Master 1 Ingéniérie Informatique

Éléments de logique du 1$^{\text{er}}$ ordre

Notes pour le cours de

Preuves assistées par ordinateur

Alexandre Miquel

(Version provisoire – 17/04/2008)

Chapitre 1

Le langage de la logique du premier ordre

Ce chapitre est consacré à la définition du langage de la logique du premier ordre. On y donne les règles de construction des termes et des formules, et on y définit quelques notions-clé telles que les notions de variable libre et de variable liée, les notions d'alpha-conversion et de renommage, ainsi que les opérations de substitution dans les termes et dans les formules.

Ce chapitre ne traite que de syntaxe. La notion de démonstration ne sera abordée qu'au chapitre 2 (en déduction naturelle).

1.1 Généralités

Le langage du calcul des prédicats est défini à partir de deux classes d'expressions :

- Les *termes*, qui expriment les objets du discours, comme par exemple l'expression $3 \times x + 1$. Suivant la théorie considérée, les termes peuvent désigner des nombres, des ensembles, des listes, des arbres, etc.
- Les *formules*, qui expriment les propositions susceptibles d'être énoncées par le discours, comme par exemple l'expression

$$\forall x \, \exists y \; (x = 2 \times y \;\vee\; x = 2 \times y + 1)$$

(« pour tout x, il existe y tel que ou bien $x = 2y$, ou bien $x = 2y + 1$ »).[1]

D'un point de vue syntaxique, les termes et les formules sont des arbres finis construits à partir d'un jeu de symboles parmi lesquels on trouve :

- Des variables (x, y, z, etc.), qui désignent des objets indéterminés ;

[1] Par rapport à la syntaxe des langues naturelles, les termes correspondent aux groupes nominaux (par ex. : « le second fils de ma petite sœur » tandis que les formules correspondent aux phrases (par ex. : « la souris verte mange le chat angora »). La distinction entre terme et formule est ainsi complètement calquée sur la distinction entre groupe nominal et groupe verbal qui structure les langues indo-européennes. Il est tout à fait possible d'envisager une présentation de la logique qui ne repose pas sur une telle distinction, suivant l'esprit des langues asiatiques. Par exemple, la *théorie des types* (de Church ou de Martin-Löf) est une présentation alternative de la logique où il n'y a que des objets (représentés par des termes), et où les propositions sont des termes d'un type particulier.

– Des symboles de constante (0, 2, π, $\mathbb{N}$, etc.), qui désignent des objets particuliers dont le sens est défini par la théorie considérée ;
– Des symboles de fonction (+, $\times$, $\sqrt{\ }$, log, cos, etc.), qui permettent de construire des objets complexes à partir des variables et des constantes ;
– Des symboles de prédicat (=, $\leq$, >, $\in$, etc.), qui permettent d'exprimer des relations (i.e. des propositions élémentaires) entre les objets ;
– Des connecteurs ($\wedge$, $\vee$, $\Rightarrow$), qui permettent de combiner les propositions élémentaires en des propositions plus complexes ;
– Des quantificateurs ($\forall$, $\exists$), qui permettent d'exprimer des propositions universelles (« pour tout... ») et existentielles (« il existe... »).

Certains de ces symboles, tels que les variables, les connecteurs et les quantificateurs, se retrouvent dans toutes les langages du premier ordre, tandis que d'autres, tels que les symboles de constante, de fonction et de prédicat, dépendent en général de la théorie considérée. On a l'habitude de regrouper les déclarations des symboles caractéristiques d'une théorie en une signature :

Définition 1 (Signature) — *Une* signature Σ *est donnée par deux ensembles de symboles disjoints, à savoir :*
– *un ensemble* Σ_F *de* symboles de fonction *(notés f, g, h, etc.) ;*
– *un ensemble* Σ_P *de* symboles de prédicat *(notés p, q, r, etc.).*

*On suppose en outre que chaque symbole s (de fonction ou de prédicat) défini dans Σ est muni d'un entier naturel qui indique le nombre d'arguments auquel ce symbole est susceptible d'être appliqué. Cet entier est appelé l'*arité[2] *du symbole s, et est noté* $\mathrm{ar}_\Sigma(s)$.

Dans tout ce qui suit, on considérera qu'un symbole de constante n'est rien d'autre qu'un symbole de fonction d'arité nulle. La définition ci-dessus n'a donc pas besoin de mentionner les symboles de constante qui ne sont qu'un cas particulier des symboles de fonction. Plus généralement, on parlera de symbole (de fonction ou de prédicat) *unaire*, *binaire*, *ternaire*, n*-aire* pour désigner un symbole (de fonction ou de prédicat) d'arité 1, 2, 3, n, etc.

Les deux exemples qui suivent présentent les signatures de deux théories du premier ordre très importantes en logique :

Exemple 1 (Signature de l'arithmétique de Peano) — *La signature de l'arithmétique de Peano est définie par :*
– *un symbole de constante noté 0 (« zéro ») ;*
– *un symbole de fonction s (« successeur ») d'arité 1 ;*
– *deux symboles de fonction + (« plus ») et $\times$ (« fois ») d'arité 2 ;*
– *un symbole de prédicat = (« égale ») d'arité 2.*

Exemple 2 (Signature de la théorie des ensembles) — *La signature de la théorie des ensembles est définie par :*
– *aucun symbole de constante ou de fonction ;*
– *deux symboles de prédicat = (« égale ») et $\in$ (« appartient à ») d'arité 2.*

[2]Le mot *arité* est formé à partir du suffixe *-aire* des mots *unaire*, *binaire*, *ternaire*.

1.2 Les termes

1.2.1 Définition

Soit Σ une signature fixée. La définition des termes (dans Σ) repose sur la présence d'un ensemble auxiliaire de symboles qu'on appelle des *variables* (notation : x, y, z, etc.), en nombre infini dénombrable.

Définition 2 (Terme) — *On appelle* terme *(ou Σ-terme) toute expression construite par application finie des règles suivantes :*

- *si x est une variable, alors x est un terme ;*
- *si f est un symbole de fonction d'arité n (dans Σ) et si $t_1, \ldots, t_n$ sont n termes, alors $f(t_1, \ldots, t_n)$ est un terme ;*

Dans ce qui suit, les termes sont notés avec les lettres t, u, etc.

Si c est un symbole de constante — c'est-à-dire un symbole de fonction d'arité 0 — on note c plutôt que $c()$ le terme obtenu en appliquant le symbole de fonction c à la suite de termes vide. Certains symboles de fonction binaires tels que $+$ ou $\times$ s'utilisent en notation infixe ; aussi a-t-on l'habitude d'écrire $t + u$ plutôt que $+(t, u)$, et $t \times u$ voire tu plutôt que $\times(t, u)$.

Exemples 3
1. *Dans l'arithmétique de Peano, on utilise les abréviations $1 = s(0)$, $2 = s(1)$, $3 = s(2)$, $4 = s(3)$, $5 = s(4)$, etc. Ainsi le terme $+(\times(s(s(0)),\ x),\ s(0))$ se note-t-il plus simplement $2x + 1$.*
2. *Dans le langage de la théorie des ensembles (qui est dépourvu de symbole de constante ou de fonction), les seuls termes sont les variables.*

1.2.2 Ensemble des variables libres

Étant donné un terme t, on appelle *ensemble des variables libres* de t et on note $FV(t)$ l'ensemble (fini) de toutes les variables qui ont au moins une occurrence dans le terme t.[3] Formellement, cet ensemble est défini par récurrence sur la structure de t à l'aide des équations suivantes :

$$\begin{array}{rcl} FV(x) & = & \{x\} \\ FV(f(t_1, \ldots, t_n)) & = & FV(t_1) \cup \cdots \cup FV(t_n) \end{array}$$

On dit qu'un terme t est *clos* lorsque $FV(t) = \varnothing$. Dans le cas contraire, on dit que t est un terme *ouvert*.

1.2.3 Substitution

Étant donnés un terme t, une variable x et un terme u, on note $t\{x := u\}$ le terme obtenu en remplaçant dans t toutes les occurrences (libres) de la variable x

[3] Dans un terme, toutes les occurrences d'une variable sont libres, car les termes ne comprennent pas de construction permettant de lier une variable (c'est-à-dire de la rendre muette). La nécessité de distinguer les occurrences libres d'une variable de ses occurrences liées n'apparaît donc que dans les formules, en raison de la présence des symboles lieurs $\forall$ et $\exists$.

par le terme u. Formellement, le terme $t\{x := u\}$ est défini par récurrence sur la structure du terme t par les équations :

$$\begin{array}{rcll} x\{x := u\} & = & u & \\ y\{x := u\} & = & y & \text{(si } y \neq x) \\ f(t_1, \ldots, t_n)\{x := u\} & = & f(t_1\{x := u\}, \ldots, t_n\{x := u\}) & \end{array}$$

On notera que la substitution est inopérante dans un terme t qui n'a pas d'occurrence libre de la variable x, c'est-à-dire : $t\{x := u\} = t$ si $x \notin FV(t)$.

En règle générale les substitutions ne commutent pas entre elles, en ce sens que les termes $t\{x := u\}\{y := v\}$ et $t\{y := v\}\{x := u\}$ peuvent être différents. En revanche, on démontre aisément le résultat suivant :

Lemme 1 (Lemme de substitution) — *Pour tous termes t, u, v et pour toutes variables x, y telles que $x \neq y$ et $x \notin FV(v)$, on a*

$$t\{x := u\}\{y := v\} \;=\; t\{y := v\}\{x := u\{y := v\}\}\,.$$

Preuve. Par récurrence sur la structure du terme t. □

La propriété de commutation $t\{x := u\}\{y := v\} = t\{y := v\}\{x := u\}$ est donc réalisée dans le cas particulier où $x \notin FV(v)$ et $y \notin FV(u)$.

1.3 Les formules

1.3.1 Définition

Soit Σ une signature quelconque.

Définition 3 (Formule) — *On appelle* formule *(ou Σ-formule) toute expression construite par application finie des règles suivantes :*

- *si p est un symbole de prédicat d'arité n (dans Σ) et si $t_1, \ldots, t_n$ sont n termes, alors l'expression $p(t_1, \ldots, t_n)$ est une formule ;*
- *les expressions $\top$ (« vrai ») et $\bot$ (« faux ») sont des formules ;*
- *si A est une formule, alors l'expression $\neg A$ (« non A ») est une formule ;*
- *si A et B sont des formules, alors les expressions $A \wedge B$ (« A et B »), $A \vee B$ (« A ou B »), $A \Rightarrow B$ (« A implique B ») et $A \Leftrightarrow B$ (« A si et seulement si B ») sont des formules ;*
- *si x est une variable et B une formule, alors les expressions $\forall x\, A$ (« pour tout x, A ») et $\exists x\, A$ (« il existe x tel que A ») sont des formules.*

On appelle *formule atomique* toute formule qui est soit l'une des deux unités $\top$ et $\bot$, soit une formule de la forme $p(t_1, \ldots, t_n)$, où p est un symbole de prédicat d'arité n et où $t_1, \ldots, t_n$ sont n termes. Plus généralement, on parle de négation (resp. de disjonction, de conjonction, d'implication, d'équivalence logique, de formule universelle, de formule existentielle) pour désigner toute formule de la forme $\neg A$ (resp. $A \wedge B$, $A \vee B$, $A \Rightarrow B$, $A \Leftrightarrow B$, $\forall x\, A$, $\exists x\, A$).

1.3.2 Variables libres, variables liées

La gestion des variables est plus complexe dans les formules que dans les termes en raison de la présence des quantificateurs $\forall$ et $\exists$ qui rendent muette la variable sur laquelle ils portent dans la formule (universelle ou existentielle) qu'ils construisent. Pour étudier ce phénomène, il est nécessaire de travailler non pas au niveau de la variable, mais au niveau de ses occurrences dans la formule, c'est-à-dire au niveau des positions où cette variable apparaît dans la formule considérée. En pratique, on distinguera deux types d'occurrences d'une même variable x dans une formule A, à savoir :

- Les occurrences de x qui figurent dans le scope d'une quantification de la forme $\forall x$ ou $\exists x$, et qu'on appelle des *occurrences liées* de x ;
- Les occurrences de x qui ne figurent dans le scope d'aucune quantification de la forme $\forall x$ ou $\exists x$, et qu'on appelle des *occurrences libres* de x.[4]

On fera attention au fait qu'une même variable peut avoir simultanément des occurrences libres et liées dans une même formule, comme par exemple la variable x dans la formule $p(x) \wedge \exists x\, q(x)$.

Formellement, les notions d'occurrence libre et d'occurrence liée d'une variable x dans une formule A sont définies récursivement sur la structure de A de la manière suivante :

- Toutes les occurrences de la variable x dans l'un des termes $t_1, \ldots, t_n$ sont des occurrences libres dans la formule $p(t_1, \ldots, t_n)$. Aucune variable n'a d'occurrence liée dans la formule $p(t_1, \ldots, t_n)$.
- Aucune variable n'a d'occurrence libre ou liée dans les formules $\top$ et $\bot$.
- Toutes les occurrences libres (resp. liées) de la variable x dans la formule A sont libres (resp. liées) dans la formule $\neg A$.
- Toutes les occurrences libres (resp. liés) de la variable x dans la formule A ou dans la formule B sont libres (resp. liées) dans les formules $A \vee B$, $A \wedge B$, $A \Rightarrow B$ et $A \Leftrightarrow B$.
- Toutes les occurrences libres et liées de la variable x dans la formule A sont liées dans les formules $\forall x\, A$ et $\exists x\, A$. En revanche, toutes les occurrences libres (resp. liées) d'une variable y distincte de x dans la formule A restent libres (resp. liées) dans les formules $\forall x\, A$ et $\exists x\, A$.

La distinction des deux formes d'occurrence d'une même variable permet de définir l'ensemble des variables libres d'une formule A (noté $FV(A)$) comme l'ensemble de toutes les variables qui ont au moins une occurrence libre dans A. De manière équivalente, l'ensemble $FV(A)$ des variables libres de A peut être défini par récurrence sur la structure de A à l'aide des équations suivantes :

$$\begin{array}{rcll}
FV(p(t_1, \ldots, t_n)) & = & FV(t_1) \cup \cdots \cup FV(t_n) & \\
FV(U) & = & \varnothing & (U \in \{\top; \bot\}) \\
FV(\neg A) & = & FV(A) & \\
FV(A \diamond B) & = & FV(A) \cup FV(B) & (\diamond \in \{\wedge; \vee; \Rightarrow; \Leftrightarrow\}) \\
FV(Qx\, A) & = & FV(A) \setminus \{x\} & (Q \in \{\forall; \exists\})
\end{array}$$

On dit qu'une formule A est *close* lorsque $FV(A) = \varnothing$. Dans le cas contraire, on dit que A est une formule *ouverte*. Étant donnée une formule A dont

[4] En toute rigueur, on devrait également distinguer une troisième forme d'occurrence de la variable x dans une formule, à savoir l'occurrence de x qui assure la liaison dans la quantification $\forall x$ ou $\exists x$, et que l'on peut qualifier d'*occurrence liante*. En pratique, on ignore systématiquement cette troisième forme d'occurrence dans laquelle la variable x n'apparaît pas en position de sous-terme dans la formule considérée.

l'ensemble des variables libres est donné par $FV(A) = \{x_1; \ldots; x_n\}$, on appelle *clôture universelle* de A la formule close

$$\forall x_1 \cdots \forall x_n\, A$$

obtenue en quantifiant universellement la formule A par rapport à toutes ses variables libres. La clôture universelle d'une formule A est définie à une permutation près des variables $x_1, \ldots, x_n$; en toute rigueur, on devrait donc parler d'*une* clôture universelle de la formule A plutôt que de *la* clôture universelle de A. Cependant, on verra au chapitre 2 que les différentes façons de clore universellement une même formule donnent lieu à des formules logiquement équivalentes. (De manière analogue on définit la notion de clôture existentielle, qui est en pratique beaucoup moins utilisée.)

Alpha-conversion et renommage Dans une formule universelle ou existentielle, on dit que la variable sur laquelle porte la quantification est muette car il est possible de changer son nom sans changer le sens de la formule considérée. Ainsi la formule $\exists x\ (x > 3)$ a-t-elle exactement le même sens que la formule $\exists y\ (y > 3)$, tandis que les formules $x > 3$ et $y > 3$ ont des sens différents (en supposant évidemment que x et y sont des variables distinctes).

On appelle *α-conversion* et on note $A =_\alpha A'$ la relation (d'équivalence) qui exprime que deux formules A et A' sont identiques aux noms de variables liées près. La définition formelle de cette relation est délicate, car un changement de nom malencontreux peut complètement changer le sens de la formule. Par exemple, la formule $\forall x\ \exists y\ p(x, y)$ est équivalente (au sens de l'α-conversion) aux formules $\forall x\ \exists z\ p(x, z)$ et $\forall z\ \exists y\ p(z, y)$, mais pas à la formule $\forall z\ \exists z\ p(z, z)$. En revanche, elle est équivalente à la formule $\forall y\ \exists x\ p(y, x)$. Dans la sous-section qui suit, on s'en tiendra à cette définition intuitive de la relation d'α-conversion, dont la définition formelle ne sera donnée qu'à la sous-section 1.3.4.

1.3.3 Substitution

Comme pour les termes, il est possible de définir une opération de substitution dans les formules, notée $A\{x := u\}$, qui consiste à remplacer dans A toutes les occurrences libres de la variable x par le terme u. La définition de cette opération est délicate car des liaisons indésirables peuvent être créées par une substitution maladroite : c'est le problème de la capture de variable.

Ce problème se comprend facilement à l'aide d'un exemple. Soit la formule $A \equiv \exists y\ (x = 2y)$ qui dans l'arithmétique de Peano exprime que « x est pair ». Normalement la formule $A\{x := 5y\}$ (obtenue en remplaçant dans A la variable x par le terme $5y$) devrait exprimer que « $5y$ est pair ». Or la formule $\exists y\ (5y = 2y)$ obtenue en effectuant naïvement la substitution exprime à l'évidence quelque chose de très différent. Le problème vient ici du fait que la variable libre y dans le terme $5y$ est « capturée » (i.e. devient liée) par la quantification $\forall y$ au cours du remplacement de la variable x par le terme $5y$ dans la formule A. Pour éviter cette capture, il est nécessaire de renommer dans A la variable y responsable de la liaison fautive à l'aide d'une variable fraîche z[5]. On obtient de la sorte une formule $A' \equiv \exists z\ (x = 2z)$ où la variable y a disparu, mais qui

[5]C'est-à-dire une variable n'apparaissant nulle part dans les termes et formules qui posent problème. Ici, il suffit de prendre n'importe quelle variable distincte de x et de y.

reste équivalente à A au sens de l'α-conversion. Une fois le renommage effectué, la substitution s'effectue sans danger sur la formule A' et a pour résultat la formule $\exists z\ (5y = 2z)$ qui a effectivement le sens recherché (« $5y$ est pair »).

L'exemple ci-dessus montre l'importance de l'α-conversion et du renommage dans la définition de la substitution dans les formules. Aussi pour définir cette notion on procède en deux temps. D'abord on définit une substitution simple notée $A(x := u)$, sans renommage, mais qui est une opération partielle dont le résultat est indéfini dès qu'il y a un risque de capture. Dans un second temps on définit la substitution générale $A\{x := u\}$ à partir de la substitution simple, en faisant précéder cette dernière d'une phase de renommage pour anticiper toutes les captures possibles (c'est-à-dire en renommant toutes les variables liées de A par des variables fraîches, du moins distinctes de toutes les variables libres du terme u.) Contrairement à la substitution simple, la substitution générale (qui est une opération totale) n'est pas définie sur les formules, mais sur les classes d'équivalence de formules pour la relation d'α-conversion.

Formellement, la substitution simple $A(x := u)$ est définie par récurrence sur la structure de A à l'aide des équations suivantes :

$$\begin{array}{rcll}
p(t_1,\ldots,t_n)(x := u) & = & p(t_1(x := u),\ldots,t_n(x := u)) & \\
U(x := u) & = & U & (U \in \{\top;\bot\}) \\
(\neg A)(x := u) & = & \neg(A(x := u)) & \\
(A \diamond B)(x := u) & = & A(x := u) \diamond B(x := U) & (\diamond \in \{\wedge;\vee;\Rightarrow;\Leftrightarrow\}) \\
(Qx\,A)(x := u) & = & Qx\,A & (Q \in \{\forall;\exists\}) \\
(Qy\,A)(x := u) & = & Qy\,(A(x := u)) & (\text{si } y \neq x \text{ et } y \notin FV(u))
\end{array}$$

On notera que la notation $(Qy\,A)(x := u)$ (pour $Q \in \{\forall;\exists\}$) n'est pas définie dans le cas où $y \in FV(u)$, c'est-à-dire précisément dans le cas où la quantification Qy est en mesure de « capturer » la variable y dans le terme u.

On démontre alors aisément le lemme de substitution suivant, dont la preuve s'effectue par récurrence structurelle sur A :

Lemme 2 (Lemme de substitution) — *Pour toute formule A, pour tous termes u, v et pour toutes variables x, y telles que $x \neq y$ et $x \notin FV(v)$, on a :*

$$A(x := u)(y := v) \ = \ A(y := v)(x := u(y := v))$$

(dès que ces deux formules sont définies).

Pour définir la substitution générale, nous devons maintenant nous attaquer au problème de la définition formelle de l'α-conversion.

1.3.4 Définition de l'α-conversion

L'opération de substitution simple $A(x := u)$ définie à la sous-section précédente permet maintenant de définir formellement la relation d'α-conversion $A =_\alpha A'$ à partir d'un ensemble de règles d'inférence qui est le suivant :

$$\frac{}{p(t_1,\ldots,t_n) \ =_\alpha \ p(t_1,\ldots,t_n)} \qquad \frac{}{U \ =_\alpha \ U} \qquad (U \in \{\top;\bot\})$$

$$\frac{A =_\alpha A'}{\neg A \ =_\alpha \ \neg A'} \qquad \frac{A =_\alpha A' \qquad B =_\alpha B'}{A \diamond B \ =_\alpha \ A' \diamond A'} \qquad (\diamond \in \{\wedge;\vee;\Rightarrow;\Leftrightarrow\})$$

$$\frac{A(x := z) \ =_\alpha \ A'(y := z)}{Qx\,A \ =_\alpha \ Qy\,A'} \qquad (z \text{ var. fraîche}, \quad Q \in \{\forall;\exists\})$$

Les règles ci-dessus sont dirigées par la syntaxe et donnent immédiatement un algorithme pour tester si deux formules A et A' sont α-convertibles.

Elles expriment en effet que deux formules A et A', pour satisfaire au test d'α-conversion, doivent d'abord avoir la même construction de tête (le même symbole de prédicat, le même connecteur ou le même quantificateur). Dans le cas où A et A' sont des formules atomiques, le test d'α-conversion se ramène à une égalité ordinaire. Dans le cas où A et A' sont construites à l'aide du même connecteur, le test d'α-conversion s'effectue membre à membre, en appelant récursivement la procédure de test. Enfin, dans le cas où A et A' sont des formules quantifiées de la forme $A = Qx\, A_0$ et $A' = Qy\, A'_0$, le test d'α-conversion s'effectue en choisissant une variable fraîche z quelconque, et en appelant récursivement la procédure de test sur les formules $A_0(x := z)$ et $A'_0(x := z)$ construites par une opération de substitution simple. Évidemment, les formules $A_0(x := z)$ et $A'_0(y := z)$ qui résultent de la substitution simple sont ici toujours définies en raison du fait qu'une variable fraîche z ne risque pas d'être capturée par une quantification dans A_0 ou dans A'_0.

On vérifie alors que :

1. La relation $A =_\alpha A'$ définie ci-dessus est une relation d'équivalence.
2. Si $A =_\alpha A'$ et si les substitutions simples $A(x := u)$ et $A'(x := u)$ sont toutes les deux définies, alors $A(x := u) =_\alpha A'(x := u)$.
3. Étant donnée une formule A, une variable x et un terme u, il existe une formule $A' =_\alpha A$ telle que la substitution simple $A'(x := u)$ est définie.

Les points 2 et 3 permettent de définir la substitution dans les formules sous sa forme générale en posant

$$A\{x := u\} \;=\; A'(x := u)\,,$$

où A' désigne n'importe quelle formule α-convertible à A pour laquelle la substitution simple $A'(x := u)$ est définie (et dont l'existence est assurée par le 3e point ci-dessus). Bien entendu, la formule $A\{x := u\}$ est définie à α-conversion près : si la formule $A'(x := u)$ dépend évidemment du choix de A', sa classe d'équivalence n'en dépend pas (d'après le 2e point).

Il est facile de déduire du lemme 2 un lemme de substitution pour la substitution généralisée où les deux membres de l'égalité (exprimée à α-conversion près) sont cette fois-ci toujours définis :

Lemme 3 (Lemme de substitution) — *Pour toute formule A, pour tous termes u, v et pour toutes variables x, y telles que $x \neq y$ et $x \notin FV(v)$, on a :*

$$A\{x := u\}\{y := v\} \;=_\alpha\; A\{y := v\}\{x := u\{y := v\}\}\,.$$

Dans la suite de ce cours, on travaillera toujours à α-équivalence près, en renommant silencieusement les variables liées en cas de nécessité.

Chapitre 2

La déduction naturelle

Ce chapitre est consacré à la définition du concept de *démonstration* (ou *preuve*) en logique du premier ordre. Nous adoptons ici une présentation basée sur la déduction naturelle[1] — un système de déduction issu notamment des travaux de Gentzen (1909–1945) et de Prawitz (1936→) — étendue avec une règle de raisonnement par l'absurde pour recouvrer toute l'expressivité de la logique classique.

2.1 Preuves en déduction naturelle

Un système de déduction — qu'il s'agisse de la déduction naturelle, du calcul des séquents ou de tout autre système de déduction — sert à organiser le raisonnement logique afin de distinguer parmi les formules celles qui sont logiquement *valides*, c'est-à-dire vraies indépendamment de l'interprétation que l'on peut donner aux symboles de fonction et de prédicat à partir desquels ces formules sont construites. Pour une formule donnée, cette distinction s'effectue en construisant une *dérivation* (ou *démonstration*, ou *preuve*) — c'est-à-dire essentiellement un enchaînement fini d'inférences logiques aboutissant à la formule considérée — qui constitue littéralement un « certificat de validité » pour la formule ainsi démontrée.

L'intérêt d'une telle dérivation est non seulement qu'elle est finie (et généralement de taille raisonnable, au moins pour les mémoires des ordinateurs), mais surtout que sa correction formelle est vérifiable en temps fini à l'aide d'une procédure finie.[2] Par comparaison, le calcul de la valeur de vérité d'une formule, dont la complexité exponentielle est déjà prohibitive dans le simple calcul propositionnel (c'est-à-dire en l'absence de quantificateurs), n'est pas réalisable en temps fini dès qu'on aborde le calcul des prédicats. (Cependant nous verrons au chapitre 4 qu'il est possible de *définir* mathématiquement la valeur de vérité

[1] Plus précisément : sur la déduction naturelle avec des hypothèses explicites. Dans sa présentation originelle, la déduction naturelle ne manipule que des formules (les « buts »), les hypothèses étant manipulées au niveau des arbres de dérivation à l'aide d'un mécanisme complexe d'introduction et d'effacement.

[2] Les algorithmes de *vérification de preuve*, qui permettent de décider si une structure arborescente finie d constitue une dérivation d'une formule A ou non, sont à la base des logiciels d'aide à la démonstration tels que Coq, Isabelle/HOL, PVS, PhoX, etc.

de n'importe quelle formule du calcul des prédicats, même s'il n'est en général pas possible de la *calculer* effectivement.)

Les premiers systèmes de déduction (dûs à Hilbert, 1862–1943) définissaient les démonstrations comme des suites finies de formules liées par des inférences logiques, suites qui commençant généralement par l'énoncé d'un certain nombre d'axiomes de la théorie pour finalement s'achever sur la formule à démontrer. Cette structure, dont le seul avantage était de rappeler l'écriture linéaire des démonstrations qui prévaut dans les mathématiques informelles, montra vite ses limites dès qu'on chercha à étudier ses propriétés, sinon à l'utiliser.

Les progrès considérables de la théorie de la démonstration à partir de la fin des années 1930 apportèrent deux modifications essentielles à la structure initiale des démonstrations. La première (déjà envisagée par certains auteurs du temps de Hilbert) fut d'écrire les démonstrations sous forme arborescente et non sous forme linéaire, de manière à regrouper les différentes formules qui participent d'une même inférence logique autour d'un même nœud de l'arbre. (Dans une démonstration linéaire, les prémisses d'une inférence logique peuvent être arbitrairement éloignées de la conclusion.) Mais la modification essentielle fut sans nul doute le remplacement par Gentzen (1909–1945) des formules par des *séquents* dans les nœuds des arbres de démonstration, introduisant ainsi l'une des notions clés de la théorie de la démonstration moderne.

2.1.1 La notion de séquent

Un *séquent* est un couple noté

$$\Gamma \vdash \Delta \qquad\qquad (\text{« } \Gamma \text{ thèse } \Delta \text{ »})$$

où Γ et Δ sont des listes finies de formules (ou éventuellement des multi-ensembles finis de formules). La liste Γ figurant à gauche du signe $\vdash$ (« thèse ») se nomme le *subséquent* tandis que la liste Δ figurant à droite se nomme le *conséquent*. Intuitivement, un séquent $\Gamma \vdash \Delta$ exprime le jugement selon lequel

> *Si toutes les formules du subséquent Γ sont vraies, alors l'une au moins des formules du conséquent Δ est vraie.*[3]

D'un point de vue logique, le symbole $\vdash$ (« thèse ») a donc une signification très proche de l'implication, et le séquent

$$A_1, \dots, A_n \vdash B_1, \dots, B_k$$

exprime sensiblement la même chose que la formule

$$(A_1 \wedge \cdots \wedge A_n) \Rightarrow (B_1 \vee \cdots \vee B_k)\,.$$

D'un point de vue technique cependant, la notion de séquent facilite considérablement la déstructuration de la formule à démontrer en introduisant un « curseur » (le signe $\vdash$) séparant un ensemble d'hypothèses (toutes placées au même niveau) d'un ensemble de buts (tous également au même niveau) dont l'un au moins est à démontrer sous ces hypothèses.

On distingue plusieurs formes de séquents, notamment :

[3] La subtilité résidant dans le fait qu'en général, on ne sait pas laquelle des formules du conséquent est vraie, comme par exemple dans le séquent dérivable $\vdash A, \neg A$ (ici A désigne n'importe quelle formule ; par exemple la conjecture de Goldbach).

– Les séquents droits, qui sont de la forme $\vdash \Delta$ (c'est-à-dire les séquents dont le subséquent est vide). Un séquent droit $\vdash \Delta$ exprime que l'une au moins des formules de Δ est vraie (même si en général on ne sait pas laquelle). Un cas particulier de séquent de cette forme est le séquent $\vdash A$ qui exprime simplement que la formule A est valide.
– Les séquents gauches, qui sont de la forme $\Gamma \vdash$ (c'est-à-dire les séquents dont le conséquent est vide). Un séquent gauche $\Gamma \vdash$ exprime que l'ensemble des hypothèses dans Γ est contradictoire. Un cas très simple de séquent de cette forme est le séquent vide $\vdash$ (sans subséquent ni conséquent), qui exprime la contradiction (et n'est donc jamais dérivable).

Dans ce chapitre consacré à la déduction naturelle, nous n'utiliserons qu'une forme particulière de séquent qui est la suivante :

Définition 4 (Séquent intuitionniste) — *On appelle* séquent intuitionniste *tout séquent dont le conséquent comporte exactement une seule formule, c'est-à-dire tout séquent de la forme* $\Gamma \vdash A$.

Un séquent intuitionniste $\Gamma \vdash A$ exprime donc que « la formule A est vraie sous les hypothèses Γ ».

Notations sur les listes de formules Dans le cadre de ce cours, les listes de formules sont notées par simple juxtaposition de leurs éléments séparés par des virgules, i.e. $A_1, \dots, A_n$. La liste vide est notée $\emptyset$ et est même omise en position de subséquent et/ou de conséquent.

La concaténation de deux listes Γ et Δ (dans cet ordre) est notée Γ, Δ. Les notations $FV(\Gamma)$ (ensemble des variables libres) et $\Gamma\{x := u\}$ (substitution) sont étendues aux listes de formules en posant

$$\begin{array}{rcl} FV(A_1, \dots, A_n) & = & FV(A_1) \cup \dots \cup FV(A_1) \\ (A_1, \dots, A_n)\{x := u\} & = & A_1\{x := u\}, \dots, A_n\{x := u\} \end{array}$$

2.1.2 La notion de règle d'inférence

Comme n'importe quel système de déduction, la déduction naturelle est définie par un ensemble de *règles d'inférence* qui sont les briques à partir desquelles on construit les dérivations. Chaque règle d'inférence est de la forme

$$\frac{P_1 \quad \cdots \quad P_n}{C}$$

où $P_1, \dots, P_n$ sont des séquents appelés les *prémisses* de la règle tandis que C est un séquent appelé *conclusion*. Certaines règles comportent en outre une condition de bord (généralement indiquée à droite du trait d'inférence) qui en restreignent l'usage. Chaque règle peut se lire dans deux directions :

De haut en bas (sens de la causalité) Si chacun des séquents $P_1, \dots, P_n$ est dérivable, alors le séquent C est dérivable également.

De bas en haut (sens de la recherche de preuve) Une méthode possible pour dériver C consiste à dériver chacun des séquents $P_1, \dots, P_n$.

2.1.3 Règles d'inférence de la déduction naturelle

Les règles d'inférence de la déduction naturelle classique, données à la figure 2.1, se répartissent en quatre groupes :

La règle axiome C'est la règle qui permet d'utiliser les hypothèses dans un raisonnement, c'est-à-dire de dériver n'importe quel séquent $\Gamma \vdash A$ dont le conséquent A figure parmi les hypothèses Γ. Avec la règle $\top$-intro, c'est la seule règle du système qui comporte zéro prémisse.

Les règles d'introduction Ce sont les règles qui décrivent le fonctionnement des connecteurs à travers leurs causes. On les appelle des règles d'introduction car elles font apparaître un connecteur quand on les lit dans le sens de la causalité (c'est-à-dire de haut en bas). Dans le sens de la recherche de preuve au contraire (c'est-à-dire de bas en haut), on les utilise pour simplifier le but en faisant disparaître le connecteur de tête.

Les règles d'élimination Ce sont les règles qui décrivent le fonctionnement des connecteurs à travers leurs conséquences. On les appelle des règles d'élimination car elles font disparaître un connecteur quand on les lit dans le sens de la causalité (c'est-à-dire de haut en bas). Dans le sens de la recherche de preuve au contraire (c'est-à-dire de bas en haut), elles font apparaître un connecteur et sont pour cette raison souvent d'un usage plus délicat que les règles d'introduction.

La règle de raisonnement par l'absurde C'est la règle de déduction qui permet de dériver une formule A à partir d'une dérivation que sa négation $\neg A$ est contradictoire. (Cette règle ne doit pas être confondue avec la règle d'introduction de la négation, qui permet de dériver la formule $\neg A$ à partir d'une dérivation que la formule A est contradictoire.) La règle de raisonnement par l'absurde a un statut particulier en déduction naturelle puisque c'est elle qui permet de recouvrer toute l'expressivité de la logique classique, les autres règles ne permettant de dériver que des tautologies intuitionnistes (cf fin de la sous-section 2.1.4).

2.1.4 Dérivations

Définition 5 (Dérivation) — *On appelle* dérivation *(ou* preuve, *ou* démonstration*) tout arbre fini dont les nœuds sont étiquetés par des séquents, et qui est construit par application finie de la règle de construction suivante :*

– *Si $d_1, \ldots, d_n$ sont des dérivations des séquents $P_1, \ldots, P_n$*

$$\begin{array}{ccccc} \vdots\, d_1 & \quad & \vdots\, d_2 & \quad \cdots \quad & \vdots\, d_n \\ P_1 & & P_2 & & P_n \end{array}$$

et si C est un séquent tel que l'inférence $\frac{P_1 \cdots P_n}{C}$ est une instance d'une des règles d'inférence de la déduction naturelle classique, alors l'arbre fini

$$d = \left\{ \begin{array}{c} \begin{array}{ccc} \vdots\, d_1 & & \vdots\, d_n \\ P_1 & \cdots & P_n \end{array} \\ \hline C \end{array} \right.$$

obtenu en connectant par un trait d'inférence les dérivations $d_1, \ldots, d_n$ au séquent C constitue une dérivation du séquent C.

(Axiome)	$\overline{\Gamma \vdash A}$	$(A \in \Gamma)$
($\top$-intro)	$\overline{\Gamma \vdash \top}$	
($\bot$-élim)	$\dfrac{\Gamma \vdash \bot}{\Gamma \vdash A}$	
($\neg$-intro)	$\dfrac{\Gamma, A \vdash \bot}{\Gamma \vdash \neg A}$	
($\neg$-élim)	$\dfrac{\Gamma \vdash A \qquad \Gamma \vdash \neg A}{\Gamma \vdash \bot}$	
($\wedge$-intro)	$\dfrac{\Gamma \vdash A \qquad \Gamma \vdash B}{\Gamma \vdash A \wedge B}$	
($\wedge$-élim$_{1,2}$)	$\dfrac{\Gamma \vdash A \wedge B}{\Gamma \vdash A} \qquad \dfrac{\Gamma \vdash A \wedge B}{\Gamma \vdash B}$	
($\vee$-intro$_{1,2}$)	$\dfrac{\Gamma \vdash A}{\Gamma \vdash A \vee B} \qquad \dfrac{\Gamma \vdash B}{\Gamma \vdash A \vee B}$	
($\vee$-élim)	$\dfrac{\Gamma \vdash A \vee B \qquad \Gamma, A \vdash C \qquad \Gamma, B \vdash C}{\Gamma \vdash C}$	
($\Rightarrow$-intro)	$\dfrac{\Gamma, A \vdash B}{\Gamma \vdash A \Rightarrow B}$	
($\Rightarrow$-élim)	$\dfrac{\Gamma \vdash A \Rightarrow B \qquad \Gamma \vdash A}{\Gamma \vdash B}$	
($\Leftrightarrow$-intro)	$\dfrac{\Gamma, A \vdash B \qquad \Gamma, B \vdash A}{\Gamma \vdash A \Leftrightarrow B}$	
($\Leftrightarrow$-élim$_{1,2}$)	$\dfrac{\Gamma \vdash A \Leftrightarrow B \qquad \Gamma \vdash A}{\Gamma \vdash B} \qquad \dfrac{\Gamma \vdash A \Leftrightarrow B \qquad \Gamma \vdash B}{\Gamma \vdash A}$	
($\forall$-intro)	$\dfrac{\Gamma \vdash A}{\Gamma \vdash \forall x\, A}$	$(x \notin FV(\Gamma))$
($\forall$-élim)	$\dfrac{\Gamma \vdash \forall x\, A}{\Gamma \vdash A\{x := t\}}$	
($\exists$-intro)	$\dfrac{\Gamma \vdash A\{x := t\}}{\Gamma \vdash \exists x\, A}$	
($\exists$-élim)	$\dfrac{\Gamma \vdash \exists x\, A \qquad \Gamma, A \vdash B}{\Gamma \vdash B}$	$(x \notin FV(\Gamma, B))$
(Absurde)	$\dfrac{\Gamma, \neg A \vdash \bot}{\Gamma \vdash A}$	

FIG. 2.1 – Règles d'inférence de la déduction naturelle

Lorsqu'un séquent S admet une dérivation d (c'est-à-dire une dérivation dont la racine est étiquetée par S), on dit que le séquent S est *dérivable*. Par exemple, le séquent $\forall x\,(p(x) \Rightarrow q(x)),\ p(3) \vdash q(3)$ est dérivable et l'arbre ci-dessous en donne une dérivation :

$$\dfrac{\dfrac{\dfrac{}{\forall x\,(p(x) \Rightarrow q(x)),\ p(3)\ \vdash\ \forall x\,(p(x) \Rightarrow q(x))}}{\forall x\,(p(x) \Rightarrow q(x)),\ p(3)\ \vdash\ p(3) \Rightarrow q(3)} \qquad \dfrac{}{\forall x\,(p(x) \Rightarrow q(x)),\ p(3)\ \vdash\ p(3)}}{\forall x\,(p(x) \Rightarrow q(x)),\ p(3)\ \vdash\ q(3)}$$

(Les règles utilisées sont, de haut en bas et de gauche à droite, les règles axiome, $\forall$-élim, axiome, $\Rightarrow$-élim) Une formule A telle que le séquent $\vdash A$ est dérivable est qualifié de *tautologie*.

Dérivabilité intuitionniste Les notions de preuve et de dérivabilité définies ci-dessus sont celles de la logique classique, qui est la logique couramment utilisée en mathématiques et qui correspond à la notion de valeur de vérité booléenne. Les notions correspondantes en logique intuitionniste (ou logique constructive) s'obtiennent simplement en prohibant l'usage de la règle de raisonnement par l'absurde (cf Fig. 2.1). Ainsi, une dérivation est intuitionniste lorsqu'elle ne comporte aucune instance de la règle de raisonnement par l'absurde, et une tautologie est intuitionniste dès qu'elle admet une dérivation intuitionniste.

La logique intuitionniste (introduite par Brouwer, 1881–1966) rejette certains principes de raisonnements tels que le tiers-exclu $A \vee \neg A$ ou la loi de Peirce $((A \Rightarrow B) \Rightarrow A) \Rightarrow A$. En logique intuitionniste, une formule A n'est en général pas équivalente à sa double négation $\neg\neg A$: si l'implication $A \Rightarrow \neg\neg A$ reste une tautologie au sens intuitionniste, ce n'est pas le cas de sa réciproque $\neg\neg A \Rightarrow A$ (élimination de la double négation) qui n'est dérivable qu'en logique classique. On notera d'ailleurs que la règle de raisonnement par l'absurde n'est qu'une formulation cachée du principe d'élimination de la double négation.

2.2 Propriétés

2.2.1 Règles structurelles

Soient Γ et Γ' deux contextes. On dit que Γ est *inclus* dans Γ' et on écrit $\Gamma \subset \Gamma'$ lorsque toute formule qui figure dans Γ figure également dans Γ'. Cette notion ne tient compte ni de l'ordre des formules ni du nombre de leurs occurrences dans les deux contextes.

Proposition 1 (Affaiblissement généralisé) — *Si le séquent $\Gamma \vdash A$ est dérivable en logique classique (resp. en logique intuitionniste), alors pour tout contexte Γ' tel que $\Gamma \subset \Gamma'$ le séquent $\Gamma' \vdash A$ est dérivable en logique classique (resp. en logique intuitionniste).*

Preuve. Par récurrence sur la preuve du séquent $\Gamma \vdash A$. □

Cette proposition a trois conséquences importantes.

La première est que la dérivabilité (en logique classique comme en logique intuitionniste) est conservée par l'ajout d'hypothèses. Autrement dit, la règle d'inférence suivante

(Affaiblissement) $$\dfrac{\Gamma \vdash A}{\Gamma, B \vdash A}$$

est *admissible*, en ce sens que l'ajout de cette règle d'inférence aux règles de la figure 2.1 ne change pas la notion de dérivabilité, et cela quelle que soit la logique (classique ou intuitionniste) considérée. Dans ce qui suit, on distinguera les règles d'inférence admissibles des règles d'inférence primitives en les signalant à l'aide d'un double trait d'inférence.

La deuxième conséquence de la proposition précédente est que la dérivabilité est invariante par permutation d'hypothèses, ce que l'on peut formuler en disant que la règle d'inférence suivante est admissible

(Permutation) $$\frac{B_1, \ldots, B_n \vdash A}{B_{\sigma(1)}, \ldots, B_{\sigma(n)} \vdash A}$$

pour tout $n \geq 0$ et pour toute permutation σ des entiers $1..n$.

Enfin, la troisième conséquence de la règle d'affaiblissement généralisé est que la dérivabilité ne dépend pas du nombre d'occurrences d'une même formule dans le contexte. Autrement dit, la règle suivante est admissible :

(Contraction) $$\frac{\Gamma, B, B \vdash A}{\Gamma, B \vdash A}$$

La vocation des variables, qui est de représenter des termes quelconques, apparaît dans le résultat suivant qui exprime que la dérivabilité est préservée lorsqu'on substitue un terme arbitraire à une variable :

Proposition 2 (Substitutivité) — *Si le séquent $\Gamma \vdash A$ est dérivable en logique classique (resp. en logique intuitionniste), alors pour toute variable x et pour tout terme t le séquent $\Gamma\{x := t\} \vdash A\{x := t\}$ est dérivable en logique classique (resp. en logique intuitionniste).*

Preuve. Par récurrence sur la structure de la dérivation $\Gamma \vdash A$. □

2.2.2 Règles inversibles

Les connecteurs $\wedge$, $\Rightarrow$ et le quantificateur $\forall$ sont *inversibles* en ce sens que leurs règles d'introduction peuvent se lire à l'envers, en déduisant chacune des prémisses à partir de la conclusion. Autrement dit :

Proposition 3 (Inversion) — *Les règles d'inférence*

$$\frac{\Gamma \vdash A \wedge B}{\Gamma \vdash A} \quad \frac{\Gamma \vdash A \wedge B}{\Gamma \vdash B} \qquad \frac{\Gamma \vdash A \Rightarrow B}{\Gamma, A \vdash B} \qquad \frac{\Gamma \vdash \forall x\, A}{\Gamma \vdash A}\ {\scriptstyle (x \notin FV(\Gamma))}$$

sont admissibles en logique classique comme en logique intuitionniste.

Preuve. L'admissibilité des deux premières règles est triviale, puisqu'il s'agit des règles d'élimination de la conjonction. L'admissibilité de la règle d'inversion de l'implication se démontre à l'aide du fragment de dérivation

$$\dfrac{\dfrac{\dfrac{\vdots}{\Gamma \vdash A \Rightarrow B}}{\Gamma, A \vdash A \Rightarrow B}\,(\text{Aff.}) \qquad \dfrac{}{\Gamma, A \vdash A}\,(\text{Ax.})}{\Gamma, A \vdash B}\,(\Rightarrow\text{-élim})$$

en utilisant la règle admissible d'affaiblissement. Enfin, l'admissibilité de la règle d'inversion du quantificateur universel n'est qu'une instance particulière de sa règle d'élimination utilisée avec le terme $t \equiv x$. (On notera que la condition de bord n'est pas utilisée.) $\square$

Intuitivement, une règle inversible est une règle qui ne perd aucune information lors du passage des prémisses à la conclusion. Dans la perspective d'une recherche automatique de preuve, l'inversibilité des connecteurs $\wedge$ et $\Rightarrow$ et du quantificateur $\forall$ est primordiale : elle signifie que lorsqu'on cherche à construire une dérivation d'une formule de la forme $A \wedge B$, $A \Rightarrow B$ ou $\forall x\, A$, il est toujours possible de décomposer la formule en appliquant la règle d'introduction correspondante sans courir le risque de devoir revenir en arrière (pour essayer ensuite une autre règle) dans le cas où la preuve d'une des prémisses échouerait. L'inversibilité du connecteur (ou quantificateur) principal entraîne en effet que si la formule admet une dérivation (en logique classique comme en logique intuitionniste), alors elle admet nécessairement une dérivation qui se termine par la règle d'introduction associée. Cet argument est très important en pratique car il permet de réduire considérablement l'espace de recherche.

La situation est très différente pour la disjonction et le quantificateur existentiel dont les règles d'introduction ne sont pas inversibles. Intuitivement, la disjonction nécessite de faire un choix sur la règle d'introduction à utiliser (faut-il prouver A ou faut-il prouver B ?) tandis que le quantificateur existentiel nécessite de choisir un témoin approprié (pour quel t a-t-on $A(t)$?) En logique classique, le problème est d'autant plus grand qu'il existe des disjonctions prouvables (dans le contexte vide) dont aucun des deux membres ne l'est,[4] de même qu'il existe des formules existentielles prouvables pour lesquelles il n'est pas possible d'exhiber le moindre témoin.[5]

2.2.3 Décomposition des connecteurs $\neg$ et $\Leftrightarrow$

En déduction naturelle, il est commode de considérer les connecteurs $\neg A$ (négation) et $\Leftrightarrow$ (équivalence logique) non pas comme des connecteurs primitifs mais comme des constructions définies par

$$\neg A \;\equiv\; A \Rightarrow \bot \qquad \text{et} \qquad A \Leftrightarrow B \;\equiv\; (A \Rightarrow B) \wedge (B \Rightarrow A)$$

Ce point de vue est justifié par le fait que les règles d'introduction et d'élimination de la négation et de l'équivalence logique peuvent elles-mêmes être décomposées à l'aide des règles d'introduction et d'élimination de l'implication et de la conjonction de la manière suivante :

$$\neg\text{-intro} \;=\; \left\{ \begin{array}{c} \vdots \\ \dfrac{\Gamma, A \vdash \bot}{\Gamma \vdash A \Rightarrow \bot}\ {\scriptstyle(\Rightarrow\text{-intro})} \\ \Gamma \vdash \neg A \end{array} \right.$$

$$\neg\text{-élim} \;=\; \left\{ \dfrac{\begin{array}{c} \vdots \\ \Gamma \vdash \neg A \\ \Gamma \vdash A \Rightarrow \bot \end{array} \qquad \begin{array}{c} \vdots \\ \Gamma \vdash A \end{array}}{\Gamma \vdash \bot}\ {\scriptstyle(\Rightarrow\text{-élim})} \right.$$

[4] Par exemple : « hypothèse du continu $\vee$ $\neg$ hypothèse du continu ».

[5] Par exemple : « $\exists x\, [(x = 1 \wedge \text{hypothèse du continu}) \vee (x = 0 \wedge \neg\text{hypothèse du continu})]$ ».

$$
\Leftrightarrow\text{-intro} \;=\; \left\{ \begin{array}{c}
\dfrac{\dfrac{\begin{array}{c}\vdots\\ \Gamma, A \vdash B\end{array}}{\Gamma \vdash A \Rightarrow B}\,(\Rightarrow\text{-intro}) \qquad \dfrac{\begin{array}{c}\vdots\\ \Gamma, B \vdash A\end{array}}{\Gamma \vdash B \Rightarrow A}\,(\Rightarrow\text{-intro})}{\Gamma \vdash (A \Rightarrow B) \wedge (B \Rightarrow A)}\,(\wedge\text{-intro}) \\
\Gamma \vdash A \Leftrightarrow B
\end{array} \right.
$$

$$
\Leftrightarrow\text{-élim}_1 \;=\; \left\{ \dfrac{\dfrac{\begin{array}{c}\vdots\\ \Gamma \vdash A \Leftrightarrow B\\ \Gamma \vdash (A \Rightarrow B) \wedge (B \Rightarrow A)\end{array}}{\Gamma \vdash A \Rightarrow B}\,(\wedge\text{-élim}_1) \qquad \begin{array}{c}\vdots\\ \Gamma \vdash A\end{array}}{\Gamma \vdash B}\,(\Rightarrow\text{-élim}) \right.
$$

$$
\Leftrightarrow\text{-élim}_2 \;=\; \left\{ \dfrac{\dfrac{\begin{array}{c}\vdots\\ \Gamma \vdash A \Leftrightarrow B\\ \Gamma \vdash (A \Rightarrow B) \wedge (B \Rightarrow A)\end{array}}{\Gamma \vdash B \Rightarrow A}\,(\wedge\text{-élim}_2) \qquad \begin{array}{c}\vdots\\ \Gamma \vdash B\end{array}}{\Gamma \vdash A}\,(\Rightarrow\text{-élim}) \right.
$$

En d'autres termes, tout instance de l'une des règles $\neg$-intro, $\neg$-élim, $\Leftrightarrow$-intro, $\Leftrightarrow$-élim$_1$ ou $\Leftrightarrow$-élim$_2$ peut être remplacée dans n'importe quelle dérivation par le fragment de dérivation ci-dessus qui la simule. Aussi les connecteurs $\neg$ et $\Leftrightarrow$ ainsi que leurs règles de déduction sont-ils en réalité superflus.

2.2.4 Caractérisation des connecteurs par leurs règles

La décomposition des connecteurs $\neg$ et $\Leftrightarrow$ illustre une propriété plus générale de la déduction naturelle qui est que chacun des connecteurs et des quantificateurs de la logique du premier ordre est caractérisé par ses règles d'introduction et d'élimination dans ce système. Pour comprendre ce phénomène, prenons par exemple la conjonction et ajoutons au langage des formules un nouveau connecteur binaire $\heartsuit$ muni des mêmes règles d'introduction et d'élimination que la conjonction, à savoir :

$$
\frac{\Gamma \vdash A \qquad \Gamma \vdash B}{\Gamma \vdash A \heartsuit B} \qquad\qquad \frac{\Gamma \vdash A \heartsuit B}{\Gamma \vdash A} \qquad \frac{\Gamma \vdash A \heartsuit B}{\Gamma \vdash B}
$$

Ces règles étant posées, il est facile de démontrer dans le système de déduction étendu avec ces règles que la nouvelle construction $A \heartsuit B$ est logiquement équivalente à la conjonction $A \wedge B$, ainsi qu'en témoigne la dérivation

$$
\dfrac{\dfrac{\dfrac{\overline{A \wedge B \vdash A \wedge B}}{A \wedge B \vdash A} \qquad \dfrac{\overline{A \wedge B \vdash A \wedge B}}{A \wedge B \vdash B}}{A \wedge B \vdash A \heartsuit B} \qquad \dfrac{\dfrac{\overline{A \heartsuit B \vdash A \heartsuit B}}{A \heartsuit B \vdash A} \qquad \dfrac{\overline{A \heartsuit B \vdash A \heartsuit B}}{A \heartsuit B \vdash B}}{A \heartsuit B \vdash A \wedge B}}{\vdash (A \wedge B) \Leftrightarrow (A \heartsuit B)}
$$

La même propriété vaut pour toutes les autres constructions : les unités $\top$, $\bot$, la négation $\neg$, les connecteurs binaires $\vee$, $\Rightarrow$ et les quantificateurs $\forall$, $\exists$.

2.2.5 Formule associée à un séquent

À chaque séquent $A_1, \ldots, A_n \vdash B$ il est possible d'associer une formule $A_1 \wedge \cdots \wedge A_n \Rightarrow B$ (sans hypothèses), qui a exactement la même signification en termes de dérivabilité :

Proposition 4 — *En logique classique comme en logique intuitionniste, les assertions suivantes sont équivalentes :*

1. $A_1, \ldots, A_n \vdash B$ *est dérivable,*
2. $\vdash A_1 \wedge \cdots \wedge A_n \Rightarrow B$ *est dérivable,*
3. $\vdash A_1 \Rightarrow \cdots \Rightarrow A_n \Rightarrow B$ *est dérivable,*

où $A_1 \Rightarrow \cdots \Rightarrow A_n \Rightarrow B$ *désigne la formule* $A_1 \Rightarrow (\cdots \Rightarrow (A_n \Rightarrow B) \cdots)$.

Preuve. L'équivalence de 1 avec 3 se démontre par récurrence sur le nombre d'hypothèses, le passage de la longueur n à la longueur $n+1$ s'effectuant à l'aide des règles d'introduction et d'inversion de l'implication. Enfin, l'équivalence de 2 avec 3 (qui se démontre également par récurrence sur n) découle de la tautologie intuitionniste $(A \Rightarrow (B \Rightarrow C)) \Leftrightarrow ((A \wedge B) \Rightarrow C)$. □

2.3 Équivalences remarquables

Dans cette section, on présente un certain nombre d'équivalences (intuitionnistes et classiques) remarquables. Pour chaque équivalence, on mentionne la logique minimale qui permet de la dériver : LJ pour la logique intuitionniste, et LK pour la logique classique.

2.3.1 Propriétés algébriques des connecteurs $\wedge$ et $\vee$

Les connecteurs $\wedge$ et $\vee$ sont associatifs et commutatifs (LJ) :

$$\begin{array}{rclcrcl} (A \wedge B) \wedge C & \Leftrightarrow & A \wedge (B \wedge C) & \qquad & A \wedge B & \Leftrightarrow & B \wedge A \\ (A \vee B) \vee C & \Leftrightarrow & A \vee (B \vee C) & & A \vee B & \Leftrightarrow & B \vee A \end{array}$$

Les unités $\top$ et $\perp$ se comportent comme des éléments neutres vis à vis des connecteurs $\wedge$ et $\vee$ respectivement, et comme des éléments absorbants vis à vis des connecteurs $\vee$ et $\wedge$ respectivement (LJ) :

$$\begin{array}{rclcrcl} A \wedge \top & \Leftrightarrow & A & \qquad & A \wedge \perp & \Leftrightarrow & \perp \\ A \vee \perp & \Leftrightarrow & A & & A \vee \top & \Leftrightarrow & \top \end{array}$$

Les connecteurs $\wedge$ et $\vee$ sont distributifs l'un par rapport à l'autre (LJ)

$$\begin{array}{rcl} A \wedge (B \vee C) & \Leftrightarrow & (A \wedge B) \vee (A \wedge C) \\ A \vee (B \wedge C) & \Leftrightarrow & (A \vee B) \wedge (A \vee C) \end{array}$$

et distribuent avec les quantificateurs $\forall$ et $\exists$ respectivement (LJ) :

$$\begin{array}{rcl} \forall x\,(A \wedge B) & \Leftrightarrow & (\forall x\, A) \wedge (\forall x\, B) \\ \exists x\,(A \vee B) & \Leftrightarrow & (\exists x\, A) \vee (\exists x\, B) \end{array}$$

L'implication est réflexive et transitive (LJ) :

$$A \Rightarrow A, \qquad\qquad (A \Rightarrow B) \wedge (B \Rightarrow C) \Rightarrow (A \Rightarrow C)\,.$$

Elle distribue à droite avec la conjonction (LJ)

$$A \Rightarrow (B \wedge C) \quad \Leftrightarrow \quad (A \Rightarrow B) \wedge (A \Rightarrow C)$$

et vérifie l'isomorphisme de Curry (LJ)

$$((A \wedge B) \Rightarrow C) \Leftrightarrow (A \Rightarrow (B \Rightarrow C))$$

(en rappelant que $A \Rightarrow B \Rightarrow C$ désigne la formule $A \Rightarrow (B \Rightarrow C)$) qui se généralise immédiatement à un nombre arbitraire d'hypothèses :

$$A_1 \wedge \cdots \wedge A_n \Rightarrow B \quad \Leftrightarrow \quad A_1 \Rightarrow \cdots \Rightarrow A_n \Rightarrow B\,.$$

2.3.2 Lois de Morgan

Les lois de Morgan sont des équivalences classiques basées sur le fait qu'en logique classique, la négation est une involution. Ces lois font apparaître deux dualités : l'une entre les connecteurs $\wedge$ et $\vee$, l'autre entre les quantificateurs $\forall$ et $\exists$. Pour chacune de ces lois classiques, nous indiquons également entre parenthèses la ou les implications qui demeurent valides en logique intuitionniste :

$$\begin{array}{rcll}
A & \Leftrightarrow & \neg\neg A & \qquad (\text{LJ}: \Rightarrow) \\[1ex]
\neg\top & \Leftrightarrow & \bot & \qquad (\text{LJ}: \Leftrightarrow) \\
\neg\bot & \Leftrightarrow & \top & \qquad (\text{LJ}: \Leftrightarrow) \\[1ex]
\neg(A \wedge B) & \Leftrightarrow & (\neg A) \vee (\neg B) & \qquad (\text{LJ}: \Leftarrow) \\
\neg(A \vee B) & \Leftrightarrow & (\neg A) \wedge (\neg B) & \qquad (\text{LJ}: \Leftrightarrow) \\[1ex]
\neg(\forall x\, A) & \Leftrightarrow & \exists x\, (\neg A) & \qquad (\text{LJ}: \Leftarrow) \\
\neg(\exists x\, A) & \Leftrightarrow & \forall x\, (\neg A) & \qquad (\text{LJ}: \Leftrightarrow)
\end{array}$$

Nous n'avons pas fait figurer ici les lois de Morgan de l'implication, dans la mesure où en logique classique, ce connecteur et sa négation se décomposent de la manière suivante (LK) :

$$\begin{array}{rcll}
A \Rightarrow B & \Leftrightarrow & (\neg A) \vee B & \qquad (\text{LJ}: \Leftarrow) \\
\neg(A \Rightarrow B) & \Leftrightarrow & A \wedge (\neg B) & \qquad (\text{LJ}: \Leftarrow)
\end{array}$$

Notons pour terminer les deux tautologies classiques suivantes (LK) :

$$\begin{array}{rl}
\text{(Tiers exclus)} & \qquad A \vee \neg A \\
\text{(Loi de Pierce)} & \qquad ((A \Rightarrow B) \Rightarrow A) \Rightarrow A
\end{array}$$

2.3.3 Formes prénexes

Définition 6 (Forme prénexe) — *On dit d'une formule A qu'elle est en* forme prénexe *si elle est de la forme*

$$A \quad \equiv \quad Q_1 x_1 \ldots Q_n x_n \; A'$$

où $Q_1, \ldots, Q_n \in \{\forall;\ \exists\}$ sont des quantificateurs, et où A' est une formule sans quantificateurs.

Proposition 5 (Mise en forme prénexe) — *Pour toute formule A, il existe une formule A' en forme prénexe qui est classiquement équivalente à A.*

Preuve. La mise sous forme prénexe s'effectue en utilisant les équivalences ci-dessous comme des règles de réécriture (orientées de la gauche vers la droite) pour faire remonter en tête de formule n'importe quel quantificateur Q situé sous un connecteur (en notant $\overline{Q}$ le quantificateur dual de Q) :

$$\begin{array}{rcll}
\neg(Qx\,A) & \Leftrightarrow & \overline{Q}x\,(\neg A) & \\[1ex]
(Qx\,A) \wedge B & \Leftrightarrow & Qx\,(A \wedge B) & \quad (x \notin FV(B)) \\
A \wedge (Qx\,B) & \Leftrightarrow & Qx\,(A \wedge B) & \quad (x \notin FV(A)) \\[1ex]
(Qx\,A) \vee B & \Leftrightarrow & Qx\,(A \vee B) & \quad (x \notin FV(B)) \\
A \vee (Qx\,B) & \Leftrightarrow & Qx\,(A \vee B) & \quad (x \notin FV(A)) \\[1ex]
(Qx\,A) \Rightarrow B & \Leftrightarrow & \overline{Q}x\,(A \Rightarrow B) & \quad (x \notin FV(B)) \\
A \Rightarrow (Qx\,B) & \Leftrightarrow & Qx\,(A \Rightarrow B) & \quad (x \notin FV(A))
\end{array}$$

On notera que la condition de bord $x \notin FV(B)$ (resp. $x \notin FV(A)$) peut toujours être satisfaite dans la mesure où dans la formule $Qx\,A$ (resp. $Qx\,B$) on a renommé la variable x en une variable distincte des variables libres de la formule B (resp. de la formule A). À l'aide de ces équivalences la proposition se démontre par une récurrence double, sur la structure de la formule A et sur le nombre de quantificateurs qu'elle comporte. $\square$

Chapitre 3

Théories du premier ordre

3.1 La notion de théorie

3.1.1 Théories et théorèmes

Définition 7 — *On appelle* théorie (du premier ordre) *tout couple* $\mathscr{T} = (\Sigma, \mathcal{A})$ *formé par une signature* Σ *et par un ensemble* $\mathcal{A}$ *de formules closes construites sur* Σ*, qu'on appelle les* axiomes *de la théorie* $\mathscr{T}$*.*

Étant donnée une théorie $\mathscr{T}$, on appelle *langage de* $\mathscr{T}$ et on note $\mathcal{L}(\mathscr{T})$ l'ensemble des formules construites sur la signature de $\mathscr{T}$. L'ensemble des formules closes du langage de $\mathscr{T}$ est noté $\mathcal{L}_0(\mathscr{T})$, tandis que l'ensemble des axiomes de $\mathscr{T}$, qui est un sous-ensemble de $\mathcal{L}_0(T)$, est noté $\mathrm{Ax}(\mathscr{T})$. Par conséquence logique (au sens de la déduction naturelle), les axiomes de $\mathscr{T}$ définissent un ensemble plus vaste de formules (closes) qu'on appelle les théorèmes de $\mathscr{T}$:

Définition 8 (Théorème) — *Étant donnée une théorie* $\mathscr{T}$*, on dit qu'une formule close* $A \in \mathcal{L}_0(\mathscr{T})$ *est un* théorème de $\mathscr{T}$ *s'il existe des axiomes* $B_1, \dots, B_n$ *tels que le séquent* $B_1, \dots, B_n \vdash A$ *est dérivable. L'ensemble des théorèmes de* $\mathscr{T}$ *est noté* $\mathrm{Th}(\mathscr{T})$*.*

Autrement dit : un théorème de $\mathscr{T}$ est une formule close démontrable sous un ensemble fini d'hypothèses prises dans les axiomes de $\mathscr{T}$.

Étant données une théorie $\mathscr{T}$, une liste de formules Γ et une formule A construites dans le langage de $\mathscr{T}$, il est commode d'utiliser la notation $\mathscr{T}, \Gamma \vdash A$ pour exprimer qu'il existe une liste finie d'axiomes $\Delta \subseteq \mathrm{Ax}(\mathscr{T})$ telle que le séquent $\Delta, \Gamma \vdash A$ est dérivable. Avec cette notation, une formule $A \in \mathcal{L}_0(\mathscr{T})$ est un théorème de $\mathscr{T}$ si et seulement si $\mathscr{T} \vdash A$.

3.1.2 Inclusion et équivalence de théories

Étant données deux théories $\mathscr{T}_1$ et $\mathscr{T}_2$, on dit que $\mathscr{T}_1$ est *incluse* dans $\mathscr{T}_2$ (ou encore que $\mathscr{T}_2$ est une *extension* de $\mathscr{T}_1$) et on note $\mathscr{T}_1 \subseteq \mathscr{T}_2$ si

1. $\mathcal{L}(\mathscr{T}_1) \subseteq \mathcal{L}(\mathscr{T}_2)$ (inclusion de langages) et
2. $\mathrm{Th}(\mathscr{T}_1) \subseteq \mathrm{Th}(\mathscr{T}_2)$ (inclusion des ensembles de théorèmes).

Pour s'assurer de la deuxième condition, il suffit en pratique de vérifier que tous les axiomes de $\mathscr{T}_1$ sont des théorèmes de $\mathscr{T}_2$:

Proposition 6 — *Soient $\mathscr{T}_1$ et $\mathscr{T}_2$ deux théories. Les deux assertions suivantes sont équivalentes :*

1. $\mathscr{T}_1 \subseteq \mathscr{T}_2$
2. $\mathcal{L}(\mathscr{T}_1) \subseteq \mathcal{L}(\mathscr{T}_2)$ *et* $\mathrm{Ax}(\mathscr{T}_1) \subseteq \mathrm{Th}(\mathscr{T}_2)$.

On dit que deux théories $\mathscr{T}_1$ et $\mathscr{T}_2$ sont équivalentes et on note $\mathscr{T}_1 \approx \mathscr{T}_2$ si $\mathscr{T}_1 \subseteq \mathscr{T}_2$ et $\mathscr{T}_2 \subseteq \mathscr{T}_1$. Deux théories sont donc équivalentes si elles définissent le même langage et le même ensemble de théorèmes. En revanche, deux théories équivalentes peuvent être construites sur des systèmes d'axiomes assez différents (c'est notamment fréquemment le cas dans les diverses présentations de la théorie des ensembles). Pour vérifier que deux théories construites sur un même langage sont équivalentes, on se contente en pratique de vérifier que tout axiome de l'une est un théorème de l'autre, et vice-versa (d'après la Prop. 6).

3.1.3 Cohérence

Définition 9 (Cohérence) — *Une théorie $\mathscr{T}$ est* contradictoire *(ou* incohérente*) lorsque la formule $\bot$ est un théorème de $\mathscr{T}$. Dans le cas contraire (i.e. lorsque $\bot \notin \mathrm{Th}(\mathscr{T})$), on dit que la théorie $\mathscr{T}$ est* cohérente.

Dans la mesure où la dérivabilité de la formule $\bot$ entraîne la dérivabilité de n'importe quelle formule (règle $\bot$-elim), une théorie contradictoire est une théorie dans laquelle toutes les formules closes du langage sont des théorèmes. Inversement, une théorie $\mathscr{T}$ est cohérente lorsqu'il existe au moins une formule close (formée dans le langage de $\mathscr{T}$) qui n'est pas un théorème de $\mathscr{T}$.

Comme par ailleurs toute dérivation dans $\mathscr{T}$ de la formule $\bot$ (ou de toute autre formule) n'utilise qu'un nombre fini d'axiomes et qu'un nombre fini de symboles dans la signature, il est clair que :

Proposition 7 (Compacité) — *Une théorie $\mathscr{T} = (\Sigma, \mathcal{A})$ est cohérente si et seulement si pour tout ensemble fini d'axiomes $\mathcal{A}_0 \subseteq \mathcal{A}$ formé sur une sous-signature finie $\Sigma_0 \subseteq \Sigma$, la théorie $\mathscr{T}_0 = (\Sigma_0, \mathcal{A}_0)$ est cohérente.*

Cohérence du calcul des prédicats Bien que le langage du calcul des prédicats soit plus expressif que celui du calcul propositionnel, il est assez facile de déduire la cohérence du calcul des prédicats (sans axiome) de la cohérence du calcul propositionnel. Pour cela, on commence par définir une traduction qui à chaque formule A du calcul des prédicats associe une formule A^* du calcul propositionnel, par induction sur la structure de A à l'aide des équations

$$\begin{array}{rcll}
(p(t_1,\dots,t_n))^* & = & p & \\
(U)^* & = & U & (U \in \{\top;\bot\} \\
(\neg A)^* & = & \neg A^* & \\
(A \diamond B)^* & = & A^* \diamond B^* & (\diamond \in \{\wedge;\vee;\Rightarrow;\Leftrightarrow\} \\
(Qx\,A)^* & = & A^* & (Q \in \{\forall;\exists\})
\end{array}$$

Intuitivement, cette traduction fait disparaître toute référence aux objets du premier ordre : les formules atomiques $p(t_1,\dots,t_n)$ sont remplacées par des variables propositionnelles p, et les quantifications sont supprimées. Cette traduction s'étend immédiatement en une traduction des listes de formules du calcul des prédicats vers les listes de formules du calcul propositionnel (notation : Γ^*).

La propriété remarquable de cette traduction est qu'elle transforme tout séquent $\Gamma \vdash A$ dérivable en déduction naturelle en un séquent $\Gamma^* \vdash A^*$ qui est dérivable en déduction naturelle propositionnelle (c'est-à-dire sans les règles d'introduction et d'élimination des quantificateurs) :

Proposition 8 — *Si le séquent $\Gamma \vdash A$ est dérivable en déduction naturelle, alors le séquent propositionnel $\Gamma^* \vdash A^*$ est dérivable en déduction naturelle propositionnelle.*

Preuve. Par induction sur la structure de la dérivation du séquent $\Gamma \vdash A$. □

Proposition 9 (Cohérence du calcul des prédicats) — *Quelle que soit la signature Σ, la théorie $\mathscr{T} = (\Sigma, \varnothing)$ est cohérente.*

Preuve. Supposons que le séquent $\vdash \bot$ est dérivable en déduction naturelle classique. D'après la proposition qui précède, le séquent $(\vdash \bot)^* = \vdash \bot$ est donc également dérivable en déduction naturelle propositionnelle, ce qui est impossible puisque le calcul propositionnel est cohérent.[1]. □

3.2 Exemples

3.2.1 L'arithmétique de Peano

On appelle *arithmétique de Peano* et on note PA la théorie du premier ordre dont la signature est formée par

- un symbole de constante noté 0 (« zéro ») ;
- un symbole de fonction s (« successeur ») d'arité 1 ;
- deux symboles de fonction $+$ (« plus ») et $\times$ (« fois ») d'arité 2 ;
- un symbole de prédicat $=$ (« égale ») d'arité 2.

et dont les axiomes sont les formules closes (1–15) données à la Fig. 3.1. Ces axiomes (qui sont en nombre infini) comprennent :

Les axiomes d'égalité qui expriment que l'égalité est réflexive (1), symétrique (2), transitive (3) — c'est-à-dire une relation d'équivalence — et que tous les symboles de fonction de la signature sont compatibles vis-à-vis de cette relation d'équivalence, à savoir : le successeur (4), l'addition $(5, 6)$ et la multiplication $(7, 8)$.

Les axiomes de calcul qui définissent l'addition $(9, 10)$ et la multiplication $(11, 12)$ à travers leurs règles de calcul.

[1] On rappelle que la cohérence du calcul propositionnel se démontre à l'aide des valeurs de vérité : on commence par associer une valeur de vérité quelconque (i.e. 0 ou 1) à chaque variable propositionnelle, puis on étend l'interprétation à toutes les formules et à tous les séquents (en interprétant chaque séquent comme sa formule associée). On montre ensuite par induction sur la structure de la dérivation que tout séquent dérivable en déduction naturelle propositionnelle a une valeur de vérité égale à 1. Ce qui n'est évidement pas le cas du séquent $\vdash \bot$ (dont la valeur de vérité vaut 0), qui n'est donc pas dérivable.

Réflexivité, symétrie et transitivité de l'égalité

1. $\forall x\ (x = x)$
2. $\forall x\ \forall y\ (x = y \Rightarrow y = x)$
3. $\forall x\ \forall y\ \forall z\ (x = y \wedge y = z \Rightarrow x = z)$

Compatibilité avec les symboles de fonction

4. $\forall x\ \forall x'\ (x = x' \Rightarrow s(x) = s(x'))$
5. $\forall x\ \forall x'\ \forall y\ (x = x' \Rightarrow x + y = x' + y)$
6. $\forall x\ \forall y\ \forall y'\ (y = y' \Rightarrow x + y = x + y')$
7. $\forall x\ \forall x'\ \forall y\ (x = x' \Rightarrow x \times y = x' \times y)$
8. $\forall x\ \forall y\ \forall y'\ (y = y' \Rightarrow x \times y = x \times y')$

Axiomes de calcul (addition, multiplication)

9. $\forall y\ (0 + y = y)$
10. $\forall x\ \forall y\ (s(x) + y = s(x + y))$
11. $\forall y\ (0 \times y = 0)$
12. $\forall x\ \forall y\ (s(x) \times y = (x \times y) + y)$

Injectivité et non surjectivité du successeur

13. $\forall x\ \forall y\ (s(x) = s(y) \Rightarrow x = y)$
14. $\forall x\ \neg(s(x) = 0)$

Axiomes de récurrence

15. $\forall x_1\ \cdots\ \forall x_n\ \big(A\{x := 0\}\ \wedge\ \forall x\,(A \Rightarrow A\{x := s(x)\})\ \Rightarrow\ \forall x\, A\big)$
    pour chaque formule A telle que $FV(A) = \{x; x_1; \ldots; x_n\}$

FIG. 3.1 – Axiomes de l'arithmétique de Peano

Les axiomes de Peano qui sont les axiomes arithmétiques proprement dits, à savoir : l'axiome (13) exprimant que le successeur est injectif (historiquement connu sous le nom de « 3e axiome de Peano »[2]), l'axiome (14) exprimant que zéro n'est le successeur d'aucun entier (le « 4e axiome de Peano ») et enfin le principe de récurrence (15) (ou « 5e axiome de Peano »). On notera que dans l'arithmétique de Peano, le principe de récurrence n'est pas un axiome, mais un *schéma d'axiomes*, c'est-à-dire un schéma qui permet de définir un nouvel axiome pour chaque couple (A, x) constitué d'une formule A et d'une variable $x \in FV(A)$[3] (les autres variables libres de A étant nommées $x_1, \dots, x_n$). De ce fait, l'ensemble des axiomes de l'arithmétique de Peano est infini.

Les axiomes (4–8) qui expriment que les symboles de fonction de la signature sont compatibles vis-à-vis de l'égalité s'étendent à tous les assemblages de symboles de fonction, c'est-à-dire à tous les termes :

Proposition 10 (Compatibilité) — *Pour tout terme t, on a*

$$\mathrm{PA}\ \vdash\ x = y\ \Rightarrow\ t\{z := x\} = t\{z := y\}\,.$$

Preuve. Par récurrence sur la structure du terme t. □

De même, toute formule A qui vaut pour un certain objet x vaut également pour n'importe quel objet y égal à x :

Proposition 11 (Principe de Leibniz) — *Pour toute formule A, on a*

$$\mathrm{PA}\ \vdash\ x = y\ \wedge\ A\{z := x\}\ \Rightarrow\ A\{z := y\}\,.$$

Preuve. On démontre par récurrence sur la structure de la formule A que

$$\mathrm{PA}\ \vdash\ x = y\ \Rightarrow\ (A\{z := x\} \Leftrightarrow A\{x := y\})$$

en utilisant la Prop. 10 dans le cas des formules atomiques $t = u$. □

(Nous verrons à la sous-section 3.2.2 comment généraliser ce mécanisme à une large classe de théories dites égalitaires.)

Expressivité de l'arithmétique de Peano Bien que son langage et ses axiomes soit très élémentaires, l'arithmétique de Peano est une théorie très expressive qui peut démontrer bien plus que les simples propriétés algébriques de l'addition et de la multiplication (associativité, commutativité, distributivité, etc.) À travers les abréviations

$$x \leq y \quad\equiv\quad \exists z\ (x + z = y) \qquad\qquad x < y \quad\equiv\quad s(x) \leq y$$
$$x \text{ premier} \quad\equiv\quad x > 1\ \wedge\ \forall y\ \forall z\ (x = y \times z \Rightarrow y = 1 \vee z = 1)$$

[2]La présentation originelle de l'arithmétique par Peano commençait par deux axiomes exprimant 1. que « 1 est un entier » (Peano faisant commencer les entiers à partir de 1) et 2. que « le successeur d'un entier est un entier ». Ces deux axiomes n'ont aujourd'hui plus de raison d'être dans les présentations modernes de l'arithmétique.

[3]Dans le cas où $x \notin FV(A)$, le principe de récurrence (15) associé à la formule A est en réalité une tautologie, démontrable sans l'aide d'aucun axiome.

elle peut exprimer toutes les propriétés relatives à l'ordre et à la décomposition des entiers en facteurs premiers. En particulier, le principe du minimum

$$\exists x\, A(x) \;\Rightarrow\; \exists x\, (A(x) \wedge \forall y\ (A(y) \Rightarrow x \leq y)$$

(pour toute formule $A(x)$ dépendant d'une variable x) ainsi que le théorème d'Euclide exprimant l'existence d'une infinité de nombre premiers

$$\forall x\ \exists y\ (y > x\ \wedge\ y \text{ premier})$$

sont tous les deux démontrables dans l'arithmétique de Peano.

Plus généralement, il est possible d'exprimer dans l'arithmétique de Peano toutes les structures de données finies (à travers des codages complexes dans les entiers) et de démontrer la plupart des propriétés combinatoires de ces structures. Ainsi est-il possible de représenter dans l'arithmétique non seulement les listes finies, les ensembles et multi-ensembles finis, les arbres finis, mais aussi toutes les structures syntaxiques de la logique telles que les termes, les formules, les séquents et les dérivations. En pratique, on pourrait donc formaliser intégralement les définitions et résultats présentés dans les quatre premiers chapitres de ce cours dans l'arithmétique de Peano. (Ce résultat technique est dû au logicien Kurt Gödel, et il constitue l'un des ingrédients essentiels de la démonstration de ses deux célèbres théorèmes d'incomplétude.)

3.2.2 Théories égalitaires

Les propositions 10 et 11 peuvent se généraliser à une classe plus vaste de théories, à savoir les théories égalitaires.

Définition 10 (Théorie égalitaire) — *On appelle* théorie égalitaire *toute théorie* $\mathscr{T}$ *comportant un symbole de prédicat binaire* $=$ *(« égale ») tel que*

1. $\mathscr{T} \vdash \forall x\ (x = x)$
2. $\mathscr{T} \vdash x = y \wedge A\{z := x\} \Rightarrow A\{z := y\}$ *pour toute formule* $A \in \mathcal{L}(\mathscr{T})$*.*

(Les points 1 et 2 sont généralement appelés le principe d'identité *et le* principe de Leibniz*, respectivement.)*

La principale difficulté qu'il y a à établir qu'une théorie donnée est égalitaire (ou tout simplement à en construire une) réside dans le 2e point (principe de Leibniz) qui porte sur un ensemble de formules qui est infini. Cependant, on peut considérablement restreindre le nombre de vérifications (et même le ramener à un nombre fini dans le cas où la signature est finie) en remarquant qu'une théorie est égalitaire si et seulement si chaque symbole de fonction et chaque symbole de prédicat défini dans la signature est compatible vis-à-vis de la relation d'égalité sur chaque composante :

Proposition 12 — *Une théorie* $\mathscr{T}$ *est égalitaire (de prédicat d'égalité* $=$*) si et seulement si les formules suivantes sont des théorèmes de* $\mathscr{T}$

1. $\forall x\ (x = x)$
2. $\forall x_1 \cdots \forall x_{i-1}\ \forall x_i\ \forall x_i'\ \forall x_{i+1} \cdots \forall x_n$
   $(x_i = x_i' \Rightarrow f(x_1, \ldots, x_i, \ldots, x_n) = f(x_1, \ldots, x_i', \ldots, x_n))$

3. $\forall x_1 \cdots \forall x_{i-1}\ \forall x_i\ \forall x'_i\ \forall x_{i+1} \cdots \forall x_n$
$(x_i = x'_i \wedge p(x_1, \ldots, x_i, \ldots, x_n) \Rightarrow p(x_1, \ldots, x'_i, \ldots, x_n))$

pour chaque symbole de fonction f et pour chaque symbole de prédicat p d'arité n dans la signature de $\mathscr{T}$, ainsi que pour chaque indice $1 \leq i \leq n$.

Preuve. Les conditions 1, 2 et 3 sont clairement nécessaires d'après le principe d'identité (pour 1) et le principe de Leibniz (pour 2 et 3). Réciproquement, supposons que $\mathscr{T}$ satisfait les conditions 1, 2 et 3. On commence par remarquer que la condition 3 appliquée au prédicat d'égalité lui-même donne la symétrie et la transitivité de l'égalité (en utilisant la réflexivité donnée par 1), ce qui permet de démontrer dans $\mathscr{T}$ que l'égalité est une relation d'équivalence. En utilisant la condition 2, on montre par induction sur le terme t que

$$\mathscr{T} \vdash x = y \Rightarrow t\{z := x\} = t\{z := y\}$$

(cf Prop. 10) avant de montrer finalement le principe de Leibniz

$$\mathscr{T} \vdash x = y \wedge A\{z := x\} \Rightarrow A\{z := y\}$$

par induction sur la formule A (cf Prop. 11). □

3.2.3 La théorie des ensembles de Zermelo-Fraenkel

On appelle *théorie des ensembles de Zermelo-Fraenkel* et on note ZF la théorie du premier ordre dont la signature est formée par

- aucun symbole de fonction (ou de constante) ;
- deux symboles de prédicat binaires $=$ (« égale »), $\in$ (« appartient »)

et dont les axiomes, donnés à la Fig. 3.2, sont formulés à l'aide des abréviations suivantes :

$$\begin{array}{rcl}
x \notin y & \equiv & \neg(x \in y) \\
x \subset y & \equiv & \forall z\ (z \in x \Rightarrow z \in y) \\
\forall x \in y\ A & \equiv & \forall x\ (x \in y \Rightarrow A) \\
\exists x \in y\ A & \equiv & \exists x\ (x \in y \wedge A) \\
\exists! x\ A & \equiv & \exists x\ (A \wedge \forall y\ (A\{x := y\} \Rightarrow y = x)) \\
\mathrm{zero}(x) & \equiv & \forall z\ z \notin x \\
\mathrm{succ}(x, y) & \equiv & \forall z\ (z \in y \Leftrightarrow z = x \vee z \in x)
\end{array}$$

Ces axiomes comprennent :

Les axiomes d'égalité qui expriment que l'égalité est réflexive, symétrique, transitive — c'est-à-dire une relation d'équivalence — et que la relation d'appartenance est compatible avec l'égalité (à gauche et à droite). Ces axiomes permettent à la théorie des ensembles de satisfaire aux critères 1–3 de la Prop. 12 (le critère 2 étant trivialement réalisé), et donc de constituer une théorie égalitaire au sens de la Déf. 10.

L'axiome d'extensionnalité qui exprime que deux ensembles a et b qui ont les mêmes éléments sont égaux. (Il est important de ne pas confondre cet axiome avec l'axiome de compatibilité à droite, qui exprime l'implication réciproque, à savoir : deux ensembles égaux ont les mêmes éléments.)

Axiomes d'égalité

(Reflexivité)	$\forall x\ (x = x)$
(Symétrie)	$\forall x\ \forall y\ (x = y \Rightarrow y = x)$
(Transitivité)	$\forall x\ \forall y\ \forall z\ (x = y \wedge y = z \Rightarrow x = z)$
(Compat. à gauche)	$\forall x\ \forall y\ \forall z\ (x = y \wedge x \in z \Rightarrow y \in z)$
(Compat. à droite)	$\forall x\ \forall y\ \forall z\ (x \in y \wedge y = z \Rightarrow x \in z)$

Axiomes de Zermelo

(Extensionnalité)	$\forall a\ \forall b\ (\forall x\ (x \in a \Leftrightarrow x \in b)\ \Rightarrow\ a = b)$
(Paire)	$\forall a\ \forall b\ \exists c\ \forall x\ (x \in c\ \Leftrightarrow\ x = a \vee x = b)$
(Compréhension)	$\forall z_1 \cdots \forall z_n\ \forall a\ \exists b\ \forall x\ (x \in b\ \Leftrightarrow\ x \in a \wedge A)$
	pour toute formule A telle que $FV(A) \subset \{z_1; \ldots; z_n; a; x\}$
(Union)	$\forall a\ \exists b\ \forall x\ (x \in b\ \Leftrightarrow\ \exists y\ (y \in a \wedge x \in y))$
(Parties)	$\forall a\ \exists b\ \forall x\ (x \in b\ \Leftrightarrow\ x \subseteq a)$
(Infini)	$\exists a\ (\exists x \in a\ \mathrm{zero}(x)\ \wedge\ \forall x \in a\ \exists y \in a\ \mathrm{succ}(x, y))$

Schéma de Fraenkel-Skolem

(Remplacement)	$\forall z_1 \cdots \forall z_n\ \forall a\ (\forall x \in a\ \exists! y\ A\ \ \Rightarrow\ \ \exists b\ \forall x \in a\ \exists y \in b\ A)$
	pour toute formule A telle que $FV(A) \subset \{z_1; \ldots; z_n; a; x; y\}$

FIG. 3.2 – Axiomes de la théorie des ensembles de Zermelo-Fraenkel

On notera qu'en informatique cet axiome n'a rien d'évident, car un même ensemble fini peut être représenté d'un grand nombre de manières différentes. (Par exemple, sous forme de liste, l'ensemble $\{a; b\}$ peut être représenté indifféremment par $[a; b]$, $[b; a]$, $[a; b; a]$, $[b; b; a]$, etc.) Dans cette perspective, on peut donc comprendre l'axiome d'extensionnalité comme un principe de transparence de la théorie des ensembles vis-à-vis des différentes manières de représenter (ou de construire) un même ensemble.

L'axiome de la paire qui exprime que pour tous ensembles a et b, il existe un ensemble c dont les seuls éléments sont a et b. Cet ensemble, que l'on note $\{a; b\}$, est unique d'après l'axiome d'extensionnalité. L'axiome de la paire permet également de construire le singleton $\{a\}$, en appliquant l'axiome avec $a = b$.

Le schéma de compréhension (ou de *séparation*) qui permet de construire, à partir d'un ensemble a et d'une formule $A(x)$, l'ensemble de tous les $x \in a$ tels que $A(x)$, que l'on note généralement $\{x \in a \mid A(x)\}$. (Là encore cet ensemble est unique, par extensionnalité.) Le schéma de compréhension — qui introduit dans la théorie un nouvel axiome pour chaque formule $A(x)$ — permet de définir de nombreuses autres constructions, telles que l'ensemble vide $\varnothing = \{x \in a \mid \bot\}$ (qui ne dépend pas de l'ensemble a utilisé pour le construire), l'intersection binaire $a \cap b = \{x \in a \mid x \in b\}$ ou la différence ensembliste $a \setminus b = \{x \in a \mid x \notin b\}$.

L'axiome de l'union qui permet de définir l'ensemble formé par la réunion de tous les éléments d'un ensemble a donné, ensemble que l'on note généralement $\bigcup_{x \in a} x$ ou plus simplement $\bigcup a$. Combiné avec l'axiome de la paire, l'axiome de l'union permet de définir l'union binaire $a \cup b = \bigcup\{a; b\}$ et plus généralement toutes les unions finies.

L'axiome des parties qui permet de construire l'ensemble des parties (i.e. des sous-ensembles) d'un ensemble a donné, que l'on note $\mathfrak{P}(a)$.

L'axiome de l'infini qui exprime l'existence d'un ensemble a contenant 0 (défini par $0 = \varnothing$) et clos par l'opération successeur s (définie pour tout ensemble x par $s(x) = x \cup \{x\} = \bigcup\{x; \{x; x\}\}$).[4] Cet axiome entraîne que l'ensemble a contient au moins tous les entiers naturels, mais n'empêche pas que a contienne d'autres objets. Pour définir précisément l'ensemble des entiers naturels à partir de cet ensemble a, il faut encore appliquer le schéma de compréhension avec la formule idoine et poser

$$\mathbb{N} \;=\; \big\{x \in a \mid \forall b\,(0 \in b \;\wedge\; \forall y\,(y \in b \Rightarrow s(y) \in b) \;\;\Rightarrow\;\; x \in b)\big\}\,.$$

Le schéma de remplacement qui exprime l'existence d'un ensemble b qui contient (au moins) l'image d'un ensemble a à travers une relation fonctionnelle $A(x, y)$ donnée (i.e. telle qu'il existe un unique y tel que $A(x, y)$ pour chaque élément $x \in a$). Ce schéma d'axiomes est nécessaire dans les cas où on ne connaît pas à l'avance un ensemble b contenant toutes les images possibles d'un élément de a à travers la relation $A(x, y)$, ensemble

[4] En théorie des ensembles, chaque entier naturel n est défini comme l'ensemble des entiers naturels strictement plus petits que lui : $n = \{0; \ldots; n-1\}$. Ainsi, on a $0 = \varnothing$, $1 = \{0\} = \{\varnothing\}$, $2 = \{0; 1\} = \{\varnothing; \{\varnothing\}\}$, etc. Avec cette représentation des entiers (due à von Neumann), il est facile de se convaincre que le successeur d'un entier n est défini par $s(n) = n \cup \{n\}$.

dont l'existence est pourtant nécessaire[5] pour construire l'ensemble-image recherché à l'aide du schéma de compréhension

$$\mathrm{Im}(a, A) \;=\; \{y \in b \mid \exists x \in a\ A(x, y)\}\,.$$

En pratique, on utilise généralement cet axiome avec une relation fonctionnelle $A(x, y)$ présentée sous forme skolémisée, en introduisant une notation $f(x)$ pour désigner l'unique y tel que $A(x, y)$ (pour chaque $x \in a$). L'image de l'ensemble a par la relation $y = f(x)$ (défini en combinant les schémas de remplacement et de compréhension) est alors notée $\{f(x) \mid x \in a\}$.

3.3 Extensions conservatives

Ainsi que nous l'avons vu en 3.1.2, la notion d'extension de théorie doit s'entendre à la fois au sens d'une extension du langage sous-jacent (par extension de signature) et au sens d'une extension de l'ensemble des théorèmes qu'il est possible de démontrer dans la théorie. À ce titre, il est important de souligner que le simple fait d'ajouter à une théorie $\mathscr{T}$ un nouveau symbole de fonction ou de prédicat suffit à étendre l'ensemble des théorèmes de $\mathscr{T}$, ne serait-ce que parce que l'extension de langage qui en résulte permet d'écrire — et de démontrer — de nouvelles tautologies. Cependant, la plupart des extensions $\mathscr{T}' \supseteq \mathscr{T}$ font apparaître de nouveaux théorèmes y compris dans le langage de la théorie initiale, en ce sens que l'inclusion

$$\mathrm{Th}(\mathscr{T}') \cap \mathcal{L}(\mathscr{T}) \;\supseteq\; \mathrm{Th}(\mathscr{T})$$

est stricte en règle générale (ce qui est typiquement le cas lorsqu'on ajoute des axiomes à la théorie sans modifier son langage).

3.3.1 La notion d'extension conservative

Dans certaines situations, il est désirable d'étendre une théorie avec de nouveaux symboles et de nouveaux axiomes décrivant le comportement des nouveaux symboles sans pour autant étendre la classe des théorèmes formulés dans l'ancien langage, ce qui nous amène à introduire la notion suivante :

Définition 11 (Extension conservative) — *Soient $\mathscr{T}$ une théorie et $\mathscr{T}'$ une extension de $\mathscr{T}$. On dit que $\mathscr{T}'$ est une* extension conservative *de $\mathscr{T}$ si tout théorème de $\mathscr{T}'$ formulé dans le langage de $\mathscr{T}$ est démontrable dans $\mathscr{T}$, c'est-à-dire si :*

$$\mathrm{Th}(\mathscr{T}') \cap \mathcal{L}(\mathscr{T}) \;=\; \mathrm{Th}(\mathscr{T})\,.$$

Un des intérêts de la notion d'extension conservative est qu'elle préserve la propriété de cohérence, contrairement à la notion d'extension en général :

Proposition 13 (Cohérence par conservativité) — *Si $\mathscr{T}$ est une théorie cohérente, alors toute extension conservative de $\mathscr{T}$ est cohérente également.*

[5] La nécessité d'un tel axiome semble avoir été remarquée de manière indépendante par Fraenkel et Skolem. On appelle *théorie des ensembles de Zermelo* et on note Z la théorie qui reprend les mêmes axiomes (égalité, extensionnalité, paire, compréhension, union, parties et infini) sans le schéma de remplacement.

Preuve. Si $\bot$ est prouvable dans $\mathscr{T}'$ alors $\bot$ est prouvable dans $\mathscr{T}$ également, puisque $\bot \in \mathcal{L}(\mathscr{T})$. □

La proposition suivante établit que l'ajout de symboles de fonctions et de prédicats (sans modification de l'ensemble des axiomes) définit une extension conservative de la théorie. On notera que la démonstration de ce résultat — pourtant très intuitif — n'est pas complètement immédiate dans la mesure où elle repose sur l'utilisation de fonctions de rétraction pour faire disparaître les symboles additionnels dans les arbres de dérivation.

Proposition 14 (Extension de signature) — *Soit $\mathscr{T}$ une théorie basée sur une signature Σ et sur un ensemble d'axiomes $\mathcal{A}$. Pour toute signature Σ' étendant la signature Σ, la théorie $\mathscr{T}' = (\Sigma', \mathcal{A})$ est une extension conservative de la théorie $\mathscr{T} = (\Sigma, \mathcal{A})$.*

Preuve. Il s'agit de montrer que si un séquent $\Gamma \vdash A$ formé dans Σ admet une dérivation π formée dans la signature $\Sigma' \supseteq \Sigma$, alors on peut faire disparaître dans π toutes les occurrences des symboles déclarés dans $\Sigma' \setminus \Sigma$ de manière à obtenir une dérivation π^* avec les seuls symboles déclarés dans Σ (et dont la conclusion reste le séquent $\Gamma \vdash A$). Pour cela, on considère une variable x_0 fixée, et on définit les trois correspondances suivantes :

Termes. À chaque terme t formé dans Σ', on associe un terme t^* formé avec les seuls symboles de Σ en posant :

$$\begin{array}{rcl} x^* & \equiv & x \\ \big(f(t_1,\ldots,t_n)\big)^* & \equiv & \begin{cases} f(t_1^*,\ldots,t_n^*) & \text{si } f \in \Sigma \\ x_0 & \text{sinon} \end{cases} \end{array}$$

Par construction, la transformation $t \mapsto t^*$ laisse invariants tous les termes déjà formés avec les seuls symboles de Σ (i.e. $t^* = t$). Elle est en outre substitutive, en ce sens que

$$(t\{x := u\})^* \equiv t^*\{x := u^*\}$$

pour tous termes t et u tels que $x \notin FV(t,u)$, et pour toute variable $x \neq x_0$.

Formules. De même, on associe à chaque formule $A \in \mathcal{L}(\Sigma')$ une formule $A^* \in \mathcal{L}(\Sigma)$ en posant

$$\big(p(t_1,\ldots,t_n)\big)^* \equiv \begin{cases} p(t_1^*,\ldots,t_n^*) & \text{si } p \in \Sigma \\ \bot & \text{sinon} \end{cases}$$

pour chaque formule atomique $p(t_1,\ldots,t_n)$, cette définition s'étendant aussitôt à toutes les formules de $\mathcal{L}(\Sigma')$ par simple propagation à travers les connecteurs et les quantificateurs. Là encore, il est clair que les formules déjà formées dans Σ sont invariantes par la transformation $A \mapsto A^*$, et que cette transformation est substitutive dans le sens où

$$(A\{x := t\})^* \equiv A^*\{x := t^*\}$$

pour toute formule A et pour tout terme t dans lesquels x_0 n'a aucune occurrence (libre ou liée) et pour toute variable x distincte de x_0.

Dérivations. Enfin, pour chaque dérivation π formée sur la signature Σ', on note π^* la structure arborescente obtenue en remplaçant chaque séquent $\Gamma \vdash A$ étiquetant un nœud de π par le séquent $\Gamma^* \vdash A^*$, le reste de la structure demeurant inchangé. On notera cependant que l'arbre π^* ne constitue une dérivation bien formée (sur Σ) que dans le cas où la variable x_0 n'apparaît pas dans π (cette condition est en effet essentielle pour assurer la préservation de la structure des règles d'introduction et d'élimination des quantificateurs ainsi que des conditions de bord correspondantes).

Revenons à la démonstration de la proposition et supposons que A est un théorème de $\mathscr{T}'$ formulé dans la signature Σ. D'après cette hypothèse, il existe une dérivation π (formée dans Σ') d'un séquent de la forme $\Gamma \vdash A$, où Γ est un contexte formé uniquement par des axiomes de $\mathscr{T}'$ — et donc par des axiomes de $\mathscr{T}$ puisque les deux théories ont les mêmes axiomes. Si x_0 est une variable qui n'apparaît nulle part dans la dérivation π, alors, d'après ce qui précède, l'arbre π^* constitue une dérivation bien formée dans Σ dont la conclusion est le séquent $\Gamma^* \vdash A^*$. Mais comme Γ et A sont tous deux formés dans la signature Σ, on a $\Gamma^* = \Gamma$ et $A^* = A$. Par conséquent, π^* est bien une dérivation de $\Gamma \vdash A$, et donc $A \in \mathrm{Th}(\mathscr{T})$. □

3.3.2 Constantes de Henkin

Le cas d'extension conservative étudié ci-dessus (par ajout de symboles sans toucher aux axiomes) n'est pas très intéressant dans la mesure où en pratique, on ajoute rarement un nouveau symbole à une théorie sans donner avec ce symbole un ou plusieurs axiomes qui en précisent la signification. Un cas beaucoup plus intéressant d'extension conservative est le suivant :

Définition 12 (Extension de Henkin) — *Soient $\mathscr{T}$ une théorie et T un théorème de $\mathscr{T}$ de la forme $T \equiv \exists x\, A(x)$ où $A(x)$ est une formule qui ne dépend que de la variable x. On appelle* extension de Henkin *de la théorie $\mathscr{T}$ vis-à-vis du théorème T la théorie formée en ajoutant à $\mathscr{T}$:*

- *un nouveau symbole de constante c (une* constante de Henkin*) ;*
- *un nouvel axiome : $A\{x := c\}$.*

Comme on le voit, ce mécanisme d'extension est la traduction en logique formelle du procédé habituel en mathématiques qui consiste à donner un nom (c'est-à-dire un statut de constante) à un objet vérifiant une certaine propriété sitôt que l'existence d'un tel objet a été démontrée.[6]

Proposition 15 (Conservativité de Henkin) — *Toute extension de Henkin est une extension conservative.*

Preuve. La preuve de ce résultat reprend le principal ingrédient de la preuve de la proposition 14, à savoir les traductions $t \mapsto t^*$, $A \mapsto A^*$ et $\pi \mapsto \pi^*$ qui remplacent dans les termes, les formules et les dérivations chaque occurrence

[6] C'est par ce procédé qu'on introduit toutes les constantes globales en mathématiques, comme les constantes $\varnothing$, π ou $\mathbb{R}$, qui sont respectivement les constantes de Henkin associées aux formules « x n'a aucun élément », « x est la demi-période de la fonction cosinus » ou « x est un corps totalement ordonné archimédien complet ».

d'un symbole de fonction supplémentaire — ici la constante c — par une variable x_0 choisie en dehors de l'ensemble des variables qui apparaissent dans la dérivation considérée. Supposons donc que B est un théorème de $\mathscr{T}'$ tel que $B \in \mathcal{L}(\mathscr{T})$ (i.e. la constante c n'apparaît pas dans B). Par définition, il existe un contexte Γ formé par des axiomes de $\mathscr{T}'$ ainsi qu'une dérivation π du séquent $\Gamma \vdash B$, elle-même formée dans le langage de $\mathscr{T}'$. Quitte à affaiblir le contexte Γ avec l'axiome $A(c)$ (s'il n'y figure pas déjà), on peut supposer que Γ est de la forme $\Gamma \equiv \Gamma_0, A(c)$, où Γ_0 ne contient plus que des axiomes de la théorie $\mathscr{T}$. En appliquant la transformation $\pi \mapsto \pi^*$ (Prop. 14) à la dérivation π du séquent $\Gamma_0, A(c) \vdash B$, on obtient une dérivation π^* du séquent $\Gamma_0, A(x_0) \vdash B$ qui est entièrement formée dans le langage de $\mathscr{T}$. De cette dérivation on tire alors aisément une dérivation du théorème B dans la théorie $\mathscr{T}$:

$$\dfrac{\dfrac{\vdots}{\Delta \vdash \exists x_0\, A(x_0)} \qquad \dfrac{\vdots\ \pi^*}{\Gamma_0, A(x_0) \vdash B}}{\Delta, \Gamma_0 \vdash B}\ \text{Aff.+}\exists\text{-intro}$$

où Δ est un contexte d'axiomes de $\mathscr{T}$ suffisant pour démontrer le théorème $\exists x\, A(x)$ dans la théorie $\mathscr{T}$. □

3.3.3 Extension de Skolem

Définition 13 (Extension de Skolem) — *Soient $\mathscr{T}$ une théorie et T un théorème de $\mathscr{T}$ de la forme*

$$T \;\equiv\; \forall x_1 \;\cdots\; \forall x_n\; \exists y\; A(x_1, \ldots, x_n, y)\,,$$

où $A(x_1, \ldots, x_n, y)$ est une formule du langage de $\mathscr{T}$ qui ne dépend que des variables $x_1, \ldots, x_n, y$. On appelle extension de Skolem *de la théorie $\mathscr{T}$ vis-à-vis du théorème T la théorie formée en ajoutant à $\mathscr{T}$:*

- *un nouveau symbole de constante f d'arité n* (fonction de Skolem)*;*
- *un nouvel axiome :* $\forall x_1 \;\cdots\; \forall x_n\; A(x_1, \ldots, x_n, f(x_1, \ldots, x_n))$.

L'extension de Skolem est une généralisation naturelle de la notion d'extension de Henkin dans le cas où l'objet y dont a démontré l'existence dépend d'objets $x_1, \ldots, x_n$ à travers la relation $A(x_1, \ldots, x_n)$ — on parle alors de ***skolémisation*** du théorème T. Là encore, il s'agit de la traduction en logique formelle du procédé habituel en mathématiques qui consiste à introduire une notation fonctionnelle $f(x_1, \ldots, x_n)$ pour nommer un objet y tel que $A(x_1, \ldots, x_n, y)$ sitôt qu'on a montré l'existence d'un tel objet pour tous $x_1, \ldots, x_n$.[7]

Le résultat de conservativité démontré à la proposition 15 se généralise naturellement aux extensions de Skolem :

Proposition 16 (Conservativité de Skolem) — *Toute extension de Skolem est une extension conservative.*

Il n'est pas possible de généraliser la preuve de la proposition 15 aux extensions de Skolem : la preuve de conservativité des extensions de Skolem nécessite

[7] En théorie des ensembles par exemple, les notations $\{a; b\}$ et $\mathfrak{P}(a)$ sont obtenues en skolémisant l'axiome de la paire et l'axiome de l'ensemble des parties.

en effet des techniques de théorie des modèles bien plus sophistiquées que les techniques syntaxiques utilisées pour démontrer la conservativité des extensions de Henkin. Pour cette raison, la preuve de conservativité des extensions de Skolem ne sera donnée qu'au chapitre 4, lorsque nous aurons introduit les outils nécessaires pour en faire la démonstration.

Extension de Skolem égalitaire Dans le cas où $\mathscr{T}$ est une théorie égalitaire, mais où il n'y a pas forcément unicité de l'objet y tel que $A(x_1, \ldots, x_n, y)$, il n'est pas possible (en règle générale) de démontrer que le symbole de Skolem f est compatible avec l'égalité. Par conséquent, l'extension de Skolem associée au théorème T n'est donc plus une théorie égalitaire !

Pour pouvoir introduire une extension de Skolem sans remettre en cause le statut de théorie égalitaire, il suffit en pratique de démontrer à côté du théorème d'existence

$$T \quad \equiv \quad \forall x_1 \ \cdots \ \forall x_n \ \exists y \ A(x_1, \ldots, x_n, y)$$

un théorème d'unicité de la forme

$$U \quad \equiv \quad \forall x_1 \ \cdots \ \forall x_n \ \forall y \ \forall y' \ (A(x_1, \ldots, x_n, y) \land A(x_1, \ldots, x_n, y) \Rightarrow y = y')\,,$$

qui garantit que le symbole de Skolem associé au théorème T est compatible avec l'égalité (par rapport à toutes ses composantes), et ainsi que l'extension de Skolem associée reste une théorie égalitaire.[8]

[8]Ainsi en théorie des ensembles, l'axiome d'extensionnalité implique l'unicité de la paire, de l'ensemble des parties, etc. et permet ainsi d'introduire des symboles de Skolem ($\{a; b\}$, $\mathfrak{P}(a)$, etc.) sans remettre en cause le statut de théorie égalitaire.

Chapitre 4

Sémantique du calcul des prédicats

Dans les chapitres qui précèdent, nous n'avons donné de la logique du premier ordre qu'une présentation purement syntaxique : si nous nous sommes intéressés aux mécanismes qui permettent de former les expressions du langage (les termes, les formules et les séquents) ainsi qu'aux règles de déduction qui permettent de construire les raisonnements (les démonstrations), jamais nous n'avons cherché à déterminer à quels objets pouvaient référer les expressions du langage. C'est ainsi que nous avons systématiquement parlé de *symboles* de fonction sans jamais parler de fonction, et que nous avons toujours parlé de *symboles* de prédicat sans jamais parler de prédicat ou de relation.

En pratique cependant, les expressions du langage de la logique du premier ordre sont censées décrire des objets du monde mathématique ou des objets informatiques, voire des objets du monde physique. Il est donc très naturel de se demander à quoi peuvent référer les termes et les formules d'une théorie, et c'est précisément là qu'est l'objet de la théorie des modèles. Dans ce chapitre on se limitera à une présentation sommaire de la théorie, dont nous ne retiendrons essentiellement que les résultats de correction et de complétude.

4.1 Σ-structures

4.1.1 La notion de Σ-structure

Pour interpréter les termes et les formules d'un langage du premier ordre défini à partir d'une signature Σ, il est nécessaire de se donner un ensemble non vide $\mathscr{M}$ (correspondant au domaine des quantifications universelles et existentielles) accompagné d'un certain nombre de fonctions et de relations (définies sur l'ensemble $\mathscr{M}$) destinées à interpréter les symboles de fonction et de prédicat déclarés dans Σ :

Définition 14 (Σ-structure) — *Étant donnée une signature Σ, on appelle* Σ-structure *tout ensemble non vide $\mathscr{M}$ muni*

– *pour chaque symbole de fonction f d'arité n, d'une fonction*

$$[\![f]\!]^{\mathscr{M}} \; : \; \mathscr{M}^n \to \mathscr{M}$$

appelée la dénotation *de f dans $\mathscr{M}$;*
– *pour chaque symbole de prédicat p d'arité n, d'une fonction*

$$[\![p]\!]^{\mathscr{M}} \ : \ \mathscr{M}^n \to \{0;1\}$$

appelée la dénotation *de p dans $\mathscr{M}$.*

On notera qu'avec cette définition, la relation qui interprète un symbole de prédicat p d'arité n est donnée sous la forme d'une *fonction de vérité* $[\![p]\!]^{\mathscr{M}} : \mathscr{M}^n \to \{0;1\}$ qui à chaque n-uplet $(v_1,\ldots,v_n) \in \mathscr{M}^n$ associe une valeur de vérité $[\![p]\!]^{\mathscr{M}}(v_1,\ldots,v_n) \in \{0;1\}$.

Si les symboles de fonction et de prédicat de la signature Σ sont interprétés par les fonctions $[\![f]\!]^{\mathscr{M}}$ et $[\![p]\!]^{\mathscr{M}}$ qui sont données avec la Σ-structure $\mathscr{M}$, les variables sont quant à elles interprétées à l'aide de valuations :

Définition 15 (Valuations) — *Étant donnée un ensemble $\mathscr{M}$ non vide, on appelle* valuation *sur $\mathscr{M}$ toute fonction $\rho : \mathcal{V} \to \mathscr{M}$ associant un objet $\rho(x) \in \mathscr{M}$ à chaque variable $x \in \mathcal{V}$. L'ensemble des valuations sur $\mathscr{M}$ est noté* $\mathrm{Val}_{\mathscr{M}}$.

Étant donnée une valuation $\rho \in \mathrm{Val}_{\mathscr{M}}$, une variable x et un élément $v \in \mathscr{M}$, on note $(\rho, x \leftarrow v)$ la valuation obtenue en remplaçant dans la valuation ρ la valeur associée à la variable x par la nouvelle valeur v (l'ancienne valeur étant « écrasée »), les valeurs associées aux autres variables restent inchangées. Formellement, la valuation $(\rho, x \leftarrow v)$ est définie pour toute variable y par

$$(\rho, x \leftarrow v)(y) \ = \ \begin{cases} v & \text{si } y = x \\ \rho(y) & \text{sinon} \end{cases}$$

4.1.2 Interprétation des termes

Étant données une signature Σ et une Σ-structure $\mathscr{M}$, on associe à chaque terme t formé sur la signature Σ une *fonction d'interprétation*

$$[\![t]\!]^{\mathscr{M}} \ : \ \mathrm{Val}_{\mathscr{M}} \to \mathscr{M}$$

qui à chaque valuation $\rho \in \mathrm{Val}_{\mathscr{M}}$ associe un élément $[\![t]\!]_\rho \in \mathscr{M}$ qu'on appelle la *dénotation* du terme t dans $\mathscr{M}$ (relativement à la valuation ρ). Formellement, la fonction d'interprétation associée à t est définie par récurrence sur la structure de t à l'aide des équations suivantes :

$$\begin{aligned} [\![x]\!]^{\mathscr{M}}_\rho \ &= \ \rho(x) \\ [\![f(t_1,\ldots,t_n)]\!]^{\mathscr{M}}_\rho \ &= \ [\![f]\!]^{\mathscr{M}}\big([\![t_1]\!]^{\mathscr{M}}_\rho,\ldots,[\![t_n]\!]^{\mathscr{M}}_\rho\big) \end{aligned}$$

Il est facile de vérifier que la dénotation $[\![t]\!]_\rho$ d'un terme t ne dépend que des valeurs prises par la valuation ρ sur l'ensemble des variables libres de t :

Lemme 4 — *Soient $\mathscr{M}$ une Σ-structure et t un terme construit sur la signature Σ. Pour tout couple de valuations $\rho_1, \rho_2 \in \mathrm{Val}_{\mathscr{M}}$ telles que $\rho_1(x) = \rho_2(x)$ pour tout $x \in FV(t)$, on a $[\![t]\!]^{\mathscr{M}}_{\rho_1} = [\![t]\!]^{\mathscr{M}}_{\rho_2}$.*

En particulier, la dénotation $[\![t]\!]_\rho$ d'un terme clos t ne dépend pas de la valuation ρ : on la note alors plus simplement $[\![t]\!]$.

4.1.3 Interprétation des formules

Étant données une signature Σ et une Σ-structure $\mathscr{M}$, on associe à chaque formule A formée sur la signature Σ une *fonction d'interprétation*

$$[\![A]\!]^{\mathscr{M}} \ : \ \mathrm{Val}_{\mathscr{M}} \to \{0;1\}$$

qui à chaque valuation $\rho \in \mathrm{Val}_{\mathscr{M}}$ associe un élément $[\![A]\!]_\rho \in \{0;1\}$ qu'on appelle la *valeur de vérité* de la formule A dans $\mathscr{M}$ (relativement à la valuation ρ). Formellement, la fonction d'interprétation associée à A est définie par récurrence sur la structure de A à l'aide des équations suivantes :

$$\begin{array}{rcl}
[\![\top]\!]^{\mathscr{M}}_\rho & = & 1 \\
[\![\bot]\!]^{\mathscr{M}}_\rho & = & 0 \\
[\![p(t_1,\ldots,t_n)]\!]^{\mathscr{M}}_\rho & = & [\![p]\!]^{\mathscr{M}}\big([\![t_1]\!]^{\mathscr{M}}_\rho,\ldots,[\![t_n]\!]^{\mathscr{M}}_\rho\big) \\
[\![\neg A]\!]^{\mathscr{M}}_\rho & = & 1 - [\![A]\!]^{\mathscr{M}}_\rho \\
[\![A \wedge B]\!]^{\mathscr{M}}_\rho & = & \min\big([\![A]\!]^{\mathscr{M}}_\rho,\ [\![B]\!]^{\mathscr{M}}_\rho\big) \\
[\![A \vee B]\!]^{\mathscr{M}}_\rho & = & \max\big([\![A]\!]^{\mathscr{M}}_\rho,\ [\![B]\!]^{\mathscr{M}}_\rho\big) \\
[\![A \Rightarrow B]\!]^{\mathscr{M}}_\rho & = & \max\big(1 - [\![A]\!]^{\mathscr{M}}_\rho,\ [\![B]\!]^{\mathscr{M}}_\rho\big) \\
[\![\forall x\, A]\!]^{\mathscr{M}}_\rho & = & \inf\limits_{v \in \mathscr{M}} [\![A]\!]^{\mathscr{M}}_{(\rho;x\leftarrow v)} \\
[\![\exists x\, A]\!]^{\mathscr{M}}_\rho & = & \sup\limits_{v \in \mathscr{M}} [\![A]\!]^{\mathscr{M}}_{(\rho;x\leftarrow v)}
\end{array}$$

Cette définition appelle plusieurs remarques :

1. Les connecteurs $\top$, $\bot$, $\neg$, $\wedge$ et $\vee$ s'interprètent à l'aide de leurs tables de vérité habituelles : on notera que la conjonction correspond à un calcul de minimum, la disjonction à un calcul de maximum, et la négation à un échange des valeurs 0 et 1.
2. L'implication $A \Rightarrow B$ est interprétée comme la formule $\neg A \vee B$. Là encore on retrouve la table de vérité attendue, puisque la formule $A \Rightarrow B$ prend la valeur 0 uniquement dans le cas où A prend la valeur 1 et B la valeur 0 (dans tous les autres cas, la formule $A \Rightarrow B$ prend la valeur 1).
3. La quantification universelle et la quantification existentielle s'interprètent respectivement comme une conjonction et une disjonction infinitaires (du moins lorsque $\mathscr{M}$ est infini). En pratique, ce sont ces deux constructions qui sont responsables du fait que dans le cas où $\mathscr{M}$ est infini, la fonction $A \mapsto [\![A]\!]^{\mathscr{M}}_\rho$ qui à une formule associe sa dénotation n'est pas calculable.[1]

Là encore, la dénotation $[\![A]\!]_\rho$ d'une formule A ne dépend que des valeurs prises par la valuation ρ sur l'ensemble des variables libres de A :

Lemme 5 — *Soient $\mathscr{M}$ une Σ-structure et A une formule construite sur la signature Σ. Pour tout couple de valuations $\rho_1, \rho_2 \in \mathrm{Val}_{\mathscr{M}}$ telles que $\rho_1(x) = \rho_2(x)$ pour tout $x \in FV(A)$, on a $[\![A]\!]^{\mathscr{M}}_{\rho_1} = [\![A]\!]^{\mathscr{M}}_{\rho_2}$.*

En particulier, la dénotation $[\![A]\!]_\rho$ d'une formule close A ne dépend pas de la valuation ρ : on la note alors plus simplement $[\![A]\!]$.

[1] Dans le cas où la structure $\mathscr{M}$ est finie, la complexité du calcul est exponentielle par rapport au nombre d'imbrications de quantificateurs dans la formule.

Extension aux séquents L'interprétation des formules s'étend aux séquents de la manière suivante : étant données une Σ-structure, une liste de formules $\Gamma = A_1, \ldots, A_n$ et une valuation $\rho \in \mathrm{Val}_{\mathscr{M}}$, on note

$$\begin{array}{rcl} \min[\![\Gamma]\!]^{\mathscr{M}}_{\rho} & = & \min([\![A_1]\!]^{\mathscr{M}}_{\rho}, \ldots, [\![A_n]\!]^{\mathscr{M}}_{\rho}) \\ \max[\![\Gamma]\!]^{\mathscr{M}}_{\rho} & = & \max([\![A_1]\!]^{\mathscr{M}}_{\rho}, \ldots, [\![A_n]\!]^{\mathscr{M}}_{\rho}) \end{array}$$

L'interprétation d'un séquent $\Gamma \vdash \Delta$ est alors définie par

$$[\![\Gamma \vdash \Delta]\!] \;=\; \max(1 - \min[\![\Gamma]\!]^{\mathscr{M}}_{\rho}, \max[\![\Delta]\!]^{\mathscr{M}}_{\rho})$$

(par analogie avec l'implication).

4.1.4 Validité des formules et des séquents

Définition 16 (Validité) — *Soit $\mathscr{M}$ une Σ-structure. On dit qu'une formule A (construite dans Σ) est* valide *dans $\mathscr{M}$ et on note $\mathscr{M} \models A$ lorsque pour toute valuation $\rho \in \mathrm{Val}_{\mathscr{M}}$ on a $[\![A]\!]^{\mathscr{M}}_{\rho} = 1$.*

La notion de validité s'étend aux séquents de manière évidente : ainsi on dit d'un séquent $\Gamma \vdash \Delta$ qu'il est *valide* dans une Σ-structure $\mathscr{M}$ et on note $\mathscr{M} \models \Gamma \vdash \Delta$ si pour toute valuation $\rho \in \mathrm{Val}_{\mathscr{M}}$ on a $[\![\Gamma \vdash \Delta]\!]^{\mathscr{M}}_{\rho} = 1$.

On vérifie alors aisément que tous les séquents dérivables sont valides dans n'importe quelle Σ-structure :

Proposition 17 (Validité) — *Si un séquent $\Gamma \vdash \Delta$ est dérivable (en déduction naturelle classique ou en calcul des séquents classique), alors ce séquent est valide dans n'importe quelle Σ-structure $\mathscr{M}$.*

Preuve. Par induction sur la structure de la dérivation du séquent $\Gamma \vdash \Delta$. □

En particulier, toute formule A qui est une tautologie (intuitionniste ou classique) est valide dans n'importe quelle Σ-structure.

4.2 Modèles

4.2.1 La notion de modèle

Définition 17 — *Soit $\mathscr{T}$ une théorie du premier ordre dont le langage est construit sur la signature Σ. On appelle* modèle *de la théorie $\mathscr{T}$ toute Σ-structure $\mathscr{M}$ dans laquelle tous les axiomes de $\mathscr{T}$ sont valides.*

Une conséquence immédiate de la Prop. 17 est qu'un modèle d'une théorie $\mathscr{T}$ valide non seulement les axiomes de $\mathscr{T}$, mais également tous ses théorèmes :

Proposition 18 (Validité) — *Soient $\mathscr{T}$ une théorie du premier ordre et A une formule. Si $\mathscr{T} \vdash A$, alors $\mathscr{M} \models A$ pour tout modèle $\mathscr{M}$ de la théorie $\mathscr{T}$.*

Si les théorèmes de $\mathscr{T}$ sont valides dans n'importe quel modèle de $\mathscr{T}$, un modèle donné d'une théorie $\mathscr{T}$ valide en général un ensemble de formules closes qui est plus vaste que l'ensemble des théorèmes de $\mathscr{T}$. Typiquement, si A est une formule close *indécidable* dans la théorie $\mathscr{T}$ (c'est-à-dire telle que ni A

ni $\neg A$ n'est un théorème de $\mathscr{T}$[2]) et si $\mathscr{M}$ est un modèle donné de $\mathscr{T}$, alors l'une des deux formules A ou $\neg A$ est validée par le modèle $\mathscr{M}$ bien qu'aucune d'elles ne soit un théorème de $\mathscr{T}$. En revanche, nous verrons dans la suite de ce chapitre que si une formule close A est valide dans *tous les modèles* de $\mathscr{T}$, alors cette formule est démontrable dans $\mathscr{T}$: c'est ce qu'énonce le théorème de *complétude* de la logique du premier ordre que nous établirons à la section 4.4, et qui constitue le principal résultat de ce chapitre.

4.2.2 Modèles des théories égalitaires

Soient $\mathscr{T}$ une théorie égalitaire (Déf. 10) et $\mathscr{M}$ un modèle de la théorie $\mathscr{T}$. Pour tous $v_1, v_2 \in \mathscr{M}$, on note $v_1 =_{\mathscr{M}} v_2$ lorsque $[\![=]\!]^{\mathscr{M}}(v_1, v_2) = 1$. D'après les propriétés de la relation d'égalité dans $\mathscr{T}$ (cf Prop. 12), il est immédiat que la relation binaire $=_{\mathscr{M}}$ est une relation d'équivalence sur $\mathscr{M}$:

1. $v =_{\mathscr{M}} v$
2. $v_1 =_{\mathscr{M}} v_2$ entraîne $v_2 =_{\mathscr{M}} v_1$;
3. $v_1 =_{\mathscr{M}} v_2$ et $v_2 =_{\mathscr{M}} v_3$ entraînent $v_1 =_{\mathscr{M}} v_3$;

Par ailleurs, l'interprétation de chaque symbole de fonction f (resp. de chaque symbole de prédicat p) d'arité n est une fonction compatible avec la relation d'équivalence $v_1 =_{\mathscr{M}} v_2$, en ce sens que :

4. Si $v_i =_{\mathscr{M}} v'_i$ pour tous $1 \le i \le n$, alors

$$\begin{array}{rrcl} & [\![f]\!]^{\mathscr{M}}(v_1, \ldots, v_n) & =_{\mathscr{M}} & [\![f]\!]^{\mathscr{M}}(v_1, \ldots, v_n) \\ (\text{resp.} & [\![p]\!]^{\mathscr{M}}(v_1, \ldots, v_n) & = & [\![p]\!]^{\mathscr{M}}(v_1, \ldots, v_n)) \end{array}$$

(où le symbole $=$ de la seconde ligne désigne l'égalité des valeurs de vérité).

Bien que ces propriétés expriment que la relation d'équivalence $v_1 =_{\mathscr{M}} v_2$ se comporte « presque » comme la relation d'égalité, cette relation n'est en général pas la relation d'égalité $v_1 = v_2$. En pratique, il est plus simple de supposer que la relation $v_1 =_{\mathscr{M}} v_2$ est la relation d'égalité $v_1 = v_2$ en ne travaillant qu'avec une forme particulière de modèle de $\mathscr{T}$: les modèles égalitaires.

Définition 18 (Modèle égalitaire) — *On dit qu'un modèle $\mathscr{M}$ d'une théorie égalitaire $\mathscr{T}$ est un* modèle égalitaire *de $\mathscr{T}$ si la relation d'équivalence $=_{\mathscr{M}}$ est la relation d'égalité, i.e. $v_1 =_{\mathscr{M}} v_2$ ssi $v_1 = v_2$ pour tous $v_1, v_2 \in \mathscr{M}$.*

En pratique, il est toujours possible de se ramener à un modèle égalitaire en procédant de la manière suivante : si $\mathscr{M}$ est un modèle (pas forcément égalitaire) de la théorie T, on considère l'ensemble quotient $\mathscr{M}' = \mathscr{M}/{=_{\mathscr{M}}}$ que l'on transforme en Σ-structure en posant

$$\begin{array}{rcll} [\![f]\!]^{\mathscr{M}'}([v_1], \ldots, [v_n]) & = & [[\![f]\!]^{\mathscr{M}}(v_1, \ldots, v_n)] & \in \mathscr{M}' \\ [\![p]\!]^{\mathscr{M}'}([v_1], \ldots, [v_n]) & = & [\![p]\!]^{\mathscr{M}}(v_1, \ldots, v_n) & \in \{0; 1\} \end{array}$$

[2]Le premier théorème d'incomplétude de Gödel établit que toute théorie cohérente permettant d'exprimer (au moins) l'arithmétique de Peano et dont l'ensemble des axiomes est récursivement énumérable admet des formules closes indécidables. Grâce au procédé de Gödel, on sait ainsi construire des formules closes indécidables dans l'arithmétique de Peano, dans la théorie des ensembles de Zermelo-Fraenkel, etc. (l'indécidabilité de ces formules étant bien entendu conditionnée par la cohérence de la théorie).

pour tout symbole de fonction f d'arité n, pour tout symbole de prédicat p d'arité n et pour tous $v_1, \ldots, v_n \in \mathscr{M}$, en notant $[v]$ la classe d'équivalence de l'élément $v \in \mathscr{M}$ dans l'ensemble quotient $\mathscr{M}' = \mathscr{M}/{=_{\mathscr{M}}}$.

On vérifie alors aisément que

Proposition 19 (Modèle quotient) — *La Σ-structure $\mathscr{M}'$ déduite du modèle $\mathscr{M}$ par passage au quotient est un modèle égalitaire de la théorie $\mathscr{T}$.*

Dans tout ce qui suit, on ne considérera donc que des modèles égalitaires.

4.3 Existence de modèles

4.3.1 Modèles et cohérence

Une conséquence importante de la Prop. 18 est la suivante :

Proposition 20 (Cohérence par existence de modèle) — *Toute théorie qui admet un modèle est cohérente.*

Preuve. Soit $\mathscr{T}$ une théorie admettant un modèle $\mathscr{M}$. Comme par définition de la fonction d'interprétation des formules on a $[\![\bot]\!]_{\mathscr{M}} = 0$, il ressort de la Prop. 18 que la formule $\bot$ n'est pas un théorème de $\mathscr{T}$. $\square$

Dans cette section, on se propose d'établir la réciproque, à savoir que toute théorie cohérente admet au moins un modèle. Pour cela, on utilisera le résultat suivant :

Lemme 6 — *Si Σ est une signature finie ou dénombrable, alors :*

1. *L'ensemble des termes ouverts construits dans Σ est dénombrable, et l'ensemble des termes clos est fini ou dénombrable.*
2. *L'ensemble des formules closes construites dans Σ est dénombrable.*

Preuve. À tout entier $k \geq 0$ on associe un ensemble de termes T_k défini par récurrence sur k en posant :

- $T_0 = \varnothing$
- Pour tout $k > 0$, T_k est défini comme l'ensemble des termes qui sont soit une variable, soit un terme de la forme $f(t_1, \ldots, t_n)$, où $f \in \Sigma$ est un symbole de fonction arbitraire (d'arité n), et où $t_1, \ldots, t_n \in T_{k-1}$.

On vérifie alors aisément que

1. Pour tout $k \geq 0$, $T_k \subseteq T_{k+1}$ (par récurrence sur k). D'où il ressort que la suite d'ensembles $(T_k)_{k\in\mathbb{N}}$ est croissante.
2. Tout terme t formé sur la signature Σ appartient à l'un au moins des ensembles T_k (par récurrence sur la structure de t). D'où il ressort que l'union $T_\infty = \bigcup_k T_k$ est précisément l'ensemble de tous les termes.
3. Pour tout $k > 0$, l'ensemble T_k est dénombrable (par récurrence sur k, en utilisant les propriétés des ensembles dénombrables). Par conséquent, l'union $T_\infty = \bigcup_k T_k$ est dénombrable.

En particulier, l'ensemble des termes clos, qui est un sous-ensemble de T_∞, est fini ou dénombrable.[3]

À l'aide d'une construction analogue (basée sur une suite croissante d'ensembles de formules $(F_k)_{k\geq 0}$), on démontre que l'ensemble des formules ouvertes est dénombrable. Contrairement aux termes, l'ensemble des formules closes est forcément dénombrable puisqu'il contient au moins la suite infinie de formules $\top$, $\top \wedge \top$, $\top \wedge \top \wedge \top$, etc. $\square$

4.3.2 Complétion d'une théorie

Définition 19 (Complétude) — *On dit qu'une théorie $\mathscr{T}$ est* complète *si pour toute formule close A du langage de $\mathscr{T}$, l'une (au moins) des deux formules A ou $\neg A$ est un théorème de $\mathscr{T}$.*

Toute théorie incohérente est complète de manière triviale. Inversement, le premier théorème d'incomplétude de Gödel stipule que toute théorie cohérente récursivement énumérable et contenant l'arithmétique de Peano est nécessairement incomplète. Une autre forme de complétude est la suivante :

Définition 20 (Saturation de témoins) — *On dit qu'une théorie $\mathscr{T}$ est* saturée de témoins *(ou complète au sens de Henkin) si pour tout théorème de $\mathscr{T}$ de la forme $\exists x\, A(x)$, il existe un terme clos t_0 tel que $\mathscr{T} \vdash A(t_0)$.*

L'ingrédient essentiel du théorème de complétude est le résultat suivant :

Proposition 21 (Complétion de théorie) — *Toute théorie cohérente admet une extension cohérente, complète et saturée de témoins. Si en outre la théorie initiale a une signature finie ou dénombrable, on peut construire l'extension recherchée de telle sorte qu'elle ait une signature dénombrable.*

Preuve. On ne fait ici la démonstration de ce résultat que dans le cas où la signature est finie ou dénombrable.[4] Supposons donc que $\mathscr{T}$ est une théorie cohérente dont le langage est formé sur une signature Σ finie ou dénombrable. Considérons un ensemble dénombrable $\mathcal{C}$ de « constantes fraîches » (c'est à dire non utilisées dans la signature Σ) et notons $\mathscr{T}_0$ la théorie obtenue en étendant le langage de $\mathscr{T}$ avec les constantes de l'ensemble $\mathcal{C}$ (sans ajouter d'axiome supplémentaire). D'après la Prop. 14, $\mathscr{T}_0$ est une extension conservative de $\mathscr{T}$, et donc une théorie cohérente. Comme la signature de $\mathscr{T}_0$ (à savoir : $\Sigma \cup \mathcal{C}$) est dénombrable, l'ensemble des formules closes de $\mathscr{T}_0$ est dénombrable. Considérons donc une énumération $(A_i)_{i\in\mathbb{N}}$ de l'ensemble des formules closes de $\mathscr{T}_0$.

Par récurrence sur $i \in \mathbb{N}$, nous allons construire une suite infinie de théories emboîtées

$$\mathscr{T}_0 \subseteq \mathscr{T}_1 \subseteq \mathscr{T}_2 \subseteq \cdots \subseteq \mathscr{T}_i \subseteq \mathscr{T}_{i+1} \subseteq \cdots$$

telles que pour tout $i \in \mathbb{N}$ (1) la théorie $\mathscr{T}_i$ est cohérente, et (2) les axiomes de $\mathscr{T}_i$ n'utilisent qu'un nombre fini de constantes de $\mathcal{C}$. On notera que la théorie $\mathscr{T}_0$,

[3] Ce sous ensemble peut être fini, voire vide dans le cas où la signature ne comporte aucun symbole de constante (comme c'est le cas dans la théorie des ensembles).

[4] La démonstration dans le cas général suit exactement la même technique. La seule difficulté vient du fait que dans le cas où le langage est non dénombrable, il est nécessaire d'indexer les formules closes par des ordinaux plutôt que par des entiers, ce qui fait reposer la démonstration sur des techniques d'induction transfinie qui dépassent le cadre de ce cours.

que nous avons déjà construite, vérifie ces deux propriétés. La suite de théories $(\mathscr{T}_i)_{i\in\mathbb{N}}$ est construite de proche en proche en passant de $\mathscr{T}_i$ à $\mathscr{T}_{i+1}$ en introduisant une théorie intermédiaire $\mathscr{T}_{i+1/2}$.

Passage de $\mathscr{T}_i$ ***à*** $\mathscr{T}_{i+1/2}$ Dans le cas où $\mathscr{T}_i$ décide la formule A_i (c'est-à-dire dans le cas où $\mathscr{T}_i \vdash A_i$ ou $\mathscr{T}_i \vdash \neg A_i$), on pose simplement $\mathscr{T}_{i+1/2} = \mathscr{T}_i$. Sinon, on pose $\mathscr{T}_{i+1/2} = \mathscr{T}_i + A_i$ (c'est-à-dire : $\mathscr{T}_i$ étendue avec l'axiome A_i). Dans tous les cas, la théorie $\mathscr{T}_{i+1/2}$ est cohérente et décide la formule A_i.

Passage de $\mathscr{T}_{i+1/2}$ ***à*** $\mathscr{T}_{i+1}$ Dans le cas où A_i n'est pas une formule existentielle, on pose $\mathscr{T}_{i+1} = \mathscr{T}_{i+1/2}$. Sinon, dans le cas où $A_i \equiv \exists x\, B(x)$, on choisit une constante c qui n'apparaît pas dans les axiomes de $\mathscr{T}_{i+1/2}$ et on pose $\mathscr{T}_{i+1} = \mathscr{T}_{i+1/2} + B(c)$ (c'est-à-dire : $\mathscr{T}_{i+1/2}$ étendue avec l'axiome $B(c)$). D'après les propositions 14 et 15, la théorie $\mathscr{T}_{i+1}$ est une extension conservative de la théorie $\mathscr{T}_{i+1/2}$, et donc une théorie cohérente.

Soit $\mathscr{T}_\infty = \bigcup_{i\in\mathbb{N}} \mathscr{T}_i$ la réunion des théories $\mathscr{T}_i$ ($i \in \mathbb{N}$), c'est-à-dire la théorie dont la signature est $\Sigma \cup \mathcal{C}$ et dont l'ensemble des axiomes est la réunion des ensembles des axiomes des théories $\mathscr{T}_i$. On vérifie alors facilement que :

1. La théorie $\mathscr{T}_\infty$ est une extension de $\mathscr{T}$. (Par construction.)
2. La théorie $\mathscr{T}_\infty$ est cohérente. En effet, toute démonstration dans $\mathscr{T}_\infty$ n'utilise qu'un nombre fini d'axiomes de $\mathscr{T}_\infty$, et se place donc dans l'une des théories $\mathscr{T}_i$ ($i \in \mathbb{N}$). L'incohérence de $\mathscr{T}_\infty$ induirait donc l'incohérence d'une des théories $\mathscr{T}_i$, ce qui est impossible.
3. La théorie $\mathscr{T}_\infty$ est complète. En effet, toute formule close A_i du langage de $\mathscr{T}_\infty$ est décidée dans la sous-théorie $\mathscr{T}_{i+1} \subseteq \mathscr{T}_\infty$.
4. La théorie $\mathscr{T}_\infty$ est saturée de témoins. En effet, pour toute formule close A_i de la forme $A_i \equiv \exists x B(x)$ qui est un théorème de $\mathscr{T}_\infty$, on a $\mathscr{T}_{i+1} \vdash \exists x B(x)$ (le contraire induirait une contradiction). Par construction de $\mathscr{T}_{i+1}$, il existe une constante $c \in \mathcal{C}$ telle que $\mathscr{T}_{i+1} \vdash B(c)$.

Enfin, il est clair que la signature et le langage de $\mathscr{T}_\infty$ sont dénombrables. $\square$

L'intérêt de la notion de théorie complète et saturée de témoins est que toute théorie de cette nature se transforme très facilement en un modèle syntaxique d'elle-même. En effet, étant donnée une théorie $\mathscr{T}$ de signature Σ qui est à la fois complète et saturée de témoins, on considère la Σ-structure $\mathscr{M}$ définie de la manière suivante :

- Le domaine de $\mathscr{M}$ est l'ensemble des termes clos du langage de $\mathscr{T}$.
- Chaque symbole de fonction f d'arité n est interprété par lui-même, c'est-à-dire par la fonction qui à tout n-uplet de termes clos $(t_1, \ldots, t_n) \in \mathscr{M}^n$ associe le terme clos $f(t_1, \ldots, t_n) \in \mathscr{M}$.
- Chaque symbole de prédicat p d'arité n est interprété par la fonction de vérité $p^{\mathscr{M}} : \mathscr{M}^n \to \{0; 1\}$ définie pour tout $(t_1, \ldots, t_n) \in \mathscr{M}^n$ par

$$p^{\mathscr{M}}(t_1, \ldots, t_n) \;=\; \begin{cases} 1 & \text{si } \mathscr{T} \vdash p(t_1, \ldots, t_n) \\ 0 & \text{si } \mathscr{T} \vdash \neg p(t_1, \ldots, t_n) \end{cases}$$

(la définition est complète en raison de la complétude de $\mathscr{T}$).

On notera que dans cette Σ-structure, une valuation se comporte comme une substitution destinée à clore les termes comme les formules. Ainsi, si ρ désigne une valuation et t un terme ouvert, on note $t[\rho]$ le terme clos obtenu en remplaçant dans t chaque occurrence libre d'une variable $x \in FV(t)$ par le terme clos $\rho(x)$. De même si A désigne une formule ouverte, on note $A[\rho]$ la formule close obtenue en remplaçant dans A chaque occurrence libre d'une variable $x \in FV(A)$ par le terme clos $\rho(x)$.

Ces notations étant précisées, on vérifie aisément que les formules valides dans $\mathscr{M}$ coïncident avec les formules démontrables dans $\mathscr{T}$:

Lemme 7 (Validité dans $\mathscr{M}$) — *Pour toute formule A du langage de $\mathscr{T}$ et pour toute valuation ρ, on a*

$$[\![A]\!]^{\mathscr{M}}_{\rho} \;=\; \begin{cases} 1 & \textit{si } \mathscr{T} \vdash A[\rho] \\ 0 & \textit{si } \mathscr{T} \vdash \neg A[\rho] \end{cases}$$

Preuve. Par induction sur la structure de la formule A. Dans le cas où la formule A est formée à l'aide d'un connecteur (unaire ou binaire), on utilise la complétude de $\mathscr{T}$ pour montrer que les cas où la formule A est prouvable sont donnés par la table de vérité de ce connecteur. Dans le cas où la formule A est formée à l'aide d'une quantification (universelle ou existentielle), on utilise la propriété de saturation de témoins pour conclure. □

De ce lemme il ressort de manière immédiate que $\mathscr{M}$ est un modèle de la théorie $\mathscr{T}$ puisqu'il valide toutes les formules démontrables dans $\mathscr{T}$. Ce qui nous permet de déduire le résultat suivant :

Proposition 22 (Existence d'un modèle) — *Toute théorie cohérente admet au moins un modèle. Si en outre la signature de la théorie est finie ou dénombrable, on peut faire en sorte que le modèle soit dénombrable.*

Preuve. Soit $\mathscr{T}$ une théorie cohérente. D'après la Prop. 21, la théorie $\mathscr{T}$ admet une extension $\mathscr{T}' \supseteq \mathscr{T}$ qui est complète et saturée de témoins. Et d'après ce qui précède, la théorie $\mathscr{T}'$ induit un modèle syntaxique (cf la construction ci-dessus) qui constitue également un modèle de $\mathscr{T}$, modèle qui est en outre dénombrable dans le cas où la signature de $\mathscr{T}$ est finie ou dénombrable. □

4.3.3 Le paradoxe de Löwenheim-Skolem

La proposition 22 fait apparaître une situation paradoxale[5] car elle implique que même la théorie de Zermelo-Fraenkel — qui décrit pourtant un univers fortement non dénombrable — admet elle aussi un modèle dénombrable (du moins si cette théorie est cohérente).

Pour comprendre cette bizarrerie connue des logiciens sous le nom de *paradoxe de Löwenheim-Skolem*, considérons un modèle dénombrable $\mathscr{M}$ de la théorie des ensembles de Zermelo-Fraenkel. Quitte à effectuer le quotient idoine (décrit en 3.2.2) on peut supposer sans perte de généralité que le modèle $\mathscr{M}$ est égalitaire (cf Déf. 18) — ce qui ne change rien à sa cardinalité[6] — et on note $\in_{\mathscr{M}}$ la dénotation de la relation d'appartenance dans le modèle $\mathscr{M}$.

Examinons à présent comment sont représentés les entiers, mais aussi les ensembles $\mathbb{N}$ et $\mathfrak{P}(\mathbb{N})$ dans le modèle $\mathscr{M}$.

[5]Il s'agit d'un *paradoxe* au sens étymologique du terme (à savoir : un phénomène « contraire à l'opinion »), et non d'une *contradiction* formelle (une antinomie).

[6]Le quotient peut transformer un modèle dénombrable en un modèle fini, mais il est facile de vérifier que ce cas ne peut pas se produire pour ZF, dont tous les modèles sont infinis.

Représentation des entiers Chaque entier naturel $n \in \mathbb{N}$ peut être caractérisé par une formule de la théorie des ensembles $A_n(x)$ à une variable libre, qui n'est vraie que pour cet entier. On peut définir par récurrence sur n une telle formule en posant

$$\begin{array}{rcl} A_0(x) & \equiv & \forall z\ z \notin x \\ A_{n+1}(x) & \equiv & \exists y\ \big(A_n(y)\ \wedge\ \forall z\ (z \in x \Leftrightarrow z \in y \vee z = y)\big) \end{array}$$

Intuitivement, $A_0(x)$ exprime que « x est l'ensemble vide », tandis que $A_{n+1}(x)$ exprime que « x est le successeur de l'ensemble y tel que $A_n(y)$ ».

Comme pour chaque entier $n \in \mathbb{N}$ la formule $\exists! x\, A_n(x)$ est prouvable dans ZF, il existe dans le modèle $\mathscr{M}$ un unique objet $\dot{n}$ tel que $[\![A_n(x)]\!]_{x \leftarrow \dot{n}} = 1$. Cet objet est (intuitivement) le représentant de l'entier $n \in \mathbb{N}$, et on le qualifie d'entier standard (dans le modèle).

Représentation de l'ensemble des entiers De même, l'ensemble des entiers naturels peut-être caractérisé par une formule $A_{\mathbb{N}}(x)$ de la théorie des ensembles, qui exprime que « x est l'ensemble de tous les entiers naturels ». Là encore, la formule $\exists! x\, A_{\mathbb{N}}(x)$ est prouvable dans ZF, et il existe donc un unique objet $\dot{\mathbb{N}} \in \mathscr{M}$ tel que $[\![A_{\mathbb{N}}(x)]\!]_{x \leftarrow \dot{\mathbb{N}}} = 1$. Intuitivement, l'objet $\dot{\mathbb{N}}$ est le représentant de l'ensemble des entiers dans le modèle $\mathscr{M}$

Comme pour tout entier $n \in \mathbb{N}$ la formule

$$\forall x\ \forall y\ (A_n(x) \wedge A_{\mathbb{N}}(y) \Rightarrow x \in y)$$

est prouvable dans ZF, on a au niveau du modèle la relation $\dot{n} \in_{\mathscr{M}} \dot{\mathbb{N}}$ pour chaque entier $n \in \mathbb{N}$. Autrement dit : dans le modèle $\mathscr{M}$, la dénotation $\dot{\mathbb{N}}$ contient (au sens de la relation $\in_{\mathscr{M}}$) tous les entiers standards. En revanche, il se peut que le point $\dot{\mathbb{N}} \in \mathscr{M}$ contienne également (au sens de la relation $\in_{\mathscr{M}}$) d'autres objets. Ces objets, lorsqu'ils existent, sont appelés des *entiers non standards*, et on dit que le modèle $\mathscr{M}$ est un modèle non standard de la théorie des ensembles.

Pour simplifier les choses, on supposera dans la suite de la discussion que le modèle $\mathscr{M}$ est standard (et donc que tous les entiers du modèle sont standards), ce qui nous permettra d'identifier l'ensemble des entiers naturels avec l'ensemble des entiers dans le modèle.

Représentation de l'ensemble des parties de $\mathbb{N}$ Passons à présent à la représentation de l'ensemble non dénombrable $R = \mathfrak{P}(\mathbb{N})$ dans le modèle. Là encore, cet ensemble est caractérisé par la formule $A_R(x)$ définie par

$$A_R(x)\ \equiv\ \exists y\ \big(A_{\mathbb{N}}(y)\ \wedge\ \forall z\ (z \in x \Leftrightarrow z \subseteq y)\big)\,,$$

et comme la formule $\exists! x\, A_R(x)$ est prouvable dans ZF, il existe un unique objet $\dot{R} \in \mathscr{M}$ tel que $[\![A_R(x)]\!]_{x \leftarrow \dot{R}} = 1$. De même que chaque entier standard « appartient » (au sens du modèle) à la dénotation $\dot{\mathbb{N}}$, on vérifie que chaque objet $p \in_{\mathscr{M}} \dot{R}$ est « inclus » dans $\dot{\mathbb{N}}$, en ce sens que $z \in_{\mathscr{M}} p$ entraîne $z \in_{\mathscr{M}} \dot{\mathbb{N}}$ pour tout objet $z \in \mathscr{M}$. Autrement dit, les « éléments » de $\dot{R}$ sont effectivement des parties de $\dot{\mathbb{N}}$ dans le modèle, c'est-à-dire des ensembles d'entiers (en supposant le modèle standard).

On a donc construit une fonction ϕ qui à chaque partie de $\mathbb{N}$ *au sens du modèle* $\mathscr{M}$ associe une partie de $\mathbb{N}$:

$$\begin{array}{rcl} \phi : \dot{R} & \to & \mathfrak{P}(\mathbb{N}) \\ p & \mapsto & \{n \in \mathbb{N} \mid \dot{n} \in_{\mathscr{M}} p\} \end{array}$$

(en identifiant $\dot{R}$ avec l'ensemble des objets $p \in_{\mathscr{M}} \dot{R}$). Cette fonction qui envoie les parties de $\mathbb{N}$ au sens du modèle dans les parties de $\mathbb{N}$ est injective, en raison du fait que le modèle $\mathscr{M}$ est standard (et extensionnel). En revanche, la fonction ϕ ne peut pas être surjective, précisément parce que le modèle $\mathscr{M}$ est dénombrable.

Toute cette discussion montre que dans un modèle dénombrable de ZF, il n'y a qu'un nombre dénombrable de parties de $\mathbb{N}$ qui sont représentées, ce qui résout le paradoxe. On peut aussi comprendre ce phénomène comme l'expression du fait que d'un point de vue strictement syntaxique, il n'est possible d'exprimer qu'un nombre dénombrable de parties de $\mathbb{N}$, précisément parce que le langage est dénombrable. Intuitivement, les autres parties de $\mathbb{N}$, celles dont on ne « sait » pas parler, peuvent être simplement omises dans le modèle.

On peut également noter que puisque l'ensemble des éléments de $\dot{R}$ (au sens du modèle) est dénombrable, on peut construire une bijection f entre les ensembles $\dot{R}$ et $\dot{\mathbb{N}}$. Là encore, ceci ne contredit pas le fait que la formule

« Il n'existe pas de bijection entre $R = \mathfrak{P}(\mathbb{N})$ et $\mathbb{N}$ »

est prouvable dans ZF. Simplement, la bijection f qui relie les objets $\dot{R}$ et $\dot{\mathbb{N}}$ n'est elle-même pas un objet du modèle ; il s'agit d'une bijection « externe », qui ne correspond à aucune bijection de ZF vue à travers le modèle.

4.4 Complétude

4.4.1 Le théorème de complétude

La Prop. 22 nous permet maintenant d'établir le théorème de complétude :

Théorème 1 (Complétude) — *Soient $\mathscr{T}$ une théorie du premier ordre et A une formule close du langage de $\mathscr{T}$. Si $\mathscr{M} \models A$ pour tout modèle $\mathscr{M}$ de la théorie $\mathscr{T}$, alors $\mathscr{T} \vdash A$.*

Preuve. En raisonnant par l'absurde, supposons que la formule close A n'est pas démontrable dans $\mathscr{T}$. Dans ce cas, la théorie $\mathscr{T}' = \mathscr{T} + \neg A$ est cohérente et admet donc un modèle $\mathscr{M}$ d'après la Prop. 22. Mais ce modèle $\mathscr{M}$ est aussi un modèle de $\mathscr{T}$, qui invalide la formule A. Ce qui prouve l'existence d'un modèle de $\mathscr{T}$ dans lequel A est fausse. □

4.4.2 Conservativité des extensions de Skolem

Nous sommes à présent en mesure d'établir la conservativité des extensions de Skolem (Déf. 13) énoncée en 3.3.3, p. 35 :

Proposition 16 (Conservativité de Skolem) — *Toute extension de Skolem est une extension conservative.*

Preuve. Soit $\mathscr{T}$ une théorie et T un théorème de $\mathscr{T}$ de la forme

$$T \equiv \forall x_1 \cdots \forall x_n \, \exists y \, A(x_1, \ldots, x_n, y) \, .$$

L'extension de Skolem correspondante est la théorie $\mathscr{T}'$ obtenue en ajoutant à $\mathscr{T}$ un nouveau symbole de fonction f d'arité n avec l'axiome

$$T' \equiv \forall x_1 \cdots \forall x_n \, A(x_1, \ldots, x_n, f(x_1, \ldots, x_n)) \, .$$

Supposons à présent que B est une formule close du langage de $\mathscr{T}$ qui est un théorème de $\mathscr{T}'$. Pour montrer que B est également un théorème de $\mathscr{T}$, il suffit de montrer (en utilisant le théorème de complétude) que la formule B est valide dans tous les modèles de $\mathscr{T}$. Pour cela, on se fixe un modèle $\mathscr{M}$ de $\mathscr{T}$, et pour chaque $(v_1, \ldots, v_n) \in \mathscr{M}^n$, on note $\phi(v_1, \ldots, v_n)$ un élément de $\mathscr{M}$ tel que

$$[\![A(x_1, \ldots, x_n, y)]\!]^{\mathscr{M}}_{x_1 \leftarrow v_1, \ldots, x_n \leftarrow v_n, y \leftarrow \phi(v_1, \ldots, v_n)} = 1$$

(un tel élément existe pour chaque $(v_1, \ldots, v_n) \in \mathscr{M}^n$ dans la mesure où la formule T, qui est un théorème de $\mathscr{T}$, est valide dans $\mathscr{M}$). On étend alors $\mathscr{M}$ en un modèle $\mathscr{M}'$ de la théorie $\mathscr{T}'$ en interprétant le nouveau symbole de fonction f par $[\![f]\!]^{\mathscr{M}} = \phi$ (les autres symboles étant interprétés de la même manière que dans $\mathscr{M}$). Par construction, $\mathscr{M}'$ est un modèle de $\mathscr{T}'$ et valide donc la formule B. Mais comme les symboles qui figurent dans la formule B sont interprétés de la même manière dans $\mathscr{M}$ et $\mathscr{M}'$, la valeur de vérité de B est la même dans les deux modèles. Par conséquent, la formule B est valide dans $\mathscr{M}$ également. $\square$

Table des matières