, retourne vrai si et seulement si Φ est classiquement valide.*

Preuve : Comme nous n'avons pas donné de définition précise de *calculable* ou *récursif* (cf. [Joh92, DW85]), et comme ce n'est pas notre propos, nous montrons ce théorème par réduction à partir d'un problème indécidable plus commode, le *problème de correspondance de Post* (du nom du logicien Emil Post).

Soit A un alphabet fini, contenant au moins deux lettres distinctes. A^* est l'ensemble de tous les mots (y compris le mot vide) sur A, c'est-à-dire de toutes les suites finies de lettres. Soit E une collection finie de couples (u_i, v_i), $1 \leq i \leq n$, que nous noterons $u_i \doteq v_i$. Étant donnée une suite finie i_1, i_2, ..., i_k d'indices entre 1 et n, nous pouvons former la concaténation $u_{i_1} u_{i_2} \ldots u_{i_k}$, et aussi $v_{i_1} v_{i_2} \ldots v_{i_k}$. Le problème de correspondance de Post est : étant donné E, y a-t-il une suite non vide comme ci-dessus telle que $u_{i_1} u_{i_2} \ldots u_{i_k} = v_{i_1} v_{i_2} \ldots v_{i_k}$?

Se reporter à la figure 3 pour un exemple : nous nous y donnons quatre équations (à gauche de la figure), et avons indiqué une solution au problème de Post sur la droite.

Le problème de Post est indécidable; voir [DW85]. Nous pouvons de plus supposer que E ne contient pas l'équation (ϵ, ϵ), où ϵ est le mot vide, et que l'alphabet ne contient que deux lettres distinctes a et b.

Pour coder le problème de Post en logique du premier ordre, nous créons deux symboles de fonction unaires pour représenter les lettres, et les nommons encore a et b; nous créons aussi une constante ϵ représentant le mot vide, et un symbole de prédicat binaire P représentant la convertibilité par les règles de Post. Nous codons les mots w, disons $a_{j_1} a_{j_2} \ldots a_{j_m}$, comme des termes de la forme $a_{j_m}(\ldots a_{j_2}(a_{j_1}(x)) \ldots)$, où a_{j_1}, ..., a_{j_m} sont des lettres et x est une variable; soit $w(x)$ ce dernier terme. Nous écrivons aussi w au lieu de $w(\epsilon)$. Nous pouvons alors axiomatiser le problème de Post comme suit :

- Axiomes de E: $\forall x, y \cdot P(x, y) \Rightarrow P(u_i(x), v_i(x))$, pour tout $1 \leq i \leq n$;
- Condition de succès : $P(\epsilon, \epsilon)$.

Soit maintenant C la conjonction de ces formules, et soit Φ la formule $C \Rightarrow \exists x \cdot P(a(x), a(x)) \vee P(b(x), b(x))$. Intuitivement, C dit que $P(w, w')$ est vraie soit quand $w = w' = \epsilon$ soit quand $w = w_1 u_i$ et $w' = w'_1 v_i$ pour un certain i et des termes w_1, w'_1 tels que $P(w_1, w'_1)$. La formule Φ dit que l'on peut fabriquer un mot non vide (autrement dit un mot de la forme wa ou wb pour un certain mot w éventuellement vide) à partir de ces règles. Comme E ne contient pas (ϵ, ϵ), ceci est équivalent au problème de Post.

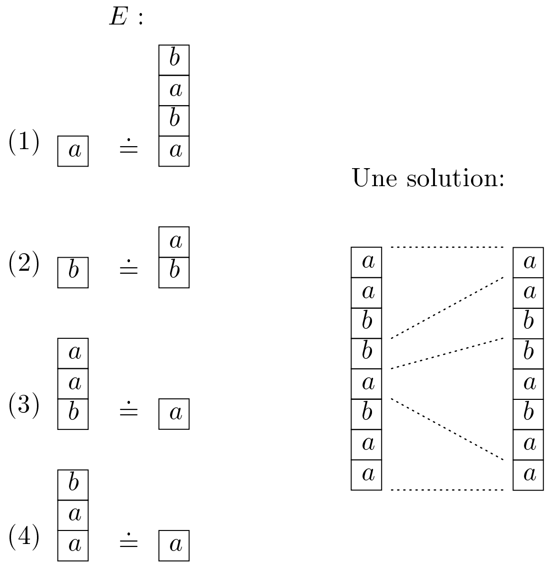

Figure 3: Un exemple du problème de correspondance de Post

Par exemple (cf. figure 3), nous avons $P(aababbaa, aababbaa)$ parce que $P(\epsilon, \epsilon)$, donc $P(aab, a)$ par l'équation (4), donc $P(aaba, aabab)$ par l'équation (1), donc $P(aabab, aababba)$ par l'équation (2), donc $P(aababbaa, aababbaa)$ par l'équation (3).

Nous montrons maintenant que Φ est valide si et seulement si le problème de Post a une solution. D'abord, si l'on met Φ en forme prénexe, on obtient une formule existentielle. Donc si Φ est valide, le théorème de Herbrand s'applique, et donc nous pouvons prouver Φ en démontrant que la disjonction de $\neg P(\epsilon, \epsilon)$, d'un nombre fini d'instances closes d'axiomes niés :

$$P(w_j, w'_j) \wedge \neg P(u_{i_j}(w_j), v_{i_j}(w'_j))$$

avec $1 \leq j \leq m$ et pour tout j, $1 \leq i_j \leq n$; et d'un nombre fini d'instances de la conclusion :

$$\begin{array}{l} P(a(v_k), a(v_k)) \\ P(b(v_k), b(v_k)) \end{array}$$

où $1 \leq k \leq p$, est propositionnellement valide. De façon équivalent, l'ensemble de clauses :

$$\begin{array}{l} P(\epsilon, \epsilon) \\ \neg P(w_j, w'_j) \vee P(u_{i_j}(w_j), v_{i_j}(w'_j)) \\ \neg P(a(v_k), a(v_k)) \\ \neg P(b(v_k), b(v_k)) \end{array}$$

où $1 \leq j \leq m$ et $1 \leq k \leq p$, est propositionnellement insatisfiable. Par la complétude de la résolution propositionnelle, il existe une dérivation de la clause vide à partir des clauses ci-dessus. Nous disons que (w, w') est une *paire de Post* si et seulement si $w = u_{j_0} u_{j_1} \ldots u_{j_l}$ et $w' = v_{j_0} v_{j_1} \ldots v_{j_l}$ pour un certain $l \geq 0$ et des indices $j_0, \ldots, j_l$ entre 1 et n. Une récurrence aisée sur la longueur de la dérivation par résolution montre que les seules clauses non vides que nous pouvons dériver sont :

- de la forme $\neg P(w_{j_0}, w'_{j_0}) \vee P(w(w_{j_0}), w'(w'_{j_0}))$, où (w, w') est une paire de Post;
- ou de la forme $\neg P(w_{j_0}, w'_{j_0})$, où il y a une paire de Post (w, w') telle que $w(w_{j_0}) = w'(w'_{j_0})$ n'est pas le mot vide;
- ou de la forme $P(w(\epsilon), w'(\epsilon))$ avec (w, w') une paire de Post.

On ne peut alors dériver la clause vide qu'en résolvant sur des clauses de l'une des deux dernières formes : ceci implique qu'il y a deux paires de Post (w_1, w_1') et (w_2, w_2') telles que $w_2(w_1(\epsilon)) = w_2'(w_1'(\epsilon))$, autrement dit telles que $w_1 w_2 = w_1' w_2'$, et que ce dernier mot n'est pas vide. Donc si Φ est valide, le problème de Post a une solution.

Réciproquement, il est clair que si le problème de Post a une solution, la formule Φ est valide.

Nous en concluons que l'espèce de problèmes du premier ordre qui code les problèmes de Post est indécidable, donc la validité en logique du premier ordre en général est indécidable. (Bien sûr, ceci est vrai de la satisfiabilité également.) □

Ceci n'a pas découragé les chercheurs d'essayer de trouver des procédures efficaces de recherche de preuve. Le théorème 18 dit que tout procédure de recherche de preuve correcte et complète doit ne pas terminer sur au moins une formule en entrée; nous pouvons quand même utiliser une telle procédure, et l'interrompre quand elle prend trop de temps, auquel cas nous ne savons pas si la formule d'entrée était valide ou non. (Plus finement, cet échec est une indication que la formule d'entrée, si elle est valide, est d'autant plus difficile à prouver qu'elle résiste plus longtemps.) Un autre point de vue est de dire que toute procédure correcte et qui termine doit être incapable de trouver des preuves de certaines propositions valides. Tous ces compromis sont acceptables, aussi longtemps que nous nous efforçons de décider la classe la plus grande de propositions en temps raisonnable. Il est clair que la procédure d'énumération décrite dans la preuve du théorème 17 est l'une des plus inefficaces que l'on puisse concevoir, et nous en examinerons d'autres dans les chapitres suivants.

D'un point de vue logique, ou méta-mathématique, les constructions de Herbrand fournissent aussi les résultats suivants :

Théorème 19 (Compacité) *Si $\Gamma \models \Phi$, alors il existe un sous-ensemble fini Δ de Γ tel que $\Delta \models \Phi$.*

Preuve : Il suffit de montrer que si Γ n'a pas de modèle, alors il existe un sous-ensemble fini Δ de Γ qui n'en a pas non plus. Supposons que toutes les formules de Γ ont été skolémisées, sans perte de généralité. Si Γ n'a pas de modèle, il n'a pas de modèle de Herbrand sur B_0, ce qui signifie que l'ensemble des instances closes des formules sans quantificateur Ψ, où $\forall x_1 \cdot \ldots \forall x_n \cdot \Psi \in \Gamma$, n'a pas de modèle, par le théorème 14, point (*iii*). Ce dernier ensemble est un ensemble propositionnellement insatisfiable de formules propositionnelles : par le théorème de compacité propositionnelle (voir la preuve du théorème 15), il existe un sous-ensemble fini insatisfiable E de ces formules. Nous pouvons alors choisir un sous-ensemble fini Δ de Γ tel que pour tout Ψ' dans E, Ψ' est une instance d'une formule Ψ telle que $\forall x_1 \cdot \ldots \forall x_n \cdot \Psi \in \Delta$. □

Ce qui a pour conséquence :

Théorème 20 (Löwenheim-Skolem, faible) *Si Φ a un modèle, alors il a des modèles de n'importe quel cardinal infini.*

Preuve : Si Φ a un modèle, alors elle a un modèle de Herbrand sur $H(B)$ pour tout B, par le théorème 14. Si tous les symboles de fonction dans $\mathcal{F}$ sont d'arité 0, et si nous choisissons B de cardinal $\alpha \geq \aleph_0$, le cardinal de $H(B)$ est celui de $\mathcal{F}$ plus α, autrement dit α puisque le cardinal de $\mathcal{F}$ est inférieur ou égal à $\aleph_0$. Sinon, s'il existe un symbole de fonction non nullaire dans $\mathcal{F}$, le cardinal de $H(B)$ est $\alpha.\aleph_0$, c'est-à-dire α puisque α est supérieur ou égal à $\aleph_0$. Comme α est un cardinal infini arbitraire, le théorème est prouvé. □

Mais nous avons triché : dans le modèle ci-dessus, nous avons introduit de nombreuses valeurs qui peuvent être en fait prouvablement égales (si nous avons un prédicat d'égalité). En fait, nous pouvons prouver le résultat surprenant suivant :

Théorème 21 (Löwenheim-Skolem) *Soit $\doteq$ un symbol de prédicat binaire. Nous disons d'un modèle qu'il est* équationnel *si et seulement si $\doteq$ est interprété comme la relation d'égalité dans ce modèle.*

Alors, si Φ a un modèle équationnel infini, il a des modèles équationnels de n'importe quel cardinal infini.

Preuve : Supposons que Φ a un modèle équationnel infini. Autrement dit, il existe une interprétation I sur un domaine D qui satisfait Φ et tous les axiomes de l'égalité $\doteq$, telle que le quotient de D par la relation d'équivalence $I(\doteq)$ est infini.

Soit β le cardinal de ce quotient, et soit α n'importe quel cardinal infini tel que $\alpha \leq \beta$. Créons α nouveaux symboles de contante c_i, $0 \leq i < \alpha$ et étendons l'interprétation I de sorte que chaque $I(c_i)$ soit dans une classe d'équivalence distincte pour $I(\doteq)$. Nous pouvons le faire, parce qu'il y a $\beta \geq \alpha$ classes d'équivalence. Soit S l'ensemble des formules contenant Φ et toutes les inéquations $\neg c_i \doteq c_j$, $i \neq j$ (vu comme une conjonction); S admet clairement I comme modèle équationnel. Mais alors I induit une interprétation de Herbrand I' sur $H(B)$, où $B = B_0 \cup \{c_i \mid 0 \leq i < \alpha\}$, telle que I' satisfait S. En particulier, I' satisfait Φ, et comme I' satisfait toutes les $\neg c_i \doteq c_j$, $i \neq j$, son quotient par $I'(\doteq)$ a un cardinal d'au moins α. Ensuite, B est de cardinal α, donc $H(B)$ est de cardinal au plus $\alpha.\aleph_0$, c'est-à-dire exactement α. En particulier, $H(B)/I'(\doteq)$ est un modèle équationnel de Φ de cardinal α. Ceci démontre ce que nous appelons la version *descendante* du théorème, à savoir que si Φ a un modèle équationnel infini, alors il a des modèles équationnels infinis de tout cardinal plus petit.

Pour montrer la version *ascendante*, nous construisons un modèle équationnel de cardinal $\alpha > \beta$ en en construisant d'abord un de cardinal β^α, puis en utilisant la version descendante pour redescendre à α. En effet, comme $\beta \geq 2$, nous avons $\alpha \leq \beta^\alpha$. Définissons D^α comme le produit de α copies distinctes de D (son cardinal est β^α), et définissons une interprétation I^α sur D^α par $I^\alpha(f)(v_1, \ldots, v_n) = (I(f)(v_{1i}, \ldots, v_{ni}))_{0 \leq i < \alpha}$ (application point à point) et en posant $I^\alpha(P)(v_1, \ldots, v_n)$ égal à la conjonction de tous les booléens $I(P)(v_{1i}, \ldots, v_{ni})$, $0 \leq i < \alpha$. Remarquons que $I^\alpha(\doteq)$ est l'égalité sur D^α.

Pour toute affectation ρ de variables vers D^α, et pour tout $0 \leq i < \alpha$, soit ρ_i l'affectation envoyant x vers la $i$ième composante de $\rho(x)$. Par une récurrence facile sur les formules, pour toute formule Ψ, nous avons $\bigwedge_{0 \leq i < \alpha} [\![\Psi]\!] I \rho_i \leq [\![\Psi]\!] I^\alpha \rho \leq \bigvee_{0 \leq i < \alpha} [\![\Psi]\!] I \rho_i$, où $\bigwedge$ est la conjonction distribuée, $\bigvee$ est la disjonction distribuée et $\leq$ est l'ordre défini sur les booléens par $\bot \leq \top$, $\bot \neq \top$. Mais lorsque $\Psi = \Phi$, la conjonction et la disjonction valent toutes les deux $\top$, donc $[\![\Phi]\!] I^\alpha \rho = \top$, autrement dit I^α est bien un modèle (équationnel) de Φ.

(Une preuve plus simple mais moins constructive de la version ascendante est : étant donné un cardinal α, ajoutons α nouvelles constantes c_i, $0 \leq i < \alpha$, et considérons l'ensemble des formules $S = \{\Phi\} \cup \{\neg c_i \doteq c_j \mid i \neq j\}$. Si S n'avait pas de modèle équationnel, par compacité il y aurait un sous-ensemble incohérent fini de S; mais tout sous-ensemble fini de S peut être interprété en envoyant chaque c_i du sous-ensemble fini vers des éléments distincts du modèle infini I. Donc S a un modèle équationnel, qui est de cardinal au moins α, et nous utilisons la version descendante pour en obtenir un de cardinal exactement α.) $\square$

En particulier, nous ne pouvons pas espérer caractériser les entiers ou les réels en logique du premier ordre, même avec une infinité d'axiomes (par compacité, ceci ce ramène au cas d'un nombre fini d'axiomes, donc au cas d'une formule). Ceci est dû au fait que, quelle que soit l'axiomatisation au premier ordre que nous choisissions, il existera des modèles non dénombrables de l'arithmétique. Le problème est analogue avec les réels : toute axiomatisation au premier ordre des réels doit avoir des modèles dénombrables, bien que les réels ne soient pas dénombrables. (La théorie des corps réels clos, aussi appelée théorie du premier ordre des réels, a en particulier comme plus petit modèle l'ensemble des nombres algébriques.)

5.3 Élimination des coupures et théorie de Herbrand syntaxique

Tout ce que nous avons effectué dans les dernières sections peut aussi être fait dans un cadre purement syntaxique, en remplaçant les notions de validité et de satisfiabilité par les notions leur correspondant en théorie de la preuve, à savoir la prouvabilité et la cohérence. C'est plus dur à faire que dans l'approche sémantique, mais nous gagnerons des intuitions supplémentaires sur ce qui se passe réellement.

L'outil de base est le théorème d'élimination des coupures (théorème 24), c'est-à-dire le fait que nous puissions transformer toute preuve de **LK** en une preuve sans coupure (Cut). Pour le démontrer, nous avons d'abord besoin de quelques définitions et quelques lemmes. Nous considérons que les séquents sont des ensembles de formules, de sorte que les contractions, les affaiblissements et les échanges seront traités implicitement; le lecteur est prié de vérifier que ceci n'invalide pas nos arguments.

Définition 22 *Dans une règle de* **LK** *:*

$$\frac{\Gamma_i \longrightarrow \Delta_i, 1 \leq i \leq n}{\Gamma \longrightarrow \Delta}$$

les séquents $\Gamma_i \longrightarrow \Delta_i$ sont appelés les prémisses, *et $\Gamma \longrightarrow \Delta$ est la* conclusion *de la règle.*

Dans une coupure (règle Cut) :

$$\frac{\Gamma \longrightarrow \Phi, \Delta \qquad \Gamma', \Phi \longrightarrow \Delta'}{\Gamma, \Gamma' \longrightarrow \Delta, \Delta'}$$

Φ est appelée la formule de coupure.

Dans toutes les règles, les formules actives *(cf. figure 2) dans les prémisses sont Φ, Φ' dans les cas $\wedge L$, $\wedge R$, $\vee L$, $\vee R$, $\Rightarrow L$, $\Rightarrow R$; Φ dans les cas $\neg L$ et $\neg R$; $\Phi[t/x]$ dans les cas $\forall L$ et $\exists R$; $\Phi[y/x]$ dans les cas $\forall R$ et $\exists L$; et Φ (la formule de coupure) dans le cas de Cut.*

Dans toutes les règles, la formule principale *dans la conclusion est Φ dans le cas de Ax, $\Phi \wedge \Phi'$ dans les cas $\wedge L$ et $\wedge R$; $\Phi \vee \Phi'$ dans les cas $\vee L$ et $\vee R$; $\Phi \Rightarrow \Phi'$ dans les cas $\Rightarrow L$ et $\Rightarrow R$; $\neg\Phi$ dans les cas $\neg L$ et $\neg R$; $\forall x \cdot \Phi$ dans les cas $\forall L$ et $\forall R$; $\exists x \cdot \Phi$ dans les cas $\exists L$ et $\exists R$.*

Soit Φ une formule. La profondeur *$d(\Phi)$ de Φ est définie par : $d(A) = 0$ si A est un atome, $d(\Phi \vee \Phi') = 1 + \max(d(\Phi), d(\Phi'))$ (et de même pour $\wedge$, $\neg$, $\Rightarrow$), $d(\forall x \cdot \Phi') = 1 + d(\Phi')$ (et de même pour $\exists$).*

Soit π une preuve dans **LK***, que nous voyons comme un arbre inversé de séquents.*

La profondeur *$d(\pi)$ de π est la profondeur de l'arbre qui le représente, autrement dit la profondeur d'un axiome Ax est 0, et la profondeur d'une preuve se terminant sur une règle de prémisses les conclusions de preuves π_i, $1 \leq i \leq n$, est $1 + \max(d(\pi_1), \ldots, d(\pi_n))$.*

Une coupure est dite maximale *s'il s'agit d'une coupure entre deux preuves sans coupure. Le* rang de coupure *$r(\pi)$ de π est défini comme étant 0 si π est sans coupure, sinon comme le maximum de $d(\pi')+d(\Phi)$, où π' parcourt toutes les sous-preuves de π se terminant sur une coupure maximale, de formule de coupure Φ.*

Le rang de coupure est une mesure de la complexité de la preuve π. Le lemme qui suit montre que nous pouvons toujours faire décroître cette complexité, le prix en étant une éventuelle augmentation de la taille de la preuve :

Lemme 23 (Élimination des coupures, une étape) *Soit π une preuve dans* **LK** *de rang de coupure r non nul. Alors il existe une preuve dans* **LK** *de rang de coupure au plus $r-1$ du même séquent.*

Preuve : Soit n le nombre de sous-preuves sans coupure π' de π qui se terminent sur une coupure (maximale), de formule de coupure Φ, telles que $d(\pi') + d(\Phi) = r$, autrement dit de rang maximal. Par hypothèse, $n \neq 0$. π' est de la forme :

$$\dfrac{\begin{array}{c}\pi_1 \\ \vdots \\ \Gamma \longrightarrow \Delta, \Phi\end{array} \qquad \begin{array}{c}\pi_2 \\ \vdots \\ \Gamma', \Phi \longrightarrow \Delta'\end{array}}{\Gamma, \Gamma' \longrightarrow \Delta, \Delta'}\ \mathrm{Cut}$$

où π_1 et π_2 sont sans coupure. L'idée est de pousser les coupures vers le haut le long de π_1 et π_2, jusqu'à ce que la coupure ci-dessus de rang maximal soit éliminée, ne laissant ainsi que des coupures de rang inférieur. Plus formellement, nous montrons que nous pouvons transformer la preuve π en une preuve du même séquent, avec au plus $n-1$ sous-preuves de rang r et aucune de rang supérieur.

Pour ce faire, nous montrons que nous pouvons transformer π' en une preuve de rang au plus $r-1$. L'argument est par récurrence sur $d(\pi_1)+d(\pi_2)$. Il y a deux cas :

- si Φ n'est pas la formule principale dans la dernière règle de π_1 (ou symétriquement, si Φ n'est pas la formule principale dans la dernière règle de π_2), alors c'est une règle R de la forme :

  $$\frac{\Gamma_i \longrightarrow \Delta_i, \Phi \quad (1 \leq i \leq k)}{\Gamma \longrightarrow \Delta, \Phi}$$

  alors nous transformons π' en :

  $$\dfrac{\dfrac{\begin{matrix}\pi'_1\\ \vdots\\ \Gamma_1 \longrightarrow \Delta_1, \Phi\end{matrix} \qquad \begin{matrix}\pi_2\\ \vdots\\ \Gamma', \Phi \longrightarrow \Delta'\end{matrix}}{\Gamma_1, \Gamma' \longrightarrow \Delta_1, \Delta'}\,\mathrm{Cut} \qquad \cdots \qquad \dfrac{\begin{matrix}\pi'_k\\ \vdots\\ \Gamma_k \longrightarrow \Delta_k, \Phi\end{matrix} \qquad \begin{matrix}\pi_2\\ \vdots\\ \Gamma', \Phi \longrightarrow \Delta'\end{matrix}}{\Gamma_k, \Gamma' \longrightarrow \Delta_k, \Delta'}\,\mathrm{Cut}}{\Gamma, \Gamma' \longrightarrow \Delta, \Delta'}\,R$$

  où $\pi'_1, \ldots, \pi'_k$ sont les sous-preuves dont les conclusions sont $\Gamma_i \longrightarrow \Delta_i, \Phi$, $1 \leq i \leq k$, et où la dernière règle est bien une instance de R, comme on peut le vérifier pour chaque règle. (Souvenez-vous que Φ n'était pas principale.)

  Par hypothèse de récurrence, (en effet, $d(\pi'_i)+d(\pi_2)$ est strictement inférieur à $d(\pi_1)+d(\pi_2)$ pour tout i) les preuves qui se terminent en les k coupures ci-dessus peuvent être transformées en des preuves de rang au plus $r-1$. Ceci fournit une preuve totale de rang au plus $r-1$.

  Remarquons que si π' était une instance de Ax, alors $k=0$ et la coupure disparaît purement et simplement, ne laissant que l'instance originale de Ax.

- si Φ est la formule principale dans les règles finales de π_1 et π_2 à la fois, nous avons encore deux cas :

  - si π_1 (symétriquement, π_2) est une instance de Ax, alors π' est de la forme :

    $$\dfrac{\dfrac{}{\Gamma, \Phi \longrightarrow \Delta, \Phi}\,\mathrm{Ax} \qquad \begin{matrix}\pi_2\\ \vdots\\ \Gamma', \Phi \longrightarrow \Delta'\end{matrix}}{\Gamma, \Gamma', \Phi \longrightarrow \Delta, \Delta'}\,\mathrm{Cut}$$

    mais nous pouvons prouver le même séquent en *affaiblissant* la preuve π_2 pour en faire une nouvelle preuve π'_2, en ajoutant systématiquement Γ à gauche et Δ à droite de chaque séquent qui apparaît dans π_2. (C'est une récurrence structurelle triviale sur π_2.) π'_2 n'a pas de coupure de rang supérieur ou égal à r : en fait, comme π_2 est sans coupure, π'_2 aussi.

– sinon, les deux preuves se terminent sur des règles dont la formule principale est la formule de coupure Φ, et ces règles ne sont ni Ax ni Cut. Maintenant, si Φ est une conjonction, ces règles finales doivent être $\wedge$L et $\wedge$R; si Φ est une disjonction, elles doivent être $\vee$L et $\vee$R; et ainsi de suite, de même pour la négation, l'implication, les quantifications universelle et existentielle.

Nous traitons de la conjonction, les autres cas propositionnels sont analogues ou plus simples. Alors π' est de la forme :

$$\dfrac{\dfrac{\begin{array}{c}\pi'_1\\ \vdots\\ \Gamma,\Phi',\Phi''\longrightarrow\Delta\end{array}}{\Gamma,\Phi'\wedge\Phi''\longrightarrow\Delta}\wedge\mathrm{L} \qquad \dfrac{\begin{array}{c}\pi'_2\\ \vdots\\ \Gamma'\longrightarrow\Delta',\Phi'\end{array}\qquad\begin{array}{c}\pi''_2\\ \vdots\\ \Gamma'\longrightarrow\Delta',\Phi''\end{array}}{\Gamma'\longrightarrow\Delta',\Phi'\wedge\Phi''}\wedge\mathrm{R}}{\Gamma,\Gamma'\longrightarrow\Delta,\Delta'}\mathrm{Cut}$$

où $\Phi = \Phi' \wedge \Phi''$. Nous la transformons en :

$$\dfrac{\dfrac{\begin{array}{c}\pi'_1\\ \vdots\\ \Gamma,\Phi',\Phi''\longrightarrow\Delta\end{array}\qquad\begin{array}{c}\pi'_2\\ \vdots\\ \Gamma'\longrightarrow\Delta',\Phi'\end{array}}{\Gamma,\Gamma',\Phi''\longrightarrow\Delta,\Delta'}\mathrm{Cut}\qquad\begin{array}{c}\pi'_3\\ \vdots\\ \Gamma'\longrightarrow\Delta',\Phi''\end{array}}{\Gamma,\Gamma'\longrightarrow\Delta,\Delta'}\mathrm{Cut}$$

(Remarquer que nous aurions pu choisir de couper sur Φ'' d'abord et ensuite sur Φ', au lieu de comme ci-dessus; ceci n'a aucune importance ici.) Les deux coupures ci-dessus sont de rangs strictement plus petits que r. En effet, la coupure du dessus est de rang $\max(d(\pi'_1)+1, d(\pi'_2)+1) + d(\Phi') < d(\pi') + d(\Phi') < d(\pi') + d(\Phi) = r$; et celle du bas est de rang $\max(d(\pi'_1)+2, d(\pi'_2)+2, d(\pi'_3)+1) + d(\Phi'') \leq d(\pi') + d(\Phi'') < d(\pi') + d(\Phi) = r$.

Examinons maintenant le cas des quantifications; nous examinons le cas des quantifications universelles, car le cas existentiel est analogue. π est de la forme :

$$\dfrac{\dfrac{\begin{array}{c}\pi'_1\\ \vdots\\ \Gamma,\Phi[t/x]\longrightarrow\Delta\end{array}}{\Gamma,\forall x\cdot\Phi\longrightarrow\Delta}\forall\mathrm{L}\qquad\dfrac{\begin{array}{c}\pi'_2\\ \vdots\\ \Gamma'\longrightarrow\Phi[y/x],\Delta'\end{array}}{\Gamma'\longrightarrow\forall x\cdot\Phi}\forall\mathrm{R}}{\Gamma,\Gamma'\longrightarrow\Delta,\Delta'}\mathrm{Cut}$$

où y n'est pas libre dans Γ', Δ'.

Soit π''_2 la preuve obtenue en remplaçant systématiquement y par t dans tout séquent de π'_2. Par une récurrence structurelle facile, et comme π'_2 est sans coupure, π''_2 est une preuve valide. (La seule difficulté serait que le remplacement de y par t pourrait briser les conditions de bord des règles $\forall$R et $\exists$L dans π'_2, mais ceci ne peut pas se produire lorsque π'_2 ne contient pas Cut.) Alors π''_2 prouve $\Gamma'[t/x] \longrightarrow \Phi[t/x], \Delta'[t/x]$, autrement dit $\Gamma' \longrightarrow \Phi[t/x], \Delta'$. Nous obtenons donc la preuve :

$$\dfrac{\begin{array}{c}\pi'_1\\ \vdots\\ \Gamma,\Phi[t/x]\longrightarrow\Delta\end{array}\qquad\begin{array}{c}\pi''_2\\ \vdots\\ \Gamma'\longrightarrow\Phi[t/x],\Delta'\end{array}}{\Gamma,\Gamma'\longrightarrow\Delta,\Delta'}\mathrm{Cut}$$

qui est moins profonde d'un niveau, de sort que la coupure restante est de rang au plus $r-1$.

Nous en concluons que le lemme est valide, par récurrence sur le nombre n de coupures maximales de rang maximal. □

Théorème 24 (Élimination des coupures) *Un séquent $\Gamma \longrightarrow \Delta$ est prouvable dans* **LK** *si et seulement s'il est prouvable dans* **LK** *sans la règle de coupure Cut.*

Preuve : Par récurrence sur le rang de coupure d'une preuve de $\Gamma \longrightarrow \Delta$, en utilisant le lemme 23. □

La preuve dit même plus : de toute preuve de $\Gamma \longrightarrow \Delta$, nous pouvons extraire *par calcul* une preuve sans coupure du même séquent.

Les propriétés suivantes sont des conséquences immédiates :

Définition 25 (Sous-formules) *Soit Φ une formule de la logique du premier ordre. L'ensemble des* sous-formules *de Φ est le plus petit ensemble de formules tel que :*

- *Φ est une sous-formule de Φ;*
- *si $\Phi' \wedge \Phi''$ est une sous-formule de Φ, alors Φ' et Φ'' sont des sous-formules de Φ (de même pour $\vee$, $\neg$, $\Rightarrow$).*
- *si $\forall x \cdot \Phi'$ est une sous-formule de Φ, alors toutes les formules de la forme $\Phi'[t/x]$, où t est n'importe quel terme, sont des sous-formules de Φ. (De même pour les quantifications existentielles.)*

L'ensemble des sous-formules est en général infini, et beaucoup plus gros que l'ensemble des nœuds dans le graphe qui représente Φ : c'est parce que toute instance de Φ' est considérée comme une sous-formule de $\forall x \cdot \Phi'$ ou de $\exists x \cdot \Phi'$. L'idée est que nous pouvons reconnaître les sous-formules d'une formule donnée, par filtrage ("pattern-matching" en anglais).

Corollaire 26 (Propriété de la sous-formule) *Dans une preuve sans coupure du séquent $\Gamma \longrightarrow \Delta$, tous les séquents consistent uniquement en des sous-formules de Γ ou de Δ.*

Preuve : Récurrence structurelle immédiate sur la preuve. □

Corollaire 27 *$\longrightarrow \exists x \cdot \Phi$ est prouvable en* **LK** *si et seulement s'il existe un nombre fini k de termes $t_1, \ldots, t_k$ tels que $\longrightarrow \Phi[t_1/x], \ldots, \Phi[t_k/x]$ soit prouvable en* **LK**.

Preuve : Soit π une preuve sans coupure de $\longrightarrow \exists x \cdot \Phi$. Soient $t_1, \ldots, t_k$ les termes tels que $\Phi[t_i/x]$ apparaît dans π. Il ne peut y avoir qu'un nombre fini de telles formules, car la preuve est finie.

Nous transformons la preuve π en une preuve de $\longrightarrow \Phi[t_1/x], \ldots, \Phi[t_k/x]$ comme suit : pour chaque séquent dans la preuve, effacer toute occurrence de $\exists x \cdot \Phi$ du côté droit des séquents, et ajouter toutes les formules $\Phi[t_1/x], \ldots, \Phi[t_k/x]$ à droite. Remarquer que $\exists x \cdot \Phi$ ne peut apparaître qu'à droite des séquents. En particulier, cette transformation préserve les axiomes Ax. Elle laisse aussi invariante toutes les autres règles, autres que Cut et les règles sur les quantificateurs. Mais la seule règle sur les quantificateurs qui pourrait poser un problème est $\exists$R, que nous remplaçons simplement par un affaiblissement, puisque la traduction de la prémisse est incluse dans celle de la conclusion. Nous obtenons finalement une preuve de $\longrightarrow \Phi[t_1/x], \ldots, \Phi[t_k/x]$. □

Corollaire 28 (Herbrand, version syntaxique) *Soit Φ une formule existentielle $\exists x_1 \cdot \ldots \exists x_n \cdot \Psi$, où Ψ est sans quantificateur.*

Si $\longrightarrow \Phi$ est prouvable dans **LK**, *alors il existe un nombre fini k, et k instances $\Psi\sigma_1, \ldots, \Psi\sigma_k$ de Ψ telles que $\longrightarrow \Psi\sigma_1, \ldots, \Psi\sigma_k$ soit prouvable dans* **LK**.

Preuve : Récurrence immédiate sur n, en utilisant le corollaire 27. □

Une autre application directe de l'élimination des coupures est :

Corollaire 29 (Cohérence) *Le système* **LK** *est cohérent, autrement dit il n'y a pas de preuve du séquent* $\longrightarrow$ *dans* **LK**.

Preuve : S'il y en avait une, il y en aurait une sans coupure. Mais l'ensemble des sous-formules de $\longrightarrow$ est vide, donc il n'en existe aucune preuve sans coupure. □

Notre prochain but est de montrer que **LK** est correct est complet par rapport à la sémantique de la logique classique du premier ordre. La version sémantique du théorème de Herbrand entraîne que ceci est vrai pour les formules existentielles, et nous savons aussi comment traduire des formules générales en formules existentielles qui aient le même statut de validité, par herbrandisation; il reste à montrer que nous pouvons faire la même chose mais dans le but de conserver le statut de prouvabilité.

Nous le faisons en remarquant que les règles de **LK** *permutent* dans les preuves sans coupure, comme Stephen C. Kleene [Kle67] l'a montré. Par exemple, nous pouvons prouver $A \wedge B \longrightarrow B \wedge A$ soit comme ceci :

$$\begin{prooftree}
\AxiomC{}
\RightLabel{Ax}
\UnaryInfC{$A, B \longrightarrow B$}
\AxiomC{}
\RightLabel{Ax}
\UnaryInfC{$A, B \longrightarrow A$}
\RightLabel{$\wedge$R}
\BinaryInfC{$A, B \longrightarrow B \wedge A$}
\RightLabel{$\wedge$L}
\UnaryInfC{$A \wedge B \longrightarrow B \wedge A$}
\end{prooftree}$$

soit comme cela :

$$\begin{prooftree}
\AxiomC{}
\RightLabel{Ax}
\UnaryInfC{$A, B \longrightarrow B$}
\RightLabel{$\wedge$L}
\UnaryInfC{$A \wedge B \longrightarrow B$}
\AxiomC{}
\RightLabel{Ax}
\UnaryInfC{$A, B \longrightarrow A$}
\RightLabel{$\wedge$L}
\UnaryInfC{$A \wedge B \longrightarrow A$}
\RightLabel{$\wedge$R}
\BinaryInfC{$A \wedge B \longrightarrow B \wedge A$}
\end{prooftree}$$

où les règles $\wedge$R et $\wedge$L ont été permutées. En général, nous avons :

Théorème 30 (Permutabilité) *Pour toutes règles R_1 et R_2 autres que Cut, nous disons que R_1* permute *sous R_2, ou que la paire R_1/R_2 permute, si et seulement si pour toute preuve de la forme :*

$$\dfrac{\dfrac{\underbrace{\begin{matrix}\pi_i^1 \\ \vdots \\ \Gamma_i^1 \longrightarrow \Delta_i^1\end{matrix}}_{1 \leq i \leq m}}{\Gamma^1 \longrightarrow \Delta^1}\,R_1 \quad \ldots \quad \dfrac{\underbrace{\begin{matrix}\pi_i^n \\ \vdots \\ \Gamma_i^n \longrightarrow \Delta_i^n\end{matrix}}_{1 \leq i \leq m}}{\Gamma^n \longrightarrow \Delta^n}\,R_1}{\Gamma \longrightarrow \Delta}\,R_2$$

où les formules principales des conclusions des règles R_1 n'est pas active dans la prémisse de la règle R_2, nous avons aussi une preuve avec les règles R_1 et R_2 permutées, autrement dit de la forme :

$$
\dfrac{\underbrace{\begin{matrix}\pi_1^j\\ \vdots\\ \Gamma_1^j \longrightarrow \Delta_1^j\end{matrix}}_{1\le j\le n}}{\Gamma_1 \longrightarrow \Delta_1}\,R_2 \quad \cdots \quad \dfrac{\underbrace{\begin{matrix}\pi_m^j\\ \vdots\\ \Gamma_m^j \longrightarrow \Delta_m^j\end{matrix}}_{1\le j\le n}}{\Gamma_m \longrightarrow \Delta_m}\,R_2
$$
$$
\overline{\qquad\qquad\Gamma \longrightarrow \Delta\qquad\qquad}\,R_1
$$

Alors, les seules paires de règles qui ne permutent pas en **LK** *sont* $\forall L/\!\!\forall R$, $\forall L/\!\!\exists L$, $\exists R/\!\!\forall R$.

Preuve : Par inspection de toutes les paires de règles possibles. Nous ne le faisons pas explicitement, car c'est fastidieux tout en n'étant pas particulièrement instructif. L'idée, comme l'exemple $\wedge$R/$\wedge$L le montre, est que comme R_1 et R_2 agissent sur des parties différentes des séquents (dans l'exemple, $\wedge$R était utilisée pour construire $B \wedge A$ à droite, alors que $\wedge$L était utilisée pour construire $A \wedge B$ à gauche), nous pouvons les appliquer dans l'ordre que nous voulons.

Ceci ne fonctionne pas sur $\forall$L/$\forall$R, parce que pour prouver $\forall x \cdot \Phi \longrightarrow \forall x \cdot \Phi \vee \Phi'$, nous devons utiliser $\forall$R en dernier, sinon nous écririons :

$$
\dfrac{\dfrac{\begin{matrix}\vdots\\ \Phi[t/x] \longrightarrow \Phi[y/x] \vee \Phi'[y/x]\end{matrix}}{\Phi[t/x] \longrightarrow \forall x \cdot \Phi \vee \Phi'}\,\forall\mathrm{R}}{\forall x \cdot \Phi \longrightarrow \forall x \cdot \Phi \vee \Phi'}\,\forall\mathrm{L}
$$

où il faut que $y \notin \mathrm{fv}(t)$ pour appliquer $\forall$R, mais en général nous ne pouvons pas montrer $\Phi[t/x] \longrightarrow \Phi[y/x] \vee \Phi'[y/x]$ à moins que $t = y$. (Prendre Φ atomique, et Φ' telle que $x \notin \mathrm{fv}(\Phi')$, alors cela ne peut être dérivé que par Ax suivi de $\vee$R.) Les autres impermutabilités sont d'une nature similaire. $\square$

Donc, pour trouver une preuve d'un séquent donné $\Gamma \longrightarrow \Delta$, nous pouvons choisir d'expanser n'importe quelle formule de Γ ou de Δ, à moins que les dépendances entre quantifications ne l'interdise. Ceci nous laisse une liberté considérable dans le choix d'une stratégie d'expansion lors de la recherche d'une preuve d'un séquent donné.

Les choses sont plus compliquées au premier ordre que dans le cas propositionnel, pour deux raisons : d'abord, les règles des quantificateurs $\forall$L et $\exists$R nous forcent à deviner un terme t parmi une infinité possible (ceci sera résolu par *unification*); ensuite, les règles de quantificateurs peuvent produire des expansions infinies, comme (lire de bas en haut, pour suivre le processus de recherche à partir du but) :

$$
\dfrac{\dfrac{\begin{matrix}\vdots\\ \dfrac{\begin{matrix}\vdots\\ \longrightarrow \Phi[t_1/x], \ldots, \Phi[t_{k-1}/x], \Phi[t_k/x], \exists x \cdot \Phi\end{matrix}}{\longrightarrow \Phi[t_1/x], \ldots, \Phi[t_{k-1}/x], \exists x \cdot \Phi}\,\exists\mathrm{R}\\ \vdots\\ \longrightarrow \Phi[t_1/x], \Phi[t_2/x], \exists x \cdot \Phi\end{matrix}}{\longrightarrow \Phi[t_1/x], \exists x \cdot \Phi}\,\exists\mathrm{R}}{\longrightarrow \exists x \cdot \Phi}\,\exists\mathrm{R}
$$

Pour éviter de telles expansions infinies, nous devrons changer de but de temps en temps, autrement dit arrêter d'essayer d'expanser $\exists x \cdot \Phi$ pour donner une chance aux autres formules dans le séquent. En d'autres termes, nous aurons besoin de stratégies d'expansion *équitables* ("fair" en anglais).

Théorème 31 (Herbrand-Skolem, syntaxique) *Une formule Φ est prouvable dans* **LK** *si et seulement si son herbrandisation est prouvable en* **LK**.

De façon duale, Φ est cohérente en **LK** *si et seulement si sa skolémisation est cohérente en* **LK**.

Preuve : Nous laissons en exercice au lecteur de vérifier que Φ est prouvable (resp. cohérente) si et seulement si n'importe laquelle de ses formes prénexes est prouvable (resp. cohérente).

Supposons donc que Φ est en forme prénexe. Si Φ n'a pas de quantificateur universel, le résultat est évident. Sinon, écrivons Φ sous la forme $Q_1x_1 \cdot \ldots \cdot Q_nx_n \cdot \forall y \cdot \Psi$, où $\Psi = \exists z_1 \cdot \ldots \cdot \exists z_m \cdot \Psi'$, et Ψ' est sans quantificateur. Nous pouvons toujours permuter toutes les règles de quantificateurs (ici, $\forall$R et $\exists$R) sous toutes les règles propositionnelles, et permuter les règles $\forall$R sous $\exists$R. Alors Φ est prouvable si et seulement si elle a une preuve π_0 où les permutations ci-dessus ont été effectuées. En remontant π_0 (le long de n'importe quelle branche) jusqu'à rencontrer une application de $\forall$R qui introduit la quantification $\forall y$, nous arrivons nécessairement à une sous-preuve π_1 de la forme :

$$\dfrac{\dfrac{\begin{array}{c}\pi \\ \vdots \\ \longrightarrow \Psi[y_1/y]\sigma_1, \Psi[y_2/y]\sigma_2, \ldots, \Psi[y_k/y]\sigma_k\end{array}}{\begin{array}{c}\longrightarrow (\forall y \cdot \Psi)\sigma_1, \Psi[y_2/y]\sigma_2, \ldots, \Psi[y_k/y]\sigma_k \\ \vdots \\ \longrightarrow (\forall y \cdot \Psi)\sigma_1, \ldots, (\forall y \cdot \Psi)\sigma_{k-1}, \Psi[y_k/y]\sigma_k\end{array}}\ \forall\mathrm{R}}{\longrightarrow (\forall y \cdot \Psi)\sigma_1, \ldots, (\forall y \cdot \Psi)\sigma_k}\ \forall\mathrm{R}$$

où $y_1, \ldots, y_k$ sont de nouvelles variables distinctes deux à deux. Nous supposerons aussi sans perte de généralité que les formules $(\forall y \cdot \Psi)\sigma_i$, $1 \leq i \leq n$ sont distinctes deux à deux.

D'une part, si Φ est prouvable, alors nous pouvons remplacer y_i dans la preuve ci-dessus par n'importe quel terme t_i, du moment que les variables libres dans t_i ne détruisent pas les conditions de bord sur les instances de $\forall$R ou $\exists$L qui sont sous π_1 dans π_0 (rappelons que ces conditions de bord disent que nous ne pouvons introduire $\forall x$ à droite ou $\exists x$ à gauche que si x n'est pas libre dans le reste du séquent). Dans ce but, il suffit de vérifier qu'aucune variable universellement quantifiée (ou aucune variable nouvelle comme les y_i) ne sont libres dans t_i. Nous pouvons choisir t_i de la forme $f(x_{i_1}, \ldots, x_{i_p})\sigma_i$ sans danger, où $x_{i_1}, \ldots, x_{i_p}$ sont les variables existentielles parmi $x_1, \ldots, x_n$, pour chaque $1 \leq i \leq k$; si l'on remplace t_i par y_i ci-dessus, on remplace toute sous-preuve de la même forme que π_1 par une autre preuve π_1' à l'intérieur de π_0 : ceci transforme π en une preuve de Φ' défini par $Q_1x_1 \cdot \ldots \cdot Q_nx_n \cdot \Psi''$, où $\Psi'' = (\exists z_1 \cdot \ldots \cdot \exists z_m \cdot \Psi')[f(x_{i_1}, \ldots, x_{i_p})/y]$ est une formule existentielle, autrement dit Φ' est une formule avec un quantificateur universel de moins que Φ.

D'autre part, si Φ' est prouvable, alors il existe une sous-preuve de $\longrightarrow \Psi[f(x_{i_1}, \ldots, x_{i_p})/y]\sigma_1, \ldots, \Psi[f(x_{i_1}, \ldots, x_{i_p})/y]\sigma_k$, où de plus les domaines de $\sigma_1, \ldots, \sigma_k$ sont inclus dans l'ensemble des variables existentielles $x_{i_1}, \ldots, x_{i_p}$. Nous pouvons aussi supposer sans perte de généralité que toutes les formules $\Psi[f(x_{i_1}, \ldots, x_{i_p})/y]\sigma_i$, $1 \leq i \leq k$, sont distinctes deux à deux. Comme f est un symbole de fonction nouveau, les seuls sous-termes dans ces formules construits avec f sont $f(x_{i_1}, \ldots, x_{i_p})\sigma_i$, $1 \leq i \leq n$. Comme la preuve peut être supposée sans coupure, elle a la propriété de la sous-formule. Ainsi, comme les formules ci-dessus sont sans quantificateur, les seuls sous-termes construits avec f dans toute la preuve sont ceux mentionnés ci-dessus. Nous remplaçons maintenant $f(x_{i_1}, \ldots, x_{i_p})\sigma_i$ par autant de variables nouvelles y_i, $1 \leq i \leq n$ dans toute la preuve : remarquons que nous ne pouvons le faire que si, chaque fois que $i \neq j$, $f(x_{i_1}, \ldots, x_{i_p})\sigma_i$ et $f(x_{i_1}, \ldots, x_{i_p})\sigma_j$ sont des termes distincts. Mais nous avons supposé que $\Psi[f(x_{i_1}, \ldots, x_{i_p})/y]\sigma_i$ et $\Psi[f(x_{i_1}, \ldots, x_{i_p})/y]\sigma_j$ étaient distinctes, donc qu'il y avait une variable libre dans $\Psi[f(x_{i_1}, \ldots, x_{i_p})/y]$ qui était remplacée par des termes différents par σ_i et par σ_j. Et toutes ces variables libres font partie de $x_{i_1}, \ldots, x_{i_p}$; donc $f(x_{i_1}, \ldots, x_{i_p})\sigma_i$ et $f(x_{i_1}, \ldots, x_{i_p})\sigma_j$ sont distincts.

Maintenant que nous avons remplacé $f(x_{i_1}, \ldots, x_{i_p})\sigma_i$ par y_i pour tout i dans la preuve, nous avons obtenu une preuve de $\longrightarrow \Psi[y_1/y]\sigma_1, \Psi[y_2/y]\sigma_2, \ldots, \Psi[y_k/y]\sigma_k$. Nous utilisons maintenant $\forall$R sur ces k formules tour à tour (nous le pouvons, parce que les y_i sont distinctes deux à deux), et obtenons une preuve de $\longrightarrow (\forall y \cdot \Psi)\sigma_1, \ldots, (\forall y \cdot \Psi)\sigma_k$, ce qui redonne la sous-preuve π_1 de départ.

Cette construction montre comment éliminer un quantificateur universel par un symbole de Herbrand. Que l'herbrandisation préserve la prouvabilité résulte d'une récurrence sur le nombre de quantifications universelles (ou de symboles de Herbrand). Que la skolémisation préserve la cohérence s'ensuit par dualité par négation. $\square$

Nous concluons par le théorème de complétude pour la logique classique du premier ordre, qui a été prouvé pour la première fois par Kurt Gödel en 1930, quoique sous une forme différente :

Théorème 32 (Correction, complétude) **LK** *est correct et complet pour la sémantique de la logique classique du premier ordre, autrement dit* $\models \Phi$ *si et seulement si* $\vdash^{\mathbf{LK}} \Phi$.

Preuve : La correction est claire. Pour montrer la complétude, supposons que $\models \Phi$. Soit $\Phi' = \exists x_1 \cdot \ldots \exists x_n \cdot \Psi$ une forme de Herbrand de Φ. Alors $\models \Phi'$. Par la version sémantique du théorème de Herbrand, il existe k instances closes $\Psi\sigma_1, \ldots, \Psi\sigma_k$ de Ψ telles que $\models \Psi\sigma_1 \vee \ldots \vee \Psi\sigma_k$. Par le théorème de complétude pour $\mathbf{LK}_0$, $\vdash^{\mathbf{LK}_0} \Psi\sigma_1, \ldots, \Psi\sigma_k$, donc $\vdash^{\mathbf{LK}} \Psi\sigma_1, \ldots, \Psi\sigma_k$ car $\mathbf{LK}_0$ est un sous-ensemble de **LK**. En utilisant $\exists$R k fois, nous en déduisons $\vdash^{\mathbf{LK}} \Phi'$, et par le théorème 31, $\vdash^{\mathbf{LK}} \Phi$. $\square$

Une preuve plus directe de la complétude d'un système de preuve pour la logique classique du premier ordre, qui n'utilise pas de forme prénexe, de Herbrand ou de Skolem, utilise ce qu'on appelle les *constantes de Henkin*, introduites par Leon Henkin en 1949. À notre avis, la preuve est au moins aussi compliquée et présente moins d'intuitions en rapport avec les façons de rechercher des preuves automatiquement. (Cf. [Man77].)

6 Digressions

Les permutabilités, l'herbrandisation et la skolémisation syntaxiques sont apparemment inutiles à notre propos. En effet, nous nous intéressons principalement à la vérité, autrement dit la validité des formules. La notion sémantique d'herbrandisation ou de skolémisation réduit le problème de la validité à celui pour les formules existentielles, et nous pourrions nous satisfaire du théorème de Herbrand (théorème 28), dont la complétude de **LK** restreint aux formules existentielles découle immédiatement.

Pourquoi avons-nous donc redérivé l'herbrandisation et la skolémisation par des moyens syntaxiques ? La réponse est, pour mieux comprendre ce que cela signifie réellement, et parce que parfois, on ne peut pas faire autrement.

6.1 Logiques non classiques

Considérons par exemple le cas des logiques intuitionniste ou linéaire. Ces logiques, comme beaucoup d'autres, sont mieux comprises au travers de leurs systèmes de preuve. Les modèles de Kripke de la logique intuitionniste fournissent un moyen rapide de vérifier que certaines formules ne sont pas intuitionnistiquement valides, mais la sémantique la plus intéressante de la logique intuitionniste est l'interprétation de Brouwer-Heyting-Kolmogorov, ou de Curry-Howard : mais cette sémantique n'est que la théorie de la preuve de la logique, retraduite en notation functionnelle (en λ-calcul).

La logique intuitionniste se présente sous la façon la plus simple en prenant le système de déduction naturelle de la logique classique du premier ordre, moins la règle :

$$\frac{\Gamma \longrightarrow \neg\neg\Phi}{\Gamma \longrightarrow \Phi}\,(\neg\neg E)$$

$$\frac{}{\Gamma, \Phi \longrightarrow \Phi}\ \text{Ax}$$

$$\frac{\Gamma, \Phi, \Phi' \longrightarrow \Delta}{\Gamma, \Phi \wedge \Phi' \longrightarrow \Delta}\ \wedge\text{L} \qquad \frac{\Gamma \longrightarrow \Phi \qquad \Gamma \longrightarrow \Phi'}{\Gamma \longrightarrow \Phi \wedge \Phi'}\ \wedge\text{R}$$

$$\frac{\Gamma, \Phi \longrightarrow \Phi'' \qquad \Gamma, \Phi' \longrightarrow \Phi''}{\Gamma, \Phi \vee \Phi' \longrightarrow \Phi''}\ \vee\text{L} \quad \frac{\Gamma \longrightarrow \Phi}{\Gamma \longrightarrow \Phi \vee \Phi'}\ \vee\text{R}_1 \quad \frac{\Gamma \longrightarrow \Phi'}{\Gamma \longrightarrow \Phi \vee \Phi'}\ \vee\text{R}_2$$

$$\frac{\Gamma \longrightarrow \Phi \qquad \Gamma, \Phi' \longrightarrow \Phi''}{\Gamma, \Phi \Rightarrow \Phi' \longrightarrow \Phi''}\ \Rightarrow\text{L} \qquad \frac{\Gamma, \Phi \longrightarrow \Phi'}{\Gamma \longrightarrow \Phi \Rightarrow \Phi'}\ \Rightarrow\text{R}$$

$$\frac{\Gamma \longrightarrow \Phi}{\Gamma, \neg\Phi \longrightarrow \Phi''}\ \neg\text{L} \qquad \frac{\Gamma, \Phi \longrightarrow \mathbf{F}}{\Gamma \longrightarrow \neg\Phi}\ \neg\text{R}$$

$$\frac{\Gamma, \Phi[t/x] \longrightarrow \Phi'}{\Gamma, \forall x \cdot \Phi \longrightarrow \Phi'}\ \forall\text{L} \qquad \frac{\Gamma \longrightarrow \Phi[y/x]}{\Gamma \longrightarrow \forall x \cdot \Phi}\ \forall\text{R}$$
$$(y \text{ non libre dans } \Gamma)$$

$$\frac{\Gamma, \Phi[y/x] \longrightarrow \Phi'}{\Gamma, \exists x \cdot \Phi \longrightarrow \Phi'}\ \exists\text{L} \qquad \frac{\Gamma \longrightarrow \Phi[t/x]}{\Gamma \longrightarrow \exists x \cdot \Phi}\ \exists\text{R}$$
$$(y \text{ non libre dans } \Gamma, \Phi')$$

$$\frac{\Gamma \longrightarrow \Phi \qquad \Gamma', \Phi \longrightarrow \Phi'}{\Gamma, \Gamma' \longrightarrow \Phi'}\ \text{Cut}$$

Figure 4: Le système **LJ** de Gentzen

(La règle $(\neg\neg I)$ peut être conservée, de toute façon elle est démontrable à partir des autres.)

Un système de séquents est obtenue en restreignant les séquents à n'avoir qu'une formule à droite, et en effectuant les ajustements nécessaires (cf. figure 4).

Et la sémantique de Kripke de la logique intuitionniste est donnée par un cadre $(W, \leq)$, où $\leq$ est un ordre (comme pour S4), un domaine non vide D, une interprétation I des symboles de fonction et de prédicats (où $I(P)$, pour chaque prédicat P d'arité m, est une fonction de D^m vers $\mathbb{P}W$ et non plus $\mathbb{B}$), et où pour toute affectation ρ :

- $w, I, \rho \models A$ si et seulement si $w' \in [\![A]\!]I\rho$ pour tout $w' \geq w$, où A est un atome; (en S4, on aurait juste écrit $w \in [\![A]\!]I\rho$;)
- $w, I, \rho \models \Phi \wedge \Phi'$ si et seulement si $w, I, \rho \models \Phi$ et $w, I, \rho \models \Phi'$;
- $w, I, \rho \models \Phi \vee \Phi'$ si et seulement si $w, I, \rho \models \Phi$ ou $w, I, \rho \models \Phi'$;
- $w, I, \rho \models \neg\Phi$ si et seulement si pour tout $w' \geq w$, non $w', I, \rho \models \Phi$; (comparer avec S4 !)
- $w, I, \rho \models \Phi \Rightarrow \Phi'$ si et seulement si pour tout $w' \geq w$, si $w', I, \rho \models \Phi$ alors $w', I, \rho \models \Phi'$; (comparer avec S4 !)
- $w, I, \rho \models \forall x \cdot \Phi$ si et seulement si pour tout $w' \geq w$, pour tout $v \in D$, $w', I, \rho[v/x] \models \Phi$;
- $w, I, \rho \models \exists x \cdot \Phi$ si et seulement s'il existe $v \in D$ tel que $w, I, \rho[v/x] \models \Phi$.

La preuve de complétude est plus compliquée que dans le cas classique, et se complique d'autant plus si l'on veut savoir si la contraction (la duplication en terme de tableaux) est éliminable ou non. En fait, la contraction n'est pas éliminable, et le mieux qu'on puisse faire sur **LJ**est de la visser dans les règles d'où elle n'est pas éliminable (notamment, $\Rightarrow$L et $\forall$L).

De plus, en **LK** toutes les stratégies étaient également acceptables pour trouver une preuve, mais en **LJ** ce n'est pas le cas. On peut le remarquer sémantiquement : certaines règles ne sont pas inversibles, autrement dit, certaines ont la propriété que la conjonction de leurs prémisses n'est pas équivalente à leur conclusion. On peut le remarquer directement sur le système de preuve, car certaines règles ne permutent pas. Il y a en fait 11 impermutabilités dans **LJ** : $\forall$L/$\forall$R, $\forall$L/$\exists$L, $\exists$R/$\forall$R (comme dans le cas classique), $\Rightarrow$L/$\Rightarrow$R, $\neg$L/$\Rightarrow$R, $\Rightarrow$L/$\neg$R, $\neg$L/$\neg$R, $\Rightarrow$L/$\vee$L, $\vee$R/$\vee$L, $\neg$L/$\vee$L et $\exists$R/$\vee$L. (Exercice.)

On peut partiellement corriger certains de ces problèmes : cf. [Dyc92] pour divers systèmes dans le cas propositionnel, certains maximisant le nombre de règles inversibles, et un évitant totalement la contraction; cf. aussi [MMO93] pour le cas intuitionniste au premier ordre. Le lecteur qui aura lu les deux articles sera sans doute convaincu que l'approche syntaxique est plus simple que l'approche sémantique. (Cf. figure 5 pour ce dernier système; les séquents sont de la forme $\Gamma \longrightarrow \Delta_1; \Delta_2$, où Γ, Δ_1 et Δ_2 sont des multi-ensembles de formules; un séquent $\Gamma \longrightarrow \Phi_1, \ldots, \Phi_m; \Psi_1, \ldots, \Psi_n$ se traduit en un séquent $\Gamma, \neg\Psi_1, \ldots, \neg\Psi_n \longrightarrow \Phi_1 \vee \ldots \vee \Phi_m$ de **LJ**, et est vérifié au monde w si et seulement si non $w, I, \rho \models \Phi$ pour au moins une formule Φ de Γ, ou $w, I, \rho \models \Phi_i$ pour un certain i, $1 \leq i \leq m$, ou il existe j, $1 \leq j \leq n$, et $w' \geq w$ tel que $w', I, \rho \models \Psi_j$.)

Le cas de la logique linéaire est encore plus frappant. Alors que l'on connait plusieurs sémantiques de la logique linéaire du premier ordre, aucune sémantique n'a été prouvée complète—du moins, aucune sémantique plus simple que le système de preuves. La plupart des résultats en logique linéaire (la cohérence, notamment) n'ont pu être déduits que par l'élimination des coupures et les propriétés des preuves par séquents sans coupure.

Du point de vue de la démonstration automatique, un système de séquents sans coupure fournit aussi directement une méthode de tableaux, et donc une méthode naturelle de recherche de preuve. (Mais en trouver une *efficace*, qui évite les répétitions durant la recherche, est beaucoup moins évident.)

6.2 Arithmétique

D'autre systèmes de preuve sont incomplets : le fait de chercher une preuve de Φ dans de tels systèmes n'est pas la même chose que de vérifier si Φ est valide. Nous serions tentés d'ajouter de nouvelles règles de preuve, dans ces conditions, pour obtenir un système de preuve complet, mais il y a des cas où c'est impossible. Ceci est une conséquence du théorème d'incomplétude de Gödel :

Théorème 33 (Gödel) *Soit D un système de preuve, dans lequel on peut interpréter l'arithmétique de Peano du premier ordre $\mathbf{PA}_1$, autrement dit il existe des formules $Z(x)$ ("x égale zéro"), $S(x,y)$ ("y est le successeur de x"), $P(x,y,z)$ ("$z = x+y$"), $T(x,y,z)$ ("$z = xy$"), $E(x,y)$ ("x égale y"), $L(x,y)$ ("x est inférieur ou égal à y") dans le langage de D qui vérifient prouvablement tous les axiomes de $\mathbf{PA}_1$.*

Alors une des assertions suivantes doit être vraie :

- *D est incomplet, autrement dit il existe une formule Φ construite au moyen des connecteurs logiques, et des quantificateurs sur les formules Z, S, P, T, E, L ci-dessus, telle que Φ soit vraie dans le modèle standard $\mathbb{N}$ de l'arithmétique, mais n'est pas prouvable dans D;*
- *ou D est incohérente, autrement dit toute formule est prouvable dans D;*
- *ou D n'est pas effectif, autrement dit il n'existe aucun algorithme décidant si une formule est (textuellement) un axiome de D ou si une règle d'inférence est une règle de D;*

$$\frac{}{\Gamma, \Phi \longrightarrow \Delta_1, \Phi; \Delta_2}\, Ax_1 \qquad \frac{}{\Gamma, \Phi \longrightarrow \Delta_1; \Delta_2, \Phi}\, Ax_2$$

$$\frac{\Gamma, \Phi, \Psi \longrightarrow \Delta_1; \Delta_2}{\Gamma, \Phi \wedge \Phi \longrightarrow \Delta_1; \Delta_2}\, \wedge L \qquad \frac{\Gamma \longrightarrow \Delta_1, \Phi; \Delta_2 \quad \Gamma \longrightarrow \Delta_1, \Psi; \Delta_2}{\Gamma \longrightarrow \Delta_1, \Phi \wedge \Psi; \Delta_2}\, \wedge R_1 \qquad \frac{\Gamma \longrightarrow \Delta_1; \Delta_2, \Phi \quad \Gamma \longrightarrow \Delta_1; \Delta_2, \Psi}{\Gamma \longrightarrow \Delta_1; \Delta_2, \Phi \wedge \Psi}\, \wedge R_2$$

$$\frac{\Gamma, \Phi \longrightarrow \Delta_1; \Delta_2 \quad \Gamma, \Psi \longrightarrow \Delta_1; \Delta_2}{\Gamma, \Phi \vee \Psi \longrightarrow \Delta_1; \Delta_2}\, \vee L \qquad \frac{\Gamma \longrightarrow \Delta_1, \Phi, \Psi; \Delta_2}{\Gamma \longrightarrow \Delta_1, \Phi \vee \Psi; \Delta_2}\, \vee R_1 \qquad \frac{\Gamma \longrightarrow \Delta_1; \Delta_2, \Phi, \Psi}{\Gamma \longrightarrow \Delta_1; \Delta_2, \Phi \vee \Psi}\, \vee R_2$$

$$\frac{\Gamma \longrightarrow \Delta_1, \Phi; \Delta_2 \quad \Gamma, \Psi \longrightarrow \Delta_1; \Delta_2}{\Gamma, \Phi \Rightarrow \Psi \longrightarrow \Delta_1; \Delta_2}\, \Rightarrow At\ L \qquad \frac{\Gamma, \Phi \longrightarrow \Psi; \Delta_2}{\Gamma \longrightarrow \Delta_1, \Phi \Rightarrow \Psi; \Delta_2}\, \Rightarrow R_1 \qquad \frac{\Gamma, \Phi \longrightarrow; \Delta_2, \Psi}{\Gamma \longrightarrow \Delta_1; \Delta_2, \Phi \Rightarrow \Psi}\, \Rightarrow R_2$$

(Φ atomique ou niée)

$$\frac{\Gamma, \Phi_1 \Rightarrow \Phi_2 \Rightarrow \Psi \longrightarrow \Delta_1; \Delta_2}{\Gamma, (\Phi_1 \wedge \Phi_2) \Rightarrow \Psi \longrightarrow \Delta_1; \Delta_2}\, \Rightarrow \wedge\ L$$

$$\frac{\Gamma, \Phi_1 \Rightarrow \Psi, \Phi_2 \Rightarrow \Psi \longrightarrow \Delta_1; \Delta_2}{\Gamma, (\Phi_1 \vee \Phi_2) \Rightarrow \Psi \longrightarrow \Delta_1; \Delta_2}\, \Rightarrow \vee\ L$$

$$\frac{\Gamma, \Phi_2 \Rightarrow \Psi \longrightarrow \Delta_1, \Phi_1 \Rightarrow \Phi_2; \Delta_2 \quad \Gamma, \Psi \longrightarrow \Delta_1; \Delta_2}{\Gamma, (\Phi_1 \Rightarrow \Phi_2) \Rightarrow \Psi \longrightarrow \Delta_1; \Delta_2}\, \Rightarrow \Rightarrow L$$

$$\frac{\Gamma, \forall x \cdot (\Phi \Rightarrow \Psi) \longrightarrow \Delta_1; \Delta_2}{\Gamma, (\exists x \cdot \Phi) \Rightarrow \Psi \longrightarrow \Delta_1; \Delta_2}\, \Rightarrow \exists L$$

$$\frac{\Gamma, (\forall x \cdot \Phi) \Rightarrow \Psi \longrightarrow \Delta_1, \forall x \cdot \Phi; \Delta_2 \quad \Gamma, \Psi \longrightarrow \Delta_1; \Delta_2}{\Gamma, (\forall x \cdot \Phi) \Rightarrow \Psi \longrightarrow \Delta_1; \Delta_2}\, \Rightarrow \forall L$$

$$\frac{\Gamma, \forall x \cdot \Phi, \Phi[t/x] \longrightarrow \Delta_1; \Delta_2}{\Gamma, \forall x \cdot \Phi \longrightarrow \Delta_1; \Delta_2}\, \forall L \qquad \frac{\Gamma \longrightarrow \Phi[y/x]; \Delta_2}{\Gamma \longrightarrow \Delta_1, \forall x \cdot \Phi; \Delta_2}\, \forall R_1 \qquad \frac{\Gamma \longrightarrow \Phi[y/x]; \Delta_2, \forall x \cdot \Phi}{\Gamma \longrightarrow \Delta_1; \Delta_2, \forall x \cdot \Phi}\, \forall R_2$$

(y non libre dans Γ, Δ_2) (y non libre dans Γ, Δ_2)

$$\frac{\Gamma, \Phi[y/x] \longrightarrow \Delta_1; \Delta_2}{\Gamma, \exists x \cdot \Phi \longrightarrow \Delta_1; \Delta_2}\, \exists L \qquad \frac{\Gamma \longrightarrow \Delta_1, \Phi[t/x]; \Delta_2}{\Gamma \longrightarrow \Delta_1, \exists x \cdot \Phi; \Delta_2}\, \exists R_1 \qquad \frac{\Gamma \longrightarrow \Delta_1; \Delta_2, \Phi[t/x], \exists x \cdot \Phi}{\Gamma \longrightarrow \Delta_1; \Delta_2, \exists x \cdot \Phi}\, \exists R_2$$

(y non libre dans $\Gamma, \Delta_1, \Delta_2$)

Figure 5: Le système de Miglioli, Moscato et Ornaghi

ce qui laisse un choix assez restreint. Il est clair que l'on préfère usuellement que D soit incomplet mais effectif et (on l'espère) cohérent. (Pour une preuve de ce théorème, voir [Joh92] ou [Man77].)

À cause du théorème d'incomplétude de Gödel, il existe des systèmes de preuve, comme ceux qui décrivent l'arithmétique de Peano du premier ordre ou la théorie des ensembles de Zermelo-Frænkel (dans laquelle on peut construire les entiers), où on ne peut pas raisonner sémantiquement ! Certains énoncés valides n'ont en effet aucune preuve dans le système correspondant. Nous sommes alors forcés de raisonner syntaxiquement, autrement dit sur le système de preuve directement.

Gerhard Gentzen a inventé les systèmes de séquents justement pour cela, et dériver des preuves de cohérence pour l'arithmétique, ce qu'il est arrivé à faire en 1934. L'idée est que les théorèmes de $\mathbf{PA}_1$ sont exactement ceux du système de séquents Z, qui est **LK** plus une infinité d'axiomes (tous les séquents atomiques clos qui sont propositionnellement dérivables des instances d'axiomes de $\mathbf{PA}_1$ sauf les axiomes de récurrence), et plus une règle appelée l'ω-règle qui exprime que l'on peut déduire $\forall n \cdot P(n)$ de $P(0)$, $P(1)$, etc., et où les preuves ont une profondeur finie (bornée). L'élimination des coupures fonctionne pour Z exactement comme pour **LK**, et nous pouvons donc dériver les propriétés usuelles à partir de l'élimination des coupures, la cohérence et la propriété de sous-formule. (La compacité échoue car l'ω-règle est infinitaire. De plus, l'élimination des coupures *échoue* dans **LK** plus l'axiome de récurrence : elle ne fonctionne que dans Z.)

Mais l'élimination des coupures fournit davantage. La preuve du théorème d'élimination des coupures est une *preuve de terminaison*, qui montre que le processus de remontée des coupures termine. Par la correspondance de Curry-Howard entre preuves et programmes, ceci signifie qu'un certain langage dont les types sont des formules du premier ordre est normalisant. En fait, ce langage est fortement normalisant, ce qui veut dire que *toutes* les stratégies de remontée des coupures terminent. Ce langage est le système T de Gödel [GLT89], et est essentiellement un λ-calcul étendu avec des constantes 0 (zéro), s (successeur), et un récurseur R qui obéit aux règles de calcul :

$$\begin{aligned} &Rfz0 \to z \\ &Rfz(sn) \to fzn(Rfzn) \end{aligned}$$

avec les types appropriés. Plus clairement, pour chaque fonction définissable f, Rfz est une fonction g définie par récursion primitive (en chaque type, même d'ordre supérieur) par $g(0) = z$ et $g(n+1) = f(z, n, g(n))$.

Des extensions de cette correspondance entre systèmes de preuve et programmes ont mené Jean-Yves Girard à montrer que $\mathbf{PA}_2$ (l'arithmétique de Peano du second ordre, avec l'axiome de récurrence complet $\forall P \cdot P(0) \Rightarrow (\forall n \cdot P(n) \Rightarrow P(s(n))) \Rightarrow \forall n \cdot P(n)$), et même $\mathbf{PA}_\omega$ (l'arithmétique de Peano d'ordre supérieur, avec l'axiome de récurrence complet et tous les raisonnements à tous les ordres) sont cohérents, en interprétant l'élimination des coupures comme la règle de réduction fondamentale d'un langage de programmation, le λ-calcul de Church restreint aux termes typables dans le système F de Girard (resp. F_ω), et en montrant que tous les termes du langage sont fortement normalisants. Le système F est essentiellement la logique propositionnelle quantifiée (intuitionniste), mais elle capture toutes les particularités de $\mathbf{PA}_2$ pour ce qui est de la typabilité.

Grâce à l'élimination des coupures, $\mathbf{PA}_1$ possède une forme de théorème de Herbrand. Il s'agit de l'*interprétation par absence de contre-exemple* ("no-counterexample interpretation" en anglais) de Georg Kreisel pour $\mathbf{PA}_1$. Alors que le théorème de Herbrand dit essentiellement que $\forall x \cdot \exists y \cdot \Psi$ est prouvable dans **LK** si et seulement si l'on peut trouver des témoins pour y sous la forme de programmes si-alors-sinon finis dépendant de x qui retournent des termes, le résultat de Kreisel dit que $\forall x \cdot \exists y \cdot \Psi$ est prouvable dans $\mathbf{PA}_1$ (ou Z) si et seulement si l'on peut trouver une fonction f prenant x en argument, définie par des programmes finis construits à l'aide de récursions primitives en n'importe quel type (celles permises par les récurseurs de Gödel), telle que $\forall x \cdot \Psi[f(x)/y]$ soit "vraie", au sens où il n'existe pas de valeur calculable pour x telle que $\Psi[f(x)/y]$ soit fausse. (Ce qui est assez tordu, on est d'accord.) C'est-à-dire que $\mathbf{PA}_1$ enrichit l'ensemble des programmes dont nous avons besoin comme témoins, en permettant une forme de récursion contrôlée, qui termine, et appelée récursion primitive. (Voir [Sho67], spécialement le chapitre 8. Voir aussi [Sch77].)

6.3 Logique d'ordre supérieur

$\mathbf{PA}_2$ et $\mathbf{PA}_\omega$ sont axiomatisables en logique du premier ordre, mais sous forme de théories relativement compliquées. Si nous souhaitons raisonner en arithmétique, il vaut mieux abandonner le domaine de la logique du premier ordre, soit en édifiant un nouveau système de preuve conçu exclusivement pour l'arithmétique, soit en se plaçant dans un cadre encore plus général.

Un de ces cadres est la *logique d'ordre supérieur*, où l'on peut quantifier non seulement sur des individus, mais aussi sur des prédicats, des fonctions, des prédicats portant sur des prédicats, et ainsi de suite. La logique d'ordre supérieur est très expressive, et on peut par exemple formaliser $\mathbf{PA}_\omega$, ainsi que l'analyse (la théorie de la droite réelle $\mathbb{R}$) du second ordre et à l'ordre supérieur, ainsi que la plupart des autres théories mathématiques sans difficulté. Le prix à payer est que pratiquement toutes les belles propriétés que nous avions auparavant sont fausses en logique d'ordre supérieur : la compacité et la complétude sont fausses, la propriété de sous-formule aussi (à moins d'accepter une notion passablement inutilisable de sous-formule), bien que l'élimination des coupures fonctionne toujours. Pire encore, on ne peut même plus effectuer d'herbrandisation ou de skolémisation à l'avance, et la recherche de preuve est considérablement compliquée par rapport au premier ordre par le fait que presque tous les sous-problèmes que l'on rencontre (l'unification en particulier) deviennent indécidables. (Voir [Hue73, Hue75] pour les problèmes, des explications, des exemples et tous les détails.)

Bien que nous n'étudions pas les logiques d'ordre supérieur, il est intéressant de donner une description rapide de la syntaxe, des règles de preuve, et des sémantiques des logiques d'ordre supérieur, de sorte que les problèmes soient plus visibles.

6.3.1 Syntaxe

La formulation la plus simple de la logique d'ordre supérieur est due à Church, qui l'a fondée sur le λ-calcul simplement typé, en ajoutant des constantes pour représenter les quantificateurs. Syntaxiquement en effet, nous pouvons utiliser l'opérateur lieur de variables λ pour représenter toutes les opérations qui lient des variables, et considérer $\forall x \cdot \Phi$ comme une abréviation de $\forall(\lambda x \cdot \Phi)$, où $\forall$ est une constante qui prend une fonction f à valeurs booléennes en argument et retourne $\top$ si et seulement si f est la fonction constante $\top$. Le besoin d'utiliser un calcul *typé* au lieu du λ-calcul non typé provient du fait que le système sans les types est incohérent. Un tel système non typé a été inventé par Gottlob Frege dans les années 1890 pour formaliser les mathématiques; mais on peut y écrire une formule qui se contredit elle-même, le paradoxe de Russell, qui est essentiellement la même chose que le paradoxe du menteur. Définissons F comme étant la fonction :

$$\lambda x \cdot \neg x(x)$$

et considérons $F(F)$. Par β-réduction, $F(F)$ est égal à $\neg F(F)$, d'où l'on déduit n'importe quelle formule.

Définissons en premier l'algèbre des types simples. Les types de base sont soit $\mathbb{B}$ (le type des formules) soit $\mathbb{T}$ (le type des termes). Nous pourrions aussi raffiner $\mathbb{T}$ en plusieurs types différents, et c'est à la base ce que les logiques à sortes font (cf. [SS89, Wal88]). Les *types simples* sont les éléments du plus petit ensemble contenant les types de base, et tel que si τ et τ' sont des types simples, alors $\tau \to \tau'$ est un type simple. Comme d'habitude, nous supposons que les flèches associent à droite, donc $\tau \to \tau' \to \tau''$ signifie $\tau \to (\tau' \to \tau'')$.

Pour définir les expressions de la logique d'ordre supérieur, fixons une *signature* Σ, qui est un ensemble de constantes pour chaque type. Nous supposons en général qu'il existe au moins une constante de chaque type. De même, nous supposons qu'il existe une infinité de variables x_τ de chaque type τ. Nous omettrons les indices de types lorsqu'ils ne sont pas nécessaires. La *syntaxe* des expressions d'ordre supérieur e est alors définie en même temps que les jugements de typage $e : \tau$ comme suit :

- x_τ est une expression de type τ (nous notons $x_\tau : \tau$);
- si c est une constante de type τ, alors c est une expression de type τ;

- si $e : \tau \to \tau'$ et $e' : \tau$, alors ee', l'*application* de e à e', est une expression de type τ';
- si $e : \tau'$, alors $\lambda x_\tau \cdot e$, l'*abstraction* de e sur x, est une expression de type $\tau \to \tau'$.

Finalement, nous supposons que Σ contient les constantes logiques $\wedge : \mathbb{B} \to \mathbb{B} \to \mathbb{B}$ (conjonction), $\vee : \mathbb{B} \to \mathbb{B} \to \mathbb{B}$ (disjonction), $\Rightarrow : \mathbb{B} \to \mathbb{B} \to \mathbb{B}$ (implication), $\neg : \mathbb{B} \to \mathbb{B}$ (négation), $\forall_\tau : (\tau \to \mathbb{B}) \to \mathbb{B}$ (quantification universelle sur les objets de type τ, pour chaque τ), et $\exists_\tau : (\tau \to \mathbb{B}) \to \mathbb{B}$ (quantification existentielle sur les objets de type τ, pour chaque τ). Les variables libres, les substitutions sont définies comme en logique du premier ordre, en prenant soin de renommer les variables liées pour éviter les captures de noms de variables.

Remarquons que nous pouvons coder les termes et les formules de la logique du premier ordre dans ce cadre. Nous traduisons les fonctions n-aires f en constantes f de type $\mathbb{T} \to \ldots \to \mathbb{T} \to \mathbb{T}$ (avec n flèches $\to$), et les symboles de prédicat n-aires P en constantes P de type $\mathbb{T} \to \ldots \to \mathbb{T} \to \mathbb{B}$ (avec n flèches $\to$ encore). Ainsi, une application $f(t_1, \ldots, t_n)$ est codée sous forme de l'application $(\ldots(ft_1)t_1 \ldots t_n)$. Si nous utilisons la convention que l'application associe à gauche, nous pouvons aussi l'écrire $ft_1 \ldots t_n$.

6.3.2 Sémantique standard

La *sémantique standard* est essentiellement ce à quoi on s'attend après avoir vu la définition, sachant ce qu'est la sémantique de la logique du premier ordre. Une *interprétation standard* est un ensemble non vide D_I muni d'une interprétation pour chaque constante, cette interprétation respectant le type de la constante. Plus précisément, nous définissons l'interprétation des types par :

- $[\![\mathbb{T}]\!]I = D_I$, $[\![\mathbb{B}]\!]I = \mathbb{B}$;
- $[\![\tau \to \tau']\!]I$ est l'espace de toutes les applications (fonctions totales) de $[\![\tau]\!]I$ vers $[\![\tau']\!]I$.

Pour que l'interprétation $I(c)$ respecte le type τ de c, nous devons imposer la contrainte : $I(c) \in [\![\tau]\!]I$, autrement dit les constantes de type $\mathbb{T}$ doivent représenter des valeurs de D_I, les constantes de type $\mathbb{T} \to \mathbb{T}$ doivent représenter des applications unaires de D_I vers D_I, et ainsi de suite.

Naturellement, nous imposons aussi la contrainte que toutes les constantes logiques $\wedge$, $\vee$, $\Rightarrow$, $\neg$, $\forall$, $\exists$ aient le sens souhaité, et :

- $[\![ee']\!]I\rho = [\![e]\!]I\rho([\![e']\!]I\rho)$;
- $[\![\lambda x_\tau \cdot e]\!]I\rho$ est l'application envoyant chaque $v \in [\![\tau]\!]I$ vers $[\![e]\!]I(\rho[v/x])$.

Dans la sémantique standard, nous pouvons définir l'arithmétique par les axiomes de Peano de base, plus l'axiome de récurrence (qui peut maintenant s'écrire sous forme d'un axiome unique, grâce à la quantification à l'ordre supérieur) :

$$\forall P_{\mathbb{T} \to \mathbb{B}} \cdot P(0) \wedge (\forall x_{\mathbb{T}} \cdot P(x) \Rightarrow P(s(x))) \Rightarrow \forall x_{\mathbb{T}} \cdot P(x)$$

et nous obtenons ainsi l'*arithmétique d'ordre supérieur*. Nous pouvons montrer que $\mathbb{N}$ est un modèle de l'arithmétique d'ordre supérieur, et que c'est le seul à isomorphisme près.

Le problème principal de l'arithmétique d'ordre supérieur est que, à cause du théorème de Gödel, et parce que nous avons un modèle ($\mathbb{N}$), tout système axiomatique cohérent et effectif doit être incomplet. (Le système ci-dessus est effectif au second ordre, quoique ce ne soit pas facile à voir, par exemple.) De plus, nous pouvons coder l'arithmétique directement en logique d'ordre supérieure sans introduire de nouveaux symboles comme s ou de nouveaux axiomes comme ceux ci-dessus, en représentant les entiers n comme des entiers de Church $\lambda f_{\mathbb{T} \to \mathbb{T}} \cdot \lambda x_{\mathbb{T}} \cdot f(f(\ldots f(x) \ldots)$ (où f est appliquée n fois à x). De nouveau, à cause du théorème de Gödel, tout système axiomatique cohérent et effectif pour la logique d'ordre supérieur est incomplet.

6.3.3 Systèmes de preuve

Nous décrivons maintenant comment le système de séquents du premier ordre **LK** peut être étendu aux ordres supérieurs. Le système obtenu est cohérent, ainsi que Girard l'a montré, donc il doit être incomplet.

Nous définissons les règles de réduction comme étant celles du λ-calcul :

(β) $(\lambda x \cdot e)e' \to e[e'/x]$

(η) $\lambda x \cdot ex \to e$ si $x \notin \mathrm{fv}(e)$

Ceci définit deux relations binaires (notées $\to$ ci-desus), que nous noterons β et η respectivement. Soit λ l'une des deux relations β ou $\beta \cup \eta$. Nous appelons $\longrightarrow_\lambda$ la plus petite relation contenant λ et stable par application de contexte (c'est-à-dire telle que si $e \longrightarrow_\lambda e'$, alors $eu \longrightarrow_\lambda eu'$, $ue \longrightarrow_\lambda u'e$, et $\lambda x \cdot e \longrightarrow_\lambda \lambda x \cdot e'$). Notons $\longrightarrow^*_\lambda$ la clôture réflexive transitive de $\longrightarrow_\lambda$, et $=_\lambda$ sa clôture réflexive symétrique transitive. Remarquons que, si $e =_\lambda e'$, alors trivialement e et e' doivent avoir les mêmes interprétations standard.

Le λ-calcul simplement typé a la propriété remarquable que nous pouvons décider si $e =_\lambda e'$ en normalisant d'abord e et e' (c'est-à-dire en les réduisant tant que c'est possible, jusqu'à obtenir une *forme normale*), et en comparant les formes normales textuellement. C'est parce que ce calcul est confluent et terminant; pour plus de détails sur le λ-calculus, consulter [Bar84].

Nous pouvons étendre **LK** simplement en remplaçant les règles d'inférence des quantificateurs par les versions typées suivantes :

$$\frac{\Gamma, \Phi[t/x] \longrightarrow \Delta}{\Gamma, \forall x_\tau \cdot \Phi \longrightarrow \Delta}\,\forall\mathrm{L} \qquad \frac{\Gamma \longrightarrow \Phi[y_\tau/x], \Delta}{\Gamma \longrightarrow \forall x \cdot \Phi, \Delta}\,\forall\mathrm{R}\ (y \text{ non libre dans } \Gamma, \Delta)$$

$$\frac{\Gamma, \Phi[y_\tau/x] \longrightarrow \Delta}{\Gamma, \exists x_\tau \cdot \Phi \longrightarrow \Delta}\,\exists\mathrm{L}\ (y \text{ non libre dans } \Gamma, \Delta) \qquad \frac{\Gamma \longrightarrow \Phi[t/x], \Delta}{\Gamma \longrightarrow \exists x \cdot \Phi, \Delta}\,\exists\mathrm{R}$$

où t est de type τ, et la règle :

$$\frac{(\Gamma \longrightarrow \Delta)[u/x] \qquad u =_\lambda v}{(\Gamma \longrightarrow \Delta)[v/x]}\,(\lambda)$$

qui signifie que nous pouvons toujours remplacer toute expression par une qui lui est λ-équivalente. En pratique, on maintient toutes les formules en forme normale pour $\longrightarrow_\lambda$, et ceci règle le cas de la règle (λ).

Mais ceci ne traite vraiment du cas de la règle (λ) que si toutes les instances en sont connues à l'avance. Considérons en effet la règle $\exists$R : pour montrer $\Gamma \longrightarrow \exists x_\tau \cdot \Phi, \Delta$, nous devons deviner une expression t de type τ telle que $\Gamma \longrightarrow \Phi[t/x], \Delta$ soit prouvable. Nous verrons que, au premier ordre, nous n'avons pas à deviner ce terme, parce qu'une procédure spéciale d'*unification* nous la fournira. L'unification au premier ordre signifie trouver une substitution σ telle que $t\sigma = t'\sigma$, étant donnés deux termes t et t'. Aux ordres supérieurs, l'*unification d'ordre supérieur* (trouver des substitutions σ envoyant chaque variable vers des expressions du même type telles que $t\sigma =_\lambda t'\sigma$, étant donnés deux termes t et t' du même type) ne suffit plus dans les systèmes de tableaux. Mais même à part ça, ce problème est indécidable dès que des symboles de fonction ou de prédicats sont autorisés (et pourvu que l'on ait deux constantes aussi).

Pour voir que le système de logique d'ordre supérieur ci-dessus présente de nombreuses autres difficultés, considérons l'exemple :

$$\neg \exists F_{\mathbb{T} \to \mathbb{T} \to \mathbb{B}} \cdot \forall P_{\mathbb{T} \to \mathbb{B}} \cdot \exists x_{\mathbb{T}} \cdot \forall y_{\mathbb{T}} \cdot Fxy \Leftrightarrow Py$$

qui exprime une forme faible du théorème de Cantor (il n'existe pas d'application surjective F de D vers $\mathbb{P}(D)$, où D est représenté par le type $\mathbb{T}$, et l'ensemble des parties $\mathbb{P}(D)$ est représenté comme l'ensemble des fonctions caractéristiques de type $\mathbb{T} \to \mathbb{B}$; la formule exprime la surjectivité en disant que pour tout sous-ensemble P de D, il existe un x tel que $Fx = P$).

Nous serions tentés de skolémiser, et donc d'essayer de réfuter :

$$F(cX)y \Leftrightarrow Xy$$

où $X_{\mathbb{T}\to\mathbb{B}}$ et $y_{\mathbb{T}}$ sont les deux variables (X représente P), F est une constante de Skolem de type $\mathbb{T} \to \mathbb{T} \to \mathbb{B}$, et c est une fonction de Skolem unaire, de type $(\mathbb{T} \to \mathbb{B}) \to \mathbb{T}$. Nous pouvons conclure en utilisant la substitution $[\lambda x_{\mathbb{T}} \cdot \neg(Fxx)/X, c(\lambda x_{\mathbb{T}} \cdot \neg(Fxx))/y]$.

Mais le problème est que, bien que la skolémisation soit correcte pour la sémantique standard, elle est *incorrecte* pour le système de preuve. La skolémisation dépend en effet de l'axiome du choix :

$$\forall P_{\tau\to\tau'\to\mathbb{B}} \cdot (\forall x_\tau \cdot \exists y_{\tau'} \cdot Pxy) \Rightarrow (\exists f_{\tau\to\tau'} \cdot \forall x_\tau \cdot Px(fx))$$

qui est improuvable dans ce système, bien qu'il soit vrai dans la sémantique standard.

La solution usuelle est d'enrichir le système de preuve par une règle exprimant l'axiome du choix. Nous avons alors le problème inverse : tenter de prouver une formule en la skolémisant puis en cherchant à trouver une preuve de la formule sans quantificateur qui en résulte est une méthode incomplète. En fait, nous devons de temps en temps utiliser l'axiome du choix (ou skolémiser certains quantificateurs seulement) en plein milieu de la recherche d'une preuve. Ceci est dû au fait que le remplacement de variables par des expressions d'ordre supérieur peut complètement changer la forme de la formule. Par exemple :

$$\forall P_{\mathbb{B}\to\mathbb{B}} \cdot P(\exists x_{\mathbb{T}} \cdot \Phi(x))$$

En appliquant la substitution $[\lambda y \cdot y/P]$, nous pouvons en déduire :

$$\exists x_{\mathbb{T}} \cdot \Phi(x)$$

où le quantificateur existentiel apparaît positivement, mais en appliquant $[\neg/P]$, nous obtenons :

$$\neg\exists x_{\mathbb{T}} \cdot \Phi(x)$$

où la quantification existentielle apparaît négativement. Il n'y a aucun moyen de décider à l'avance comment skolémiser ce quantificateur existentiel.

Le fait de remplacer P par des expressions booléennes plus compliquées peut en fait instancier la formule $P(\exists x_{\mathbb{T}} \cdot \Phi(x))$ pour en extraire n'importe quelle formule. C'est la raison pour laquelle nous disons que la propriété de la sous-formule n'est pas valide des preuves sans coupure (et pourquoi l'unification ne suffit pas dans les tableaux). L'élimination des coupures fonctionne, mais ne fournit qu'une aide minime pour automatiser la recherche de preuve, ici.

6.3.4 Sémantique générale

Nous avons vu qu'il y avait au moins quatre variantes différentes de logiques d'ordre supérieur, selon que nous admettons la β-égalité ou la $\beta\eta$-égalité, et selon que nous admettons ou non l'axiome du choix. Ces logiques ont réellement des ensembles de théorèmes différents, bien que cela ne se voie pas à l'examen de la sémantique standard.

Pour montrer plus précisément ce qui arrive dans ces divers systèmes, Leon Henkin a proposé une sémantique plus générale des logiques d'ordre supérieur en 1950 (see [And86]). L'astuce est que les dénotations des types fonctionnels peuvent devenir trop gros : par exemple, si $\mathbb{T}$ est interprété par $\mathbb{N}$, $\mathbb{T} \to \mathbb{T}$ doit être interprété par l'ensemble de toutes les fonctions de $\mathbb{N}$ vers $\mathbb{N}$ dans la sémantique standard. Mais on ne peut décrire qu'un sous-ensemble dénombrable de fonctions par des formules, et nous pouvons penser que c'est une cause raisonnable d'incomplétude.

Henkin a donc conçu une notion plus abstraite d'interprétation des types, de l'application et de l'abstraction. Une *interprétation générale* remplace l'ensemble non vide D_I par une famille

D d'ensembles D_τ indicés par les types τ, munie d'un *opérateur d'application*$@_{\tau,\tau'}$, ou @ en abrégé, de $D_{\tau\rightarrow\tau'} \times D_\tau$ vers $D_{\tau'}$. Le couple $(D, @)$ est une *structure applicative* si et seulement si toutes les expressions ont des interprétations, où l'interprétation est définie en modifiant la notion d'interprétation standard :

- $[\![ee']\!]I\rho = @([\![e]\!]I\rho, [\![e']\!]I\rho)$;
- $[\![\lambda x_\tau \cdot e]\!]I\rho$ est un élément f de $D_{\tau\rightarrow\tau'}$, où $e : \tau'$, tel que pour tout $v \in [\![\tau]\!]I$, $@(f, v) = [\![e]\!]I(\rho[v/x])$.

Une structure applicative est donc une structure qui contienne suffisamment d'éléments pouvant être interprétés comme des fonctions de D_τ vers $D_{\tau'}$.

Nous pouvons maintenant définir le concept de *structure générale*. Une structure générale est une structure applicative munie d'une *valuation*, qui est par définition une application v de $D_{I\!B}$ vers $\mathbb{B}$ qui interprète les éléments de $D_{I\!B}$ comme des booléens, et telle que $v(@([\![\neg]\!]I\rho, x))$ soit la négation de $v(x)$, $v(@(@([\![\wedge]\!]I\rho, x), y))$ soit la conjonction de $v(x)$ et $v(y)$, et ainsi de suite. En effet les structures applicatives n'imposent pas que l'interprétation de $D_{I\!B}$ soit $\mathbb{B}$.

Remarquer aussi que nous n'avons pas exigé que $f \in D_{\tau\rightarrow\tau'}$ soit unique parmi tous les éléments représentant la fonction $\lambda x_\tau \cdot e$ dans la définition des interprétations générales. Et en effet, il existe des structures applicatives non extensionnelles, autrement dit des structures applicatives dans lesquelles $\forall x \cdot fx = gx$ n'implique pas $f = g$, pour certains termes f et g où x n'est pas libre. En fait, l'η-égalité est le cas particulier où g est $\lambda y \cdot fy$ (par (β), pour tout x, $fx = (\lambda y \cdot fy)x$, donc l'extensionnalité implique $f = \lambda y \cdot fy$ dans ce cas), et il existe des structures applicatives où (η) n'est pas valide; en particulier, (η) n'est pas prouvable à partir de la β-égalité seule.

Remarquer finalement que les structures applicatives peuvent ne pas être assez riches pour satisfaire l'axiome du choix. Elles n'ont besoin que de contenir des fonctions représentant les λ-abstractions, mais pas les fonctions de choix. En particulier l'axiome du choix n'est pas prouvable à partir du système de séquents d'ordre supérieur de la dernière section.

Les structures générales jouent essentiellement le même rôle pour le système de séquents d'ordre supérieur que les interprétations du premier ordre jouent pour **LK**. Nous pouvons alors redévelopper une forme de théorie de Herbrand. En particulier, une structure générale parfaitement acceptable est celle de toutes les classes d'équivalence de λ-termes typés modulo la congruence engendrée par (β) et (η), où @ est l'application ordinaire du λ-calculus et $D_{I\!B}$ est l'ensemble des booléens de Church par exemple. Cette interprétation syntaxique est tout à fait analogue à celle de Herbrand sur l'univers des termes clos en logique du premier ordre.

Il n'est pas très surprenant que le système de séquents d'ordre supérieur de la dernière section soit correct et complet pour la sémantique des structures générales. Bien sûr, nous perdons la possibilité d'axiomatiser $\mathbb{N}$ ou $\mathbb{R}$ par des logiques d'ordre supérieur en sémantique des structures générales, à cause du théorème de Gödel.

En particulier, il reste des formules valides improuvables. Un cas particulièrement frustrant, dû à Michael Kohlhase [Koh95], est :

$$\neg cb \vee c(\neg\neg b)$$

qui est improuvable dans notre système de séquents, ou dans le premier système de tableaux de Kohlhase, qui est intensionnel (autrement dit, non extensionnel). La raison en est que la seule façon de le prouver est de montrer que $b = \neg\neg b$. Malheureusement, nous pouvons prouver $b \Leftrightarrow \neg\neg b$, et bien que nous puissions penser que l'équivalence logique et l'égalité devrait être les mêmes objets de type $I\!B$, les structures générales permettent le contraire. (Choisir $D_{I\!B}$ contenant plus de deux éléments.)

On peut le réparer en considérant les *modèles généraux*, qui sont des structures générales dans lesquelles $D_{I\!B}$ est contraint à être exactement $\mathbb{B}$, et v à être l'identité. Au système de preuve, nous ajoutons les axiomes suivants d'extensionnalité booléenne (qui était donc improuvable à partir des autres) :

$$\forall F_{I\!B} \cdot \forall G_{I\!B} \cdot (F \Leftrightarrow G) \Leftrightarrow F \doteq G$$

où $\doteq$ est l'égalité définissable de la logique d'ordre supérieur, à savoir $F \doteq G$ égale $\forall P \cdot P(F) \Rightarrow P(G)$. (Il s'agit de l'*égalité de Leibniz*, qui dit que deux objets sont égaux si et seulement s'ils ont les mêmes propriétés.) Le système de preuve qui en résulte est alors correct et complet pour les modèles généraux (mais pas pour les modèles standard, bien sûr).

Nous terminons ici notre étude des notions et chausses-trappes fondamentales de la logique d'ordre supérieur. Pour plus de détails, consulter [And86], un bon livre sur la logique et les preuves consacré pour une bonne part à un système minimal décrivant la logique d'ordre supérieur (avec la règle η, mais sans l'axiome du choix). Il existe des méthodes de preuve automatique pour les logiques d'ordre supérieur : la résolution d'ordre supérieur de Huet [Hue73] et les tableaux d'ordre supérieur de Kohlhase [Koh95] en sont deux. Ces méthodes ont toutes besoin d'une forme d'unification modulo la théorie de la β (resp. $\beta\eta$) conversion entre λ-termes typés. La procédure d'unification d'ordre supérieur de Huet est la référence [Hue75], et est en général efficace en pratique. On consultera aussi [SG89] pour une présentation complètement détaillée. On peut étendre une forme de théorie de Herbrand aux ordres supérieurs, comme montré par Miller [Mil87].

References

[And86] Peter B. Andrews. *An Introduction to Mathematical Logic and Type Theory: To Truth through Proof.* Computer Science and Applied Mathematics. Academic Press, 1986.

[Bar84] Henk P. Barendregt. *The Lambda Calculus, Its Syntax and Semantics*, volume 103 of *Studies in Logic and the Foundations of Mathematics.* North-Holland Publishing Company, Amsterdam, revised edition, 1984.

[DW85] Martin D. Davis and Elaine J. Weyuker. *Computability, Complexity and Languages.* Academic Press, New York, 1985.

[Dyc92] Roy Dyckhoff. Contraction-free sequent calculi for intuitionistic logic. *Journal of Symbolic Logic*, 57(3):795–807, 1992.

[GLT89] Jean-Yves Girard, Yves Lafont, and Paul Taylor. *Proofs and Types*, volume 7 of *Cambridge Tracts in Theoretical Computer Science.* Cambridge University Press, 1989.

[Hue73] Gérard P. Huet. A mechanization of type theory. In *Proceedings of the 3rd International Joint Conference on Artificial Intelligence*, pages 139–146, Stanford University, Stanford, California, August 1973.

[Hue75] Gérard P. Huet. A unification algorithm for typed λ-calculus. *Theoretical Computer Science*, 1:27–57, 1975.

[Joh92] Peter T. Johnstone. *Notes on Logic and Set Theory.* Cambridge University Press, 1992.

[Kle67] Stephen Cole Kleene. *Mathematical Logic.* John Wiley and Sons, 1967.

[Koh95] Michael Kohlhase. Higher-order tableaux. In *Workshop on Theorem Proving with Analytic Tableaux and Related Methods*, 1995.

[Man77] Youri I. Manin. *A Course in Mathematical Logic.* Graduate Texts in Mathematics. Springer Verlag, 1977.

[Mil87] Dale A. Miller. A compact representation of proofs. *Studia Logica*, 46(4), 1987.

[MMO93] Pierangelo Miglioli, Ugo Moscato, and Mario Ornaghi. How to avoid duplications in refutation systems for intuitionistic and Kuroda logic. In *Workshop on Theorem Proving with Analytic Tableaux and Related Methods*, 1993.

[Rob66] Abraham Robinson. *Non-standard Analysis.* Studies in logic and the foundations of mathematics. Amsterdam, North-Holland Pub. Co., 1966.

[Sch77] Helmut Schwichtenberg. Proof theory: Some applications of cut-elimination. In Jon Barwise, editor, *Handbook of Mathematical Logic*, chapter D.2, pages 867–895. North-Holland Publishing Company, 1977.

[SG89] W. Snyder and J. Gallier. Higher order unification revisited: Complete sets of tranformations. *Journal of Symbolic Computation*, 8(1 & 2):101–140, 1989. Special issue on unification. Part two. Available at `ftp://ftp.cis.upenn.edu/pub/papers/gallier/hounif.dvi.Z`.

[Sho67] Joseph R. Shoenfield. *Mathematical Logic*. Addison Wesley, 1967.

[SS89] Manfred Schmidt-Schauß. *Computational Aspects of an Order-Sorted Logic with Term Declarations*, volume 395 of *Lecture Notes in Artificial Intelligence*. Springer Verlag, 1989.

[Wal88] Christoph Walther. Many-sorted unification. *Journal of the ACM*, 35(1):1–17, 1988.