Cours de Logique

Titouan Leclercq et Werner Mérian

Cours de logique

Titouan Leclercq
titouan.leclercq@ens-lyon.fr

Werner Mérian
werner.merian99@gmail.com

Printemps 2024

Version compilée le 23 février 2024 à 19 h 20 min 40 s

Nanos gigantum umeris insidentes

Des nains sur des épaules de géants

Table des matières

Introduction générale

Il est intéressant, en étudiant la logique, de voir combien ce domaine est difficile à situer : à l'origine une part de la philosophie, puis devenue plus tard une branche mathématique, elle est de nos jours omniprésente en informatique. Autant dire que ce livre est bien trop court pour vous donner, cher lecteur, une vision réunissant tous ces domaines et se voulant complète. Les auteurs étant principalement rodés à la logique mathématique et à l'informatique, ce livre portera largement sur ces deux visions : la logique sera un outil formel d'analyse du discours mathématique en premier lieu. Cependant, il est important de rappeler que ce domaine n'est pas exempt de controverses, loin s'en faut. Ces controverses sont rarement de nature mathématique, et encore moins de nature informatique : elles appartiennent pleinement à la philosophie.

Il est donc nécessaire d'accepter dès le début de ce livre que, dans un objectif de pédagogie, et puisque les controverses peuvent nuire dans un premier temps à la compréhension de certaines notions, des choix d'ordre philosophique seront régulièrement pris au long de cet ouvrage. Lorsque cela arrivera, la position qui sera tenue sera argumentée si possible, et le lecteur est libre de ne pas adhérer à l'interprétation qui sera donnée de certains phénomènes. Le contenu du livre, lui, se trouve avant tout dans la compréhension des objets étudiés et dans les résultats démontrés.

Pour commencer, qu'est-ce que la logique ? Au sens philosophique, cela désigne l'étude du raisonnement, mais son utilisation sur les mathématiques permet d'être plus précis. La logique est l'étude du langage mathématique. C'est donc la branche qui s'intéresse en premier lieu à comment parler des mathématiques.

Cela mène à une première distinction importante : qu'est-ce que n'est pas la logique ? Elle n'est pas, du moins dans le cadre donné dans ce livre, une recherche d'une vérité pré-existante aux mathématiques. Au contraire, la logique mathématique commence par l'acceptation des mathématiques, pour mieux les étudier. Cela peut apparaître comme un raisonnement circulaire : quelle valeur prend une étude d'un système se basant sur le système lui-même ? N'a-t-on pas un raisonnement erroné à partir du moment où nous utilisons les mathématiques pour parler des mathématiques ? La réponse que nous adopterons ici est la suivante : la logique mathématique, utilisant les mathématiques pour étudier le langage mathématique, est un procédé empirique, et les conclusions qu'elle tire ne sont à proprement parler que des résultats portant sur des objets mathématiques. Cependant, de la même manière qu'une mesure d'intensité électrique fait penser à un électricien que des électrons sont en mouvement alors que c'est la théorie électrique elle-même qui permet de supposer la pertinence de cette mesure, nos résultats mathématiques nous donnent à croire que quelque chose arrive, au-delà d'un simple fait mathématique.

Ainsi, quand nous aurons prouvé qu'il ne peut exister de preuve dans ZFC que ZFC est cohérente, où ZFC est la théorie des ensembles dans laquelle toutes les mathématiques usuelles peuvent se faire, nous en extrapolons largement qu'il n'existe pas de preuve de la cohérence de ZFC. Pourtant, les deux expressions ne signifient pas strictement la même

chose, mais il est cette conviction forte chez nombre de logiciens que cette étude des mathématiques par les mathématiques nous apprend des choses sur leur nature.

Beaucoup d'auteurs, pour distinguer ces mathématiques usuelles qui sont celles que nous utilisons lorsque l'on fait de la logique des théories formelles utilisées au sein de la logique pour représenter les mathématiques usuelles, utilisent l'expression « méta-théorie » pour la première. Ainsi la logique se place dans une méta-théorie pour étudier des théories. En utilisant ce terme, la thèse des paragraphes précédents est que l'étude des théories nous renseigne sur la méta-théorie.

Notons particulièrement la différence de traitement entre les deux : nous étudierons la théorie, tandis que la méta-théorie sera considérée comme acquise. Ainsi la théorie ZFC permet d'imaginer un langage formel pour parler des ensembles, mais elle se formule elle-même dans un univers que l'on considère comme pré-existant et vérifiant beaucoup de propriétés qu'on ne saurait écrire au sein de ce même univers. Face au trilemme d'Agrippa, le choix est donc fait de prendre une posture dogmatique initiale, en gardant l'esprit ouvert sur ce que l'étude logique peut nous faire réviser sur cet univers mathématique.

Cet ouvrage est principalement basé sur un semestre de cours suivi par les deux auteurs, qui couvrait les 4 thèmes principaux de la logique :

- la théorie des ensembles
- la théorie des modèles
- la calculabilité
- la théorie de la démonstration

C'est donc naturellement que ce cours sera structuré suivant ces quatres parties, mais en ajoutant une partie préliminaire présentant les outils de base qui seront utilisés en logique : l'induction, la logique propositionnelle et le calcul des prédicats ainsi que la théorie des ensembles ordonnés. Nous pensons en effet que ces prérequis méritent d'être traités à part, à la fois pour leur importance dans toutes les autres parties et pour pouvoir s'attarder plus longuement sur des éléments qui ne sont pas toujours approfondis dans un thème donné.

Nous remercions nos professeurs Arnaud Durand, Thomas Ibarlucia, Thierry Joly et Alessandro Vignati, du master LMFI, ainsi que Daniel Hirschkoff, Pascal Koiran, Natacha Portier et Colin Riba qui ont été nos professeurs à l'ÉNS de Lyon et dont les cours ont été de magnifiques portes d'entrées vers le monde de la logique.

Glossaire

Enculer les mouches (apparition : fin du XXe siècle, variation possible : sodomiser les drosophiles) Composé de enculer et de mouche. Permet probablement d'imager une grande difficulté ou déployer de grands efforts pour un but dérisoire. Décrit également une démarche consistant à s'attarder inutilement sur des points de détail, en faisant preuve d'une méticulosité extrême, voire excessive, au détriment de l'essentiel.. viii

Honnête et fréquentable Se dit d'une fonction manipulée par un physicien, c'est-à-dire continues et dérivables autant que les nécessités de calcul l'exigeront.. viii

Poussage de symboles Activité passionnante à laquelle s'attèlent certains logiciens, qui consiste à dérouler les définitions, les notations et les abréviations des symboles jusqu'à arriver au résultat, et ce, sans avoir besoin d'ajouter aucune conjonction de coordination de la langue française.. viii

Quanteur (terme vieilli, usité seulement en logique) Version plus rare de quantificateur. Dérivé savant du latin *quantus* (qui signifie « combien ») et *-eur*. Utilisé notamment dans l'expression (uniquement française) « élimination des quanteurs ». Exemple d'utilisation :

> Le quanteur universel I.6 est un signe logique : une paire de parenthèses avec une variable à l'intérieur ; la partie de formule qui le suit et qui généralement contient cette variable sera délimitée par des parenthèses ad hoc, qui serviront à faire reconnaître l'étendue (le scope) de ce quanteur.
> — Jean Largeault, Intuitionisme et théorie de la démonstration, 1992

. viii

Sommaire

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque

egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetuer.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

Première partie

Préliminaires

Chapitre 1

Induction

Table des sous-matières

L'un des outils fondamentaux en logique est l'induction. Intuitivement, elle peut se voir comme une généralisation du principe de récurrence. Nous allons cependant adopter un formalisme différent de celui utilisé pour faire une simple récurrence. Les objets principaux sur lesquels nous utiliserons l'induction sont les ensembles inductifs et les relations inductives, que nous présenterons. Ces deux objets sont associés à des formalismes différents : le premier aux grammaire en forme de Backus-Naur, et le deuxième aux points fixes et à la théorie des treillis. Comme la théorie des treillis sera étudiée plus tard dans cette partie préliminaire, nous ne traiterons qu'un cas restreint suffisant pour le travail sur l'induction.

L'objectif de ce chapitre est de donner une justification mathématique aux procédés qui seront utilisés par la suite, et d'offrir un modèle mathématique derrière le formalisme introduit, pour le lecteur qui en aurait besoin. Le point essentiel est avant tout de comprendre comment fonctionne une preuve par induction et comment utiliser les objets inductifs, puisqu'ils seront utilisés sans arrêt par la suite. Cependant, les justifications mathématiques données sont partielles : tous les cas ne se ramènent pas à ceux traités de façon évidente. Le lecteur le plus prudent devra trouver comment adapter le formalisme donné dans ce chapitre aux multiples variantes qui seront utilisées sans le dire par la suite.

1.1 Ensemble inductif

1.1.1 Construction d'un ensemble inductif

Au niveau intuitif, les ensembles finis semblent avoir une réalité plus robuste que les ensembles infinis. Il est en effet très facile de se convaincre à partir de règles simples qu'il existe un ensemble à 3 éléments, ou à n éléments pour n aussi grand que l'on veut (bien que se convaincre qu'il existe un ensemble à 300! éléments semble légèrement plus long). Cette robustesse découle du fait qu'on peut explicitement les construire, et cette possibilité

n'existe que pour un ensemble fini. Pourtant, l'ensemble $\mathbb{N}$ tend aussi à être plus facilement accepté qu'un ensmeble tel que $\mathbb{R}/\mathbb{Q}$. Un point essentiel qui rend le premier ensemble logiquement plus simple que le deuxième est qu'il est facile à engendrer : l'ensemble $\mathbb{N}$ est constitué de l'élément 0 et de l'opération S définie par $n \mapsto n+1$, et tout autre élément de $\mathbb{N}$ peut être construit à partir de ces deux éléments. Sa structure est donc fondamentalement simple, et peut être décrite en des termes finis.

C'est exactement cette idée de structure générée par des termes finis qui est formalisée par les ensembles inductifs. Un ensemble inductif va être un ensemble obtenu par une liste de générateurs, chaque générateur ayant une arité (un nombre d'objets qu'il prend en entrée). Dans cette définition d'ensemble inductif, l'exemple canonique est bien sûr $\mathbb{N}$ lui-même, qu'on peut définir par :

- un constructeur sans argument, 0
- un constructeur à un argument, S

Avant de donner la définition d'ensemble inductif, nous allons donner un formalisme pour parler des constructeurs.

Définition 1.1.1.1 (Signature). Une signature est un couple C, α tel que $\alpha : C \to \mathbb{N}$. On appelle C l'ensemble des constructeurs et, pour $c \in C$, $\alpha(c)$ est appelé l'arité de c.

Une signature sera généralement donnée sous forme dite de Backus-Naur. Cette présentation se décompose de la façon suivante :

$$a, b, \ldots ::= \text{cas 1} \mid \text{cas 2} \mid \ldots$$

où $a, b, \ldots$ représentent les éléments que les constructeurs définissent, et où chaque cas est la définition d'un nouveau constructeur (ou d'une famille de constructeurs). Par exemple pour le cas de $\mathbb{N}$, nous avons :

$$n ::= 0 \mid S\ n$$

Il est fréquent d'employer des variables qui seront quantifiées hors de la définition à proprement parler, comme

$$\ell ::= \text{nil} \mid \text{cons}(a, \ell)$$

où $a \in A$ et A est certain ensemble fixé au préalable. Cette définition doit se lire comme l'ensemble $(\{\text{nil}\} \cup A, \alpha)$ où α est défini par :

$$\begin{array}{rccc} \alpha & : & C & \longrightarrow & \mathbb{N} \\ & & \text{nil} & \longmapsto & 0 \\ & & a(\in A) & \longmapsto & 1 \end{array}$$

Voyons maintenant comment associer à une signature un ensemble généré par les constructeurs donnés dans la signature. L'ensemble généré doit être un ensemble X contenant, pour chaque $x_1, \ldots, x_n \in X$ et $c \in C$ d'arité n, l'objet $c(x_1, \ldots, x_n)$, et ne doit contenir que les objets de cette forme. Nous procédons alors par le bas : un premier ensemble est construit par l'ensemble $C_0 = \{c \in C \mid \alpha(c) = 0\}$, puis l'ensemble C_1 est construit par $C_1 = C_0 \cup \{c(x_1, \ldots, x_n) \mid c \in C, x_1, \ldots, x_n \in C_0, \alpha(c) = n\}$ et ainsi de suite. Comme c est simplement un élément dans notre cas, écrire $c(x_1, \ldots, x_n)$ n'a pas de sens, c'est pourquoi l'on va utiliser à la place $(c, x_1, \ldots, x_n)$.

Définition 1.1.1.2 (Ensemble inductif sur une signature). Soit (C, α) une signature, on définit la suite d'ensembles $(X_i)_{i \in \mathbb{N}}$ par :

- $X_0 = \varnothing$

- $X_{n+1} = \{(c, x_1, \ldots, x_p) \mid c \in C, (x_1, \ldots, x_p) \in (X_n)^p, \alpha(c) = p\}$

L'ensemble inductif engendré par (C, α) est alors l'ensemble

$$X = \bigcup_{n \in \mathbb{N}} X_n$$

La définition donnée n'est pas exactement celle décrite plus haut, mais la proposition suivante assure que l'union finale génère bien le même ensemble avec les deux méthodes.

Proposition 1.1.1.3. *Soit (C, α) une signature, X l'ensemble inductif engendré par cette signature et (X_n) la suite précédemment construite. Alors*

$$\forall n, m \in \mathbb{N}, n \leq m \implies X_n \subseteq X_m$$

Démonstration. On procède par récurrence sur n :

- comme $X_0 = \varnothing$, il est évident que $\varnothing \subseteq X_m$ pour tout $m \in \mathbb{N}$.
- supposons que $X_n \subseteq X_m$ pour tout $m \geq n$. Alors

  $$X_{n+1} = \{(c, x_1, \ldots, x_p) \mid c \in C, (x_1, \ldots, x_p) \in (X_n)^p, \alpha(c) = p\}$$

  mais par inclusion, comme $(x_1, \ldots, x_p) \in (X_n)^p$, on en déduit que $(x_1, \ldots, x_p)$ est aussi dans $(X_m)^p$, pour tout $m \geq n$. Ainsi $(c, x_1, \ldots, x_p) \in X_{m+1}$ pour tout $m \geq n$, donc $X_{n+1} \subseteq X_m$ pour tout $m \geq n + 1$.

□

Le lemme suivant est un outil de base pour étudier des ensembles inductifs.

Lemme 1.1.1.4 (Lecture unique). *Soit une signature (C, α) et l'ensemble X engendré par cette signature. Alors pour tout élément $x \in X$, il existe $c \in C$ et $x_1, \ldots, x_p \in X$ (possiblement une famille vide si $\alpha(c) = 0$) telle que $p = \alpha(c)$ et $x = (c, x_1, \ldots, x_p)$.*

Démonstration. Soit (X_n) la suite d'ensemble définie précédemment telle que X en est l'union. Si $x \in X$, alors on trouve $n \in \mathbb{N}$ tel que $x \in X_n$. Par disjonction de cas sur ce n, on prouve le résultat :

- si $n = 0$ alors l'hypothèse $x \in X_0$ signifie qu'on a $x \in \varnothing$: par absurdité de la prémisse, la conclusion est vraie.
- si $n = m + 1$, alors $x \in X_n$ signifie que $x \in \{(c, x_1, \ldots, x_p) \mid c \in C, (x_1, \ldots, x_p) \in (X_m)^p, \alpha(c) = p\}$ d'où le résultat par définition.

□

1.1.2 Récursion et induction

Maintenant qu'une construction a été donnée d'un ensemble inductif, il faut vérifier que le comportement que l'on a décrit est en accord avec le comportement réel de l'ensemble que l'on a construit. Nous avons dit que l'ensemble engendré par (C, α) doit contenir exactement les éléments de la forme $c(x_1, \ldots, x_n)$ où $x_i \in X$ pour tout $i \in \{1, \ldots, n\}$, mais une autre façon de penser le fait que l'ensemble ne contient que des applications de constructeurs est le fait qu'une fonction partant d'un ensemble inductif est exactement spécifiée par son comportement sur les constructeurs. De telle fonctions sont appelées récursives, car elles peuvent faire appel à elles-mêmes pour s'appliquer sur les arguments d'un constructeurs, comme nous le verrons en pratique. Nous verrons ensuite que ce principe de définition récursive peut se modifier pour donner le principe d'induction, un analogue à la preuve par récurrence pour un ensemble inductif quelconque.

Théorème 1.1.2.1 (Propriété universelle des ensembles inductifs). *Soit (C, α) une signature, et X l'ensemble associé à cette signature. Soit un ensemble Y quelconque. Soit une famille de fonctions $\{f_c\}_{c\in C}$ telles que pour tout $c \in C$, $f_c : Y^{\alpha(c)} \to Y$ (avec la convention que $Y^0 = \{*\}$ est un singleton quelconque). Alors il existe une unique fonction $f : X \to Y$ telle que*

$$\forall c \in C, \forall (x_1, \dots, x_p) \in X^{\alpha(c)}, f((c, x_1, \dots, x_p)) = f_c(f(x_1), \dots, f(x_p))$$

On peut représenter l'équation précédente par le diagramme suivant, où l'égalité signifie que le diagramme commute, c'est-à-dire que les deux chemins possibles pour aller d'un coin à l'autre du carré sont égaux.

$$\begin{array}{ccc} X^{\alpha(c)} & \xrightarrow{f^{\alpha(c)}} & Y^{\alpha(c)} \\ \downarrow{\scriptstyle c} & & \downarrow{\scriptstyle f_c} \\ X & \xrightarrow{\quad f \quad} & Y \end{array}$$

Démonstration. Soit (X_n) la suite d'ensemble construite précédemment pour définir X. On va prouver par récurrence sur n la proposition suivante :

$$\forall n \in \mathbb{N}, \exists! f : X_n \to Y, \forall c \in C, \forall (x_1, \dots, x_p) \in (X_n)^{\alpha(c)}, f((c, x_1, \dots, x_p)) = f_c(f_n(x_1), \dots, f_n(x_p))$$

- Si $n = 0$, il existe une unique fonction $f_0 : \varnothing \to Y$ et elle vérifie la propriété par vacuité.
- Soit $n \in \mathbb{N}$. Supposons qu'il existe une unique fonction $f_n : X_n \to Y$ telle que

  $$\forall c \in C, \forall (x_1, \dots, x_p) \in (X_n)^{\alpha(c)}, f((c, x_1, \dots, x_p)) = f_c(f(x_1), \dots, f(x_p))$$

  On définit alors

  $$\begin{array}{rccc} f_{n+1} & : & X_{n+1} & \longrightarrow & Y \\ & & (c, x_1, \dots, x_p) & \longmapsto & f_c(f_n(x_1), \dots, f_n(x_p)) \end{array}$$

  On remarque que cette fonction vérifie bien la propriété. De plus, si une autre fonction g vérifie la propriété, alors pour $x \in X_{n+1}$, on trouve $c \in C$ et $x_1, \dots, x_p \in X_n$ tels que $x = (c, x_1, \dots, x_p)$, on a alors

  $$\begin{aligned} f_{n+1}(x) &= f((c, x_1, \dots, x_p)) \\ &= f_c(f_n(x_1), \dots, f_n(x_n)) \\ &= g((c, x_1, \dots, x_p)) \\ &= g(x) \end{aligned}$$

  Donc pour tout $x \in X_{n+1}$, $f_{n+1} = g(x)$, ce qui signifie que $f_{n+1} = g$, d'où l'unicité de f_{n+1}.

Soit $x \in X$, par définition on trouve $n \in \mathbb{N}$ tel que $x \in X_n$, et on peut donc définir $f(x) = f_n(x)$. Pour montrer que la fonction est unique, il suffit de remarquer que toute fonction $X \to Y$ vérifiant les prémisses du théorème induit une fonction sur chaque X_n, et doit donc coïncider avec chaque f_n sur X_n. □

Remarque 1.1.2.2. Pour une signature (C, α) et un ensemble associé X, chaque $c \in C$ peut maintenant s'interpréter comme une fonction $c : X^{\alpha(c)} \to X$. Nous confondrons dorénavant le constructeur et la fonction associée, et écrirons donc sans distinction $(c, x_1, \dots, x_p)$ et $c(x_1, \dots, x_p)$ pour un constructeur $c \in C$.

Cet outil nous permet maintenant de définir des fonctions dont le domaine est un ensemble inductif en utilisant sa structure.

Exemple. Donnons un premier exemple de fonction récursive : la fonction $d : n \mapsto 2n$, définie de $\mathbb{N}$ dans $\mathbb{N}$. En effet, on peut la décrire par

- $d(0) = 0$
- $d(S\ n) = S\ S\ d(n)$

nous donnant alors une définition de d grâce au théorème précédent.

Exemple. Un exemple à la fois d'ensemble inductif et de fonction récursive est le suivant. Soit A un ensemble quelconque, on définit la signature des listes sur A par la grammaire suivante :

$$\ell ::= \text{nil} \mid \text{cons}(a, \ell)$$

où $a \in A$. L'ensemble $\text{List}(A)$ est alors l'ensemble inductif associé. On définit alors la fonction $|-|$ donnant la longueur d'une liste :

- $|\,\text{nil}\,| = 0$
- $|\,\text{cons}(a, \ell)| = 1 + |\ell|$

Exercice 1.1.2.3. Soit A un ensemble. Donner une signature définissant l'ensemble $\text{BinTree}(A)$ des arbres binaires étiquetés par A, constitué d'un objet arbre vide et d'un constructeur binaire node prenant en argument un élément a de A et deux arbres g et d et retournant un nouvel arbre binaire, d'étiquette a et dont les deux sous-arbres sont g et d.

Définir une fonction $h : \text{BinTree}(A) \to \mathbb{N}$ donnant la hauteur d'un arbre, c'est-à-dire la longueur du plus long chemin entre la racine de l'arbre (la première étiquette) et une feuile (un arbre vide qui est sous-arbre). On prend comme convention que $h(\text{nil}) = 0$.

Définir une fonction $|-| : \text{BinTree}(A) \to \mathbb{N}$ donnant le nombre d'étiquettes d'un arbre.

Exercice 1.1.2.4. En utilisant la structure inductive de $\mathbb{N}$, définir la fonction $+ : \mathbb{N}\times\mathbb{N} \to \mathbb{N}$. On pourra pour cela définir la fonction $n \mapsto (m \mapsto n + m)$ et faire une fonction récursive sur l'argument n.

Un procédé similaire est celui d'induction. Là où la récursion nous permet de définir une fonction depuis un ensemble inductif, l'induction va nous permettre de faire une preuve sur un ensemble inductif. Il s'agit donc, au lieu de donner une fonction $X \to Y$, de donner un prédicat $P \subseteq X$ et de montrer que $P = X$.

Théorème 1.1.2.5 (Principe d'induction). *Soit (C, α) une signature et X l'ensemble inductif associé. Soit $P \subseteq X$ un prédicat sur X. Si pour tous $c \in C$ et $x_1, \ldots, x_p \in X^{\alpha(c)}$, la propriété*

$$x_1 \in P \text{ et } x_2 \in P \text{ et } \ldots \text{ et } x_p \in P \implies c(x_1, \ldots, x_p) \in P$$

est vérifiée, alors $P = X$.

Démonstration. Pour prouver ce résultat, il suffit de montrer que pour tout $n \in \mathbb{N}$, $X_n \subseteq P$ pour (X_n) la suite d'ensembles construisant X. En effectuant une récurrences sur n :

- Si $n = 0$, alors $\varnothing \subseteq P$.
- Supposons que $X_n \subseteq P$ pour $n \in \mathbb{N}$. Soit $x \in X_{n+1}$. Par définition de X_{n+1}, on trouve $c \in C$ et $x_1, \ldots, x_p \in X_n$ tels que $x = c(x_1, \ldots, x_p)$. Par hypothèse de récurrence, on en déduit que pour tout $i \in \{1, \ldots, p\}$, $x_i \in P$. Il vient donc, avec l'hypothèse du théorème sur P, que $c(x_1, \ldots, x_p) \in P$.
  On en conclus que $P \subseteq X_{n+1}$.

Ainsi, par récurrence, pour tout $n \in \mathbb{N}$, $X_n \subseteq P$. Cela montre alors que

$$\bigcup_{n\in\mathbb{N}} X_n \subseteq P$$

ce qu'il fallait démontrer. □

On peut affiner ce résultat en distinguant deux sortes de constructeurs : d'un côté les constructeurs d'arité 0, qui sont des constantes, et de l'autre les constructeurs d'arité supérieure à 1. Le théorème précédent nous dit que pour prouver une proposition P sur un ensemble inductif, il suffit de prouver qu'il contient les constantes et qu'il est stable par chaque constructeur. On remarque que dans le cas de $\mathbb{N}$, engendré par 0 et S, ce principe nous dit qu'une partie P contenant 0 et telle que $\forall n \in \mathbb{N}, n \in P \implies n+1 \in P$ est exactement $\mathbb{N}$: c'est le principe de récurrence.

Ainsi, puisque nous prouvons le principe d'induction à partir du principe de récurrence, et puisque le principe de récurrence est un cas particulier du principe d'induction, les deux sont logiquement équivalents. L'induction, cependant, est conceptuellement plus intéressante puisqu'elle peut s'utiliser dans plus de cas.

Exercice 1.1.2.6. On considère $\mathbb{N}$ comme un ensemble inductif, et la fonction d définie dans un exemple précédent. Montrer que pour tout $n \in \mathbb{N}$, $d(n)$ est pair. On prendra comme définition de pair le prédicat $\text{pair}(n) \triangleq \exists m \in \mathbb{N}, n = 2 \times m$.

Exercice 1.1.2.7. Soit un ensemble A. On définit $\oplus : \text{List}(A) \times \text{List}(A) \to \text{List}(A)$ par induction sur le premier argument:

- Pour tout $\ell \in \text{List}(A)$, $\text{nil} \oplus \ell = \ell$.
- Pour tous $\ell, \ell' \in \text{List}(A), a \in A$, $\text{cons}(a, \ell) \oplus \ell' = \text{cons}(a, \ell \oplus \ell')$.

Montrer que

$$\forall \ell, \ell' \in \text{List}(A), |\ell \oplus \ell'| = |\ell| + |\ell'|$$

1.2 Relation inductive

1.2.1 Construction d'une relation inductive

Une façon de considérer les ensembles inductif est, étant donnée une signature, de prendre le plus petit ensemble stable par les constructeurs de cette signature. Le problème, pour faire cela, est que nous n'avons d'ensemble sur lequelle travailler a priori. C'est pourquoi la construction précédente définissait chaque ensemble intermédiaire pour en prendre l'union. Au contraire, pour définir les relations inductives, nous avons un ensemble ambiant et nous pouvons donc travailler sur la notion de plus petit ensemble stable. Pour pouvoir mieux appréhender les relations inductives, nous allons directement introduire les règles d'inférence.

Rappelons la définition d'une relation.

Définition 1.2.1.1 (Relation). Soit X un ensemble. On appelle relation sur X une partie $R \subseteq X^n$ où $n \in \mathbb{N}$ est appelé l'arité de la relation R.

Une règle d'inférence, elle, va relier des relations.

Définition 1.2.1.2 (Règle d'inférence). On appelle règle d'inférence une présentation de la forme suivante:

$$\frac{P_1 \qquad P_2 \qquad \cdots \qquad P_n}{P}\ \mathrm{r}$$

Où $P_1, \ldots, P_n$ sont appelées les prémisses de la règle, et P est appelée la conclusion de la règle. On dit que la règle est juste si, lorsque toutes les prémisses de la règle sont vérifiées, alors la conclusion est aussi vérifiée.

Une règle peut contenir plusieurs paramètres, par exemple

$$\frac{a = b \qquad P(a)}{P(b)}$$

auquel cas la règle est juste lorsqu'elle est juste pour chaque instance possible de ces paramètres. Dans l'exemple, on suppose que a et b sont quantifiés sur un certain ensemble A, et la règle est donc juste lorsque $\forall a, b \in A, ((a = b) \text{ et } P(a)) \implies P(b)$.

Étant donnée une règle r comme précédemment avec un paramètre X non spécifié, on notera $R \models \mathrm{r}$ si la règle r est juste en prenant R pour le paramètre X. De même, on notera $R, x_1, \ldots, x_n \models \mathrm{r}$ dans le cas où l'ont remplace plusieurs paramètres.

Ainsi, à une règle de la forme

$$\frac{P_1 \qquad P_2 \qquad \cdots \qquad P_n}{P}\ \mathrm{r}$$

avec un paramètre R fixé d'arité $n \in \mathbb{N}$ et sur un ensemble E, on peut associer une fonction

$$\begin{array}{rcccl} r & : & \mathcal{P}(E^n) & \longrightarrow & \mathcal{P}(E^n) \\ & & R & \longmapsto & R \cup \{(x_1, \ldots, x_n) \in E^n \mid R, x_1, \ldots, x_n \models \mathrm{r}\} \end{array}$$

où tous les paramètres n'apparaissant pas dans la conclusion sont quantifiés de façon existentielle.

Exemple. Prenons la règle

$$\frac{x = y}{R(x, y)}\ \mathrm{refl}$$

exprimant que la règle R est réflexive. La fonction refl est alors $R \mapsto R \cup \{(x, x) \mid x \in E\}$.

Dans ce genre de cas, on écrira de façon plus compacte la règle par

$$\frac{}{R(x, x)}$$

Exemple. La règle précédente n'utilisait pas R dans ses prémisses, un exemple l'utilisant est la règle exprimant la transivité :

$$\frac{R(x, y) \qquad R(y, z)}{R(x, z)}\ \mathrm{trans}$$

Dans ce cas, l'image de R par trans est l'ensemble $R \cup \{(x, z) \in E^2 \mid \exists z \in E, R(x, y) \text{ et } R(y, z)\}$.

Supposons maintenant que nous ayons une relation R et une règle r, R ne vérifiant pas forcément r. Nous cherchons alors à construire à partir de R une relation vérifiant r. Une remarque essentielle : si R vérifie r, alors R est un point fixe de la fonction r associée. On va ainsi chercher à créer un point fixe de la fonction associée. Pour cela, nous allons donner un théorème essentiel en logique : le théorème de Knaster-Tarski.

Théorème 1.2.1.3 (Knaster-Tarski (faible)). *Soit un ensemble E quelconque, et une fonction $f : \mathcal{P}(E) \to \mathcal{P}(E)$ croissante pour $\subseteq$, c'est-à-dire telle que*

$$\forall A, B \in \mathcal{P}(E), A \subseteq B \implies f(A) \subseteq f(B)$$

Alors il existe un plus petit point fixe de f pour $\subseteq$, ou de façon équivalente il existe un élément $A \in \mathcal{P}(E)$ tel que $f(A) = A$ et tel que pour tous $B \in \mathcal{P}(E)$ tel que $f(B) = B$, $A \subseteq B$.

Démonstration. Pour commencer, définissons l'ensemble des pré-points fixes de f, qui est

$$\text{prefix}(f) = \{A \in \mathcal{P}(E) \mid f(A) \subseteq A\}$$

Soit $\alpha = \bigcap \text{prefix}(f)$, montrons que α est dans $\text{prefix}(f)$:

Pour cela, il suffit de montrer que $f(\alpha) \subseteq \alpha$, mais comme α est une intersection, il suffit de montrer que pour tout $A \in \text{prefix}(f), f(\alpha) \subseteq A$. Pour montrer cela, on remarque par transitivité de $\subseteq$ qu'il suffit de montrer que $f(\alpha) \subseteq f(A)$ pour tout $A \in \text{prefix}(f)$, mais cela est direct en utilisant le fait que f est croissante et que pour tout $A \in \text{prefix}(f)$, $\alpha \subseteq A$. Ainsi $f(\alpha) \subseteq \alpha$.

De plus, comme $f(\alpha) \subseteq \alpha$, on en déduit par croissance de f que $f(f(\alpha)) \subseteq f(\alpha)$, c'est-à-dire que $f(\alpha)$ est lui aussi un élément de $\text{prefix}(f)$: comme α en est une borne inférieure, cela signifie que $\alpha \subseteq f(\alpha)$, d'où en utilisant l'inclusion précédente, $f(\alpha) = \alpha$.

De plus, si B est un point fixe de f, alors $f(B) \subseteq B$ donc par définition de α, $\alpha \subseteq B$. □

Exercice 1.2.1.4. Soit E un ensemble quelconque et $f : \mathcal{P}(E) \to \mathcal{P}(E)$ une fonction croissante pour l'inclusion, et $A \in \mathcal{P}(A)$. Montrer que l'on peut étendre le théorème pour trouver un plus petit point fixe α tel que $A \subseteq \alpha$.

Exercice 1.2.1.5. Soit E un ensemble quelconque, et $f : \mathcal{P}(E) \to \mathcal{P}(E)$ une fonction croissante pour l'inclusion. On définit $g_f : \mathcal{P}(E) \to \mathcal{P}(E)$ par $X \mapsto X \cup f(X)$. Montrer que g_f est croissante pour l'inclusion.

Des deux exercises précédents, on peut déduire le résultat suivant :

Corollaire 1.2.1.6 (Définition d'une relation inductive). *Soit une règle* r *sur un ensemble E et une relation R. Alors il existe une plus petite relation R_r contenant R et stable par r. On dit alors que cette relation R_r est la relation définie par r sur R. Si $R = \varnothing$, on dira seulement que la relation est définie par r.*

Cela nous permet alors de définir le principe d'induction sur les relations, qui est très proche de celui sur les ensembles inductifs.

Théorème 1.2.1.7 (Induction sur une relation). *Soit E un ensemble, R une relation n-aire sur E et r une règle d'inférence incluant comme paramètre R. Soit $P \subseteq \mathcal{P}(E^n)$ un prédicat d'arité n sur E. Supposons que $R \subseteq P$ et que $P \models$ r. Alors $R' \subseteq P$.*

Démonstration. Il suffit de remarquer que P est un point fixe de r puisque $P \models$ r. Ainsi, comme P est un point fixe de r contenant R, on en déduit que P contient le plus petit point fixe de r contenant R, qui est exactement R'. □

Ce résultat est essentiel pour prouver beaucoup de résultats : étant donnée une relation R construite par induction, pour montrer un résultat de la forme $\forall x_1, \ldots, x_n \in E, R(x_1, \ldots, x_n) \implies P(x_1, \ldots, x_n)$ pour un certain prédicat P, il suffit de montrer que ce prédicat est stable par la règle le définissant.

Exercice 1.2.1.8. Soient désormais n règles $\mathrm{r}_1, \ldots, \mathrm{r}_n$ incluant toutes un paramètre R d'arité p. Montrer qu'on peut leur associer une fonction $r_{1,\ldots,n} : \mathcal{P}(E^p) \to \mathcal{P}(E^p)$ croissante. En déduire un analogue des propositions précédentes pour un nombre fini de règles.

Remarque 1.2.1.9. La plupart des relations inductives que nous construirons utilisent plusieurs règles, mais l'idée de la construction pour une seule règle suffit. L'exercice précédent sert principalement pour justifier au lecteur le plus dubitatif que le procédé fonctionne effectivement aussi pour plusieurs règles.

Exercice 1.2.1.10. Soit une relation R et une règle r. Montrer que la relation R' définie par r sur R est la même que la relation R'' définie par la règle r et la règle

$$\frac{R(x_1, \ldots, x_n)}{R''(x_1, \ldots, x_n)}$$

Remarque 1.2.1.11. On définira dorénavant des relations inductives seulement par des règles.

Exercice 1.2.1.12. Soit A un ensemble quelconque. On définit sur $\mathrm{List}(A)$ le prédicat pair par induction avec les règles suivantes :

$$\frac{}{\mathrm{pair}(\mathrm{nil})} \qquad \frac{\mathrm{pair}(\ell)}{\mathrm{pair}(\mathrm{cons}(a, \mathrm{cons}(b, \ell)))}$$

Montrer l'assertion suivante :

$$\forall \ell \in \mathrm{List}(A), \mathrm{pair}(\ell) \implies \mathrm{pair}(|\ell|)$$

avec le prédicat pair sur les entiers défini précédemment.

1.2.2 Dérivation d'arbre

Une autre utilité de ce formalisme des règles d'inférences est de permettre de définir une dérivation, qui est une preuve purement syntaxique qu'une certaine relation est vérifiée.

Définition 1.2.2.1 (Dérivation). Soit une relation R définie par des règles $r_1, \ldots, r_n$. Une dérivation de $R(x_1, \ldots, x_p)$ est un arbre dont la racine est $R(x_1, \ldots, x_p)$ et tel que chaque nœud de l'arbre est une règle parmi $r_1, \ldots, r_n$ et contient autant de sous-arbres que l'arité de la règle présente.

Exemple. Soit $A = \{0\}$, dérivons $\mathrm{pair}([0,0,0,0])$ où $[0,0,0,0]$ est une abréviation pour $\mathrm{cons}(0, \mathrm{cons}(0, \mathrm{cons}(0, \mathrm{cons}(0, \mathrm{nil}))))$:

$$\cfrac{\cfrac{\cfrac{}{\mathrm{pair}(\mathrm{nil})}}{\mathrm{pair}([0,0])}}{\mathrm{pari}([0,0,0,0])}$$

Par construction, comme une relation définie par des règles vérifie ces règles, et étant donnée la définition de « vérifier une règle », il ne fait aucun doute que si l'on peut dériver $R(x_1, \ldots, x_n)$, alors $R(x_1, \ldots, x_n)$ est vraie.

En logique, nous formalisons les preuves mathématiques comme de tels arbres, car leur structure simple permet de facilité l'étude du langage mathématique. Si, dans la réalité, on n'écrit pas une preuve comme une dérivation (nous le verrons, cette pratique est beaucoup trop laborieuse), il est communément admis qu'une preuve satisfaisante permet de savoir écrire un tel arbre mentalement.

Étudions maintenant la notion de règle admissible et dérivable, correspondant respectivement à une règle vérifiée par une relation et à une règle dont on peut syntaxiquement prouver la correspondance avec la règle.

Définition 1.2.2.2 (Règle admissible). Une règle r est admissible pour une relation R si $R \models \mathrm{r}$. La règle est admissible pour un ensemble de règles $\mathrm{r}_1, \ldots, \mathrm{r}_n$ si la relation définie par ces règles vérifie la règle r.

Définition 1.2.2.3 (Règle dérivable). Soit R une relation définie par les règles $\mathrm{r}_1, \ldots, \mathrm{r}_n$, une règle r est dérivable s'il existe une dérivation de la conclusion de r utilisant les règles $\mathrm{r}_1, \ldots, \mathrm{r}_n$ et les prémisses de r.

Exercice 1.2.2.4. Montrer qu'une règle dérivable est admissible et que si une règle est admissible pour une relation R, alors la relation définie par cette règle sur R est exactement R.

Exercice 1.2.2.5. Avec les définitions précédentes, montrer que la règle

$$\frac{\mathrm{pair}(\ell)}{\mathrm{pair}([a,b,c,d] \oplus \ell)}$$

est dérivable. Montrer que la règle

$$\frac{\mathrm{pair}(\mathrm{cons}(a, \mathrm{cons}(b, \ell)))}{\mathrm{pair}(\ell)}$$

est admissible. Est-elle dérivable ?

Exercice 1.2.2.6. Étant donnée une relation R inductive sur un ensemble E définie par des règles $\mathrm{r}_1, \ldots, \mathrm{r}_n$, donner une construction analogue au principe de récursion pour les ensembles inductifs, permettant de définir une fonction $R \to Y$ pour un ensemble Y quelconque.

Enfin, donnons le lemme d'inversion, que l'on peut voir comme un affaiblissement de l'induction sur une relation, pour une seule étape.

Théorème 1.2.2.7 (Inversion). *Soient des règles $\mathrm{r}_1, \ldots, \mathrm{r}_n$, un ensemble E et R la relation définie sur R par ces règles. Si $R(x_1, \ldots, x_p)$ pour un tuple $(x_1, \ldots, x_p) \in E$, alors il existe $i \in \{1, \ldots, n\}$ et des instances $P_1, \ldots, P_k$ des prémisses de r_i tels que*

$$\frac{P_1 \qquad \cdots \qquad P_k}{R(x_1, \ldots, x_p)}\ \mathrm{r}_i$$

est vérifiée.

Démonstration. Pour prouver ce résultat, il suffit de le prouver par induction sur la relation R. Mais alors, pour chaque cas de la règle à vérifier, le résultat est vrai par hypothèse. □

Ce résultat permet de travailler par disjonction de cas lorsque l'on étudie une relation inductive. On peut l'assimiler au lemme de lecture unique. Par exemple, il permet de dire à partir de $\mathrm{pair}(\ell)$ que ℓ est soit la liste vide, soit $\mathrm{cons}(a, \mathrm{cons}(b, \ell))$ pour deux éléments $a, b \in A$, en prenant les conventions de l'exercice 1.2.2.5.

Chapitre 2

Logique propositionnelle

Table des sous-matières

POUR commencer l'étude de la logique, nous allons étudier sa forme la plus élémentaire : la logique propositionnelle. Lorsque l'on écrit une phrase mathématique, disons par exemple

$$\forall n \in \mathbb{N}, (\exists m \in \mathbb{N}, n = 2 \times m) \text{ ou } (\exists m \in \mathbb{N}, n = 2 \times m + 1)$$

on peut séparer plusieurs parties :

- les quantificateurs.
- les connecteurs logiques, comme « et », « ou » ou « non ».
- les propositions atomiques, comme $n = 2 \times m$ ci-dessus.

La logique propositionnelle est une simplification de cette grammaire, dans laquelle les propositions atomiques sont remplacées par de simple variables, pouvant prendre la valeur Vrai ou Faux, et où l'on supprime les quantificateurs. L'étude de la logique propositionnelle, beaucoup plus simple que le calcul des prédicats, permet de donner une première idée des propriétés qui nous intéressent dans la logique.

2.1 Définitions

Fixons tout d'abord un ensemble Var dénombrable de variables propositionnelles. Nous allons construire l'ensemble des propositions comme un ensemble inductif, en ajoutant les connecteurs logiques usuels.

Définition 2.1.0.1 (Propositions). On définit l'ensemble Prop des propositions du calcul propositionnel par la grammaire suivante :

$$P, Q ::= x \mid \top \mid \bot \mid \neg P \mid P \vee Q \mid P \wedge Q \mid P \to Q$$

où $x \in$ Var est une variable propositionnelle.

Le sens des différents symboles est le suivant :

- $\top$ représente la proposition vraie.
- $\bot$ représente la proposition fausse.
- $\neg$ représente la négation logique.
- $\vee$ représente la disjonction logique.
- $\wedge$ représente la conjonction logique.
- $\rightarrow$ représente l'implication logique.

Par convention, nous donnons l'ordre de priorité (du plus prioritaire au moins prioritaire) suivant : $\neg > \wedge > \vee > \rightarrow$. Cette convention évite ainsi d'écrire certaines parenthèses. De plus, comme $\wedge$ et $\vee$ seront montrés associatifs, nous ne parenthèserons pas $P \vee Q \vee R$ par exemple. Pour $\rightarrow$, nous associons à droite, c'est-à-dire que $P \rightarrow Q \rightarrow R$ signifie $P \rightarrow (Q \rightarrow R)$. Par exemple la proposition

$$(P \rightarrow Q) \wedge (R \rightarrow Q) \rightarrow P \vee R \rightarrow Q$$

doit se lire

$$((P \rightarrow Q) \wedge (R \rightarrow Q)) \rightarrow ((P \vee R) \rightarrow Q)$$

L'objectif d'une proposition est bien sûr de lui attribuer une valeur de vérité. Evidemment, la valeur de vérite d'une proposition dépend de celle des variables. Par exemple, $(x \vee \neg y) \wedge z$ n'aura pas la même valeur de vérité suivant si z est vrai ou faux.

Pour manipuler les variables d'une propositions, il est important de définir l'ensemble des variables impliquées dans la construction d'une proposition.

Définition 2.1.0.2 (Variables libres). Par induction sur la structure inductive de Prop, on définit pour P l'ensemble $\mathrm{VL}(P)$ des variables libres de P :

- si $P = x$ et $x \in \mathrm{Var}$, alors $\mathrm{VL}(P) = \{x\}$.
- si $P = \neg Q$ alors $\mathrm{VL}(P) = \mathrm{VL}(Q)$.
- si $P = Q \vee R$ ou $P = Q \wedge R$ ou $P = Q \rightarrow R$, alors $\mathrm{VL}(P) = \mathrm{VL}(Q) \cup \mathrm{VL}(R)$.

Remarque 2.1.0.3. La dénomination de variable « libre » prendra son sens dans le prochain chapitre. L'intérêt de définir VL et non juste l'ensemble des variables est d'ignorer les variables muettes, mais dans le calcul propositionnel aucune variable n'est muette.

Une attribution de valeurs de vérité aux variables propositionnelles est appelée un environnement. On utilisera l'ensemble $\{0, 1\}$ pour signifier $\{\mathrm{Faux}, \mathrm{Vrai}\}$.

Définition 2.1.0.4 (Environnement). Un environnement est une fonction partielle $\rho : \mathrm{Var} \rightarrow \{0, 1\}$ dont le domaine (l'ensemble des valeurs $x \in \mathrm{Var}$ sur lesquelles ρ est défini) est fini.

Un environnement permet ensuite de définir la notion de valuation, qui est la valeur de vérité d'une proposition dans un environnement donné. On considère

Définition 2.1.0.5 (Valuation). Soit ρ un environnement. On définit par induction sur Prop la fonction partielle $\mathrm{Val}_\rho : \mathrm{Prop} \rightarrow \{0, 1\}$:

- pour $x \in \mathrm{Var}$, si $x \in \mathrm{Dom}(\rho)$ alors $\mathrm{Val}_\rho(x) = \rho(x)$, si $x \notin \mathrm{Dom}(\rho)$ alors $\mathrm{Val}_\rho(x)$ n'est pas défini.
- soit P une proposition, alors $\mathrm{Val}_\rho(\neg P) = 1 - \mathrm{Val}_\rho(P)$.

- soient P et Q deux propositions, alors $\mathrm{Val}_\rho(P \vee Q) = \max(\mathrm{Val}_\rho(P), \mathrm{Val}_\rho(Q))$
- soient P et Q deux propositions, alors $\mathrm{Val}_\rho(P \wedge Q) = \min(\mathrm{Val}_\rho(P), \mathrm{Val}_\rho(Q))$
- soient P et Q deux propositions, alors $\mathrm{Val}_\rho(P \to Q) = \max(1 - \mathrm{Val}_\rho(P), \mathrm{Val}_\rho(Q))$

Exercice 2.1.0.6. Soit ρ un environnement et P une proposition, montrer que $\mathrm{Val}_\rho(P)$ est définie si et seulement si $\mathrm{VL}(P) \subseteq \mathrm{Dom}(\rho)$.

Exercice 2.1.0.7. Soit P une proposition. Soient ρ et ρ' deux environnements tels que $\rho_{|\ \mathrm{VL}(P)} = \rho'_{|\ \mathrm{VL}(P)}$. Montrer que $\mathrm{Val}_\rho(P) = \mathrm{Val}_{\rho'}(P)$.

On peut donc définir la relation de vérité, dans un environnement donné, pour une proposition.

Définition 2.1.0.8 (Satisfaction). Soit ρ un environnement et $P \in \mathrm{Prop}$. On définir la relation $\rho \models P$ par

$$(\rho \models P) \triangleq (\mathrm{Val}_\rho(P) = 1)$$

On dit alors que ρ satisfait P.

Si pour tout environnement ρ tel que $\mathrm{VL}(P) \subseteq \mathrm{Dom}(\rho)$, $\rho \models P$, alors on dit que P est une tautologie, et on le note $\models P$. Si aucun environnement ne satisfait P, on dit alors que P est une antilogie.

Exercice 2.1.0.9. Montrer que P est une antilogie si et seulement si $\neg P$ est une tautologie. En déduire que P est une tautologie si et seulement si $\neg P$ est une antilogie.

Donnons un premier outil pour traiter des vérités du calcul des propositions : les tables de vérité.

Définition 2.1.0.10 (Table de vérité). Une table de vérité pour une proposition P de variables libres $x_1, \ldots, x_n$ est un tableau contenant 2^n lignes et $n+1$ colonnes, où chaque ligne énumère un environnement différent, où la $i^{\text{ème}}$ colonne représente la valeur des environnement en x_i et où la dernière colonne représente la valuation de P pour l'environnement donné.

Exemple. Voici une table de vérité pour l'expression $\neg(x_0 \vee x_1) \to \neg x_0 \wedge \neg x_1$:

Une table de vérité permet de vérifier qu'une proposition est bien une tautologie puisque, grâce à l'exercice 2.1.0.7, nous savons que tout environnement ρ contenant les variable libres d'une proposition P donnée ne définit $\mathrm{Val}_\rho(P)$ que sur les variables libres listées dans la table de vérité. Ainsi, une proposition est une tautologie si et seulement si toute la dernière colonne est remplie de 1.

Exercice 2.1.0.11. Montrer que $\neg x_0 \wedge \neg x_1 \implies \neg(x_0 \vee x_1)$ est une tautologie.

La notion de tautologie peut être considérée comme la bonne notion de vérité dans le cadre du calcul propositionnel : une proposition vraie est une proposition qui s'évalue toujours en une formule vraie. Comme $\to$ sert à signifier l'implication logique, dire que $A \to B$ est une tautologie revient à dire que chaque fois que A est vraie, B l'est aussi. Ainsi, en notant $A \leftrightarrow B$ pour $(A \to B) \wedge (B \to A)$, dire que $A \leftrightarrow B$ est une tautologie revient à dire que A et B prennent toujours la même valeur de vérité. Cette relation est l'équivalence logique : elle traduit que deux propositions ont la même valeur.

Définition 2.1.0.12 (Équivalence logique). Pour tous $P, Q \in \mathrm{Prop}$, on dit que P et Q sont logiquement équivalents, ce que l'on écrit $P \equiv Q$, si $\models P \leftrightarrow Q$.

Un affaiblissement est la relation de conséquence logique.

Définition 2.1.0.13 (Conséquence logique). Pour tous $P, Q \in \text{Prop}$, on dit que Q est conséquence logique de P, ce que l'on note $P \vDash Q$, si $\models P \to Q$.

Exercice 2.1.0.14. Vérifier à l'aide de tables de vérités les équivalences suivantes (appelées lois de De Morgan) :

$$\neg\neg x \equiv x$$
$$\neg(x \wedge y) \equiv \neg x \vee \neg y$$
$$\neg(x \vee y) \equiv \neg x \wedge \neg y$$

2.2 Calcul des séquents propositionnels et complétude

Maintenant que nous avons défini une notion satisfaisante de vérité pour une proposition, il convient de se demander quels outils permettent de l'établir. Pour l'instant, pour prouver qu'une proposition est une tautologie, la seule façon de procéder est d'en construire la table de vérité. C'est une façon largement inefficace, puisqu'elle prend une taille exponentielle en le nombre de variables libres d'une proposition, et la preuve qu'une proposition est une tautologie est assez vide de sens : dans l'exercice précédent, il n'a été question que de calcul et pas de considérations logiques.

2.2.1 Théorème de compacité

Cette sous-section se concentre sur un théorème important de la logique : le théorème de compacité. Son principe est de permettre de passer d'ensembles finis à des ensembles infinis. Si son utilité est relativement anecdotique dans le cas de la logique propositionnelle, il sera un élément essentiel en logique du premier ordre. Pour pouvoir établir ce théorème, nous devons tout d'abord généraliser la relation de satisfaction $\vDash$ au cas d'un ensemble infini de propositions.

Définition 2.2.1.1 (Environnement infini). On définit l'ensemble $\mathcal{E}$ des environnement potentiellement infinis comme l'ensemble des fonctions partielles $\text{Var} \to \{0, 1\}$. On définit la fonction $\text{Val} : (\mathcal{E}, \text{Prop}) \to \{0, 1\}$, pour tout $\rho \in \mathcal{E}$, par :

$$\text{Val}_\rho(P) = \text{Val}_{\rho_{|\,\text{VL}(P)}}(P)$$

où le deuxième Val correspond à la définition sur les environnements finis.

On définit $\rho \models P$ de façon analogue à précédemment :

$$\rho \models P \iff \text{Val}_\rho(P) = 1$$

Exercice 2.2.1.2. Montrer que si ρ est un environnement fini, alors Val_ρ est définie de la même manière avec les deux définitions.

Définition 2.2.1.3 (Satisfaction infinie). Soit $\mathcal{P}$ un ensemble (potentiellement infini) de propositions et $P \in \text{Prop}$. On définit $\mathcal{P} \vDash P$ par

$$\forall \rho \in \mathcal{E}, (\forall Q \in \mathcal{P}, \rho \models Q) \implies \rho \models P$$

Le théorème de compacité possède plusieurs expressions différentes. Pour les donner, nous allons introduire le vocabulaire nécessaire sur les ensembles (potentiellement infinis) de propositions.

Définition 2.2.1.4 (Satisfiabilité). On dit qu'un ensemble $\mathcal{P} \subseteq \text{Prop}$ est satisfiable s'il existe $\rho \in \mathcal{E}$ telle que $\forall P \in \mathcal{P}, \rho \models P$. On dit qu'un ensemble $\mathcal{P} \subseteq \text{Prop}$ est finiment satisfiable si toutes ses parties finies sont satisfiables.

Définition 2.2.1.5 (Contradiction). On dit qu'un ensemble $\mathcal{P} \subseteq \text{Prop}$ est contradictoire s'il n'existe pas d'environnement $\rho \in \mathcal{E}$ tel que $\forall P \in \mathcal{P}, \rho \models P$. Un ensemble $\mathcal{P} \subseteq \text{Prop}$ est finiment contradictoire si l'une de ses parties finies est contradictoire.

Le théorème de compacité énonce alors l'équivalence entre la version finie et la version infinie des deux caractères, et de façon équivalence, l'équivalence de la relation $\vDash$ pour un ensemble infini et pour ses parties finies. Montrons d'abord que ces trois principes sont bien équivalents.

Proposition 2.2.1.6. *Les deux propriétés suivantes sont équivalentes :*

(i) pour tout $\mathcal{P} \subseteq \text{Prop}$, $\mathcal{P}$ est satisfiable si et seulement s'il est finiment satisfiable.

(ii) pour tout $\mathcal{P} \subseteq \text{Prop}$, $\mathcal{P}$ est contradictoire si et seulement s'il est finiment contradictoire.

(iii) pour tout $\mathcal{P} \subseteq \text{Prop}$ et $P \in \text{Prop}$, $\mathcal{P} \vDash P$ si et seulement s'il existe $F \subseteq_{\text{fin}} \mathcal{P}$ tel que $F \vDash P$.

Démonstration. Pour commencer, remarquons qu'un ensemble satisfiable est finiment satisfiable, qu'un ensemble finiment contradictoire est contradictoire, et que s'il existe $F \subseteq_{\text{fin}} \mathcal{P}$ tel que $F \vDash P$, alors pour tout environnement ρ tel que $\rho \models \mathcal{P}$, on a en particulier $\rho \models F$ et donc $\rho \models P$. Il nous suffit donc de travailler sur un seul sens de l'équivalence à chaque fois. Nous allons maintenant montrer $(i) \implies (ii) \implies (iii) \implies (i)$:

- Supposons (i) et montrons (ii). Soit $\mathcal{P} \subset \text{Prop}$ contradictoire, montrons que $\mathcal{P}$ est finiment contradictoire. Par l'absurde, supposons que $\mathcal{P}$ n'est pas finiment contradictoire : toute partie finie $F \subseteq \mathcal{P}$ possède donc une valuation ρ_F telle que $\rho_F \models F$. Mais alors, en utilisant (i), on en déduit qu'il existe une valuation $\rho \models \mathcal{P}$. Pourtant $\mathcal{P}$ est contradictoire : c'est une absurdité. Ainsi, par l'absurde, on en déduit que $\mathcal{P}$ est finiment contradictoire.
- Supposons (ii) et montrons (iii). Soient $\mathcal{P} \subseteq \text{Prop}$, $P \in \text{Prop}$ tels que $\mathcal{P} \vDash P$. Montrons qu'il existe $F \subseteq_{\text{fin}} \mathcal{P}$, tel que $F \vDash P$. Comme $\mathcal{P} \vDash P$, on en déduit que $\mathcal{P} \cup \{\neg P\}$ est contradictoire. Ainsi, par (ii), on trouve un ensemble $F \subseteq_{\text{fin}} \mathcal{P} \cup \{\neg P\}$ contradictoire. Si F ne contient pas $\neg P$, alors $F \vDash P$ est vérifié par vacuité de la condition. Si $F = F' \sqcup \{\neg P\}$, alors pour tout ρ tel que $\rho \models F'$, $\rho \not\models \neg P$ donc $\rho \models P$, d'où $F' \vDash P$ et $F' \subseteq_{\text{fin}} \mathcal{P}$.
- Supposons (iii) et montrons (i). Soit $\mathcal{P} \subseteq \text{Prop}$ un ensemble finiment satisfiable, montrons que $\mathcal{P}$ est satisfiable. Par l'absurde, supposons que $\mathcal{P}$ n'est pas satisfiable. Cela signifie que $\mathcal{P} \vDash \bot$: par (iii) on trouve donc $F \subseteq_{\text{fin}} \mathcal{P}$ tel que $F \vDash \bot$: F est donc contradictoire, ce qui contredit le fait que $\mathcal{P}$ est finiment satisfiable. Par l'absurde, on en déduit que $\mathcal{P}$ est satisfiable.

□

Il ne nous reste plus qu'à prouver le résultat en lui-même.

Théorème 2.2.1.7 (Compacité de la logique propositionnelle). *Pour tout $\mathcal{P} \subseteq \text{Prop}$, si $\mathcal{P}$ est finiment satisfiable alors $\mathcal{P}$ est satisfiable.*

Démonstration. Soit $\mathcal{P} \subseteq \mathrm{Prop}$ finiment satisfiable, prouvons que $\mathcal{P}$ est satisfiable. Pour construire une valuation satisfaisant $\mathcal{P}$ entièrement, nous allons construire une suite de valuations partielles, $(\rho_n)_{n\in\mathbb{N}}$, telle que $(\rho_n)_{|\{0,\dots,m\}} = \rho_m$ pour tous $m \leq n$. Tout d'abord, comme $\mathcal{P}$ est dénombrable, on peut se donner une énumération $\mathcal{P} = \{P_i\}_{i\in\mathbb{N}}$.

- ρ_0 est la fonction partielle nulle part définie.
- Soit $n \in \mathbb{N}$. Supposons donnée ρ_n et construisons alors ρ_{n+1}. A FINIR

A FINIR □

2.2.2 Calcul des séquents

Ce qui manque à notre système, c'est une syntaxe : un système simple qui va nous permettre de justifier des tautologies. Cette syntaxe se base sur la notion de séquent : un séquent est une paire de listes de propositions, que l'on écrira $\Gamma \vdash \Delta$, exprimant que $\bigwedge \Gamma \to \bigvee \Delta$ est une tautologie. L'intérêt de la relation $\vdash$ à définir est de donner un système facile de preuve : pour prouver que P est une tautologie, il suffit de prouver $\mathrm{nil} \vdash \mathrm{cons}(P, \mathrm{nil})$ en suivant les règles de base définissant $\vdash$.

Notation 2.2.2.1. Dorénavant, pour utiliser des listes, nous écrirons a, ℓ pour signifie $\mathrm{cons}(a, \ell)$ et nous n'écrirons pas nil lorsque le contexte permet clairement de comprendre qu'il s'agit d'une liste. Par exemple, on confondra la simple proposition P et la liste $\mathrm{cons}(\mathrm{nil}, P)$. La notation Γ, Δ correspond à la concaténation, qui serait notée avec les conventions précédentes $\Gamma \oplus \Delta$.

Définition 2.2.2.2 (Calcul des séquents). On définit la relation $\vdash \subseteq \mathrm{List}(\mathrm{Prop})^2$ par les règles suivantes:

$$\frac{P \in \Gamma}{\Gamma \vdash P}\,\mathrm{ax} \qquad \frac{\Gamma, P \vdash \Delta \qquad \Theta \vdash \Xi, P}{\Gamma, \Theta \vdash \Delta, \Xi}\,\mathrm{cut}$$

$$\frac{\Gamma, P, Q, \Gamma' \vdash \Delta}{\Gamma, Q, P, \Gamma' \vdash \Delta}\,\mathrm{le} \qquad \frac{\Gamma \vdash \Delta, P, Q, \Delta'}{\Gamma \vdash \Delta, Q, P, \Delta'}\,\mathrm{re}$$

$$\frac{\Gamma, P, P \vdash \Delta}{\Gamma, P \vdash \Delta}\,\mathrm{lc} \qquad \frac{\Gamma \vdash \Delta, P, P}{\Gamma \vdash \Delta, P}\,\mathrm{rc}$$

$$\frac{\Gamma \vdash \Delta}{\Gamma, P \vdash \Delta}\,\mathrm{lw} \qquad \frac{\Gamma \vdash \Delta}{\Gamma \vdash \Delta, P}\,\mathrm{rw}$$

$$\frac{\Gamma, \top \vdash \Delta}{\Gamma \vdash \Delta}\,\top \qquad \frac{\Gamma \vdash \Delta, \bot}{\Gamma \vdash \Delta}\,\bot$$

$$\frac{\Gamma \vdash \Delta, P}{\Gamma, \neg P \vdash \Delta}\,\mathrm{l}\neg \qquad \frac{\Gamma, P \vdash \Delta}{\Gamma \vdash \Delta, \neg P}\,\mathrm{r}\neg$$

$$\frac{\Gamma, P \vdash \Delta \qquad \Theta, Q \vdash \Xi}{\Gamma, \Theta, P \vee Q \vdash \Delta, \Xi}\,\mathrm{l}\vee \qquad \frac{\Gamma \vdash \Delta, P}{\Gamma \vdash \Delta, P \vee Q}\,\mathrm{r}\vee_1 \qquad \frac{\Gamma \vdash \Delta, Q}{\Gamma \vdash \Delta, P \vee Q}\,\mathrm{r}\vee_2$$

$$\frac{\Gamma, P \vdash \Delta}{\Gamma, P \wedge Q \vdash \Delta}\,\mathrm{l}\wedge_1 \qquad \frac{\Gamma, Q \vdash \Delta}{\Gamma, P \wedge Q \vdash \Delta}\,\mathrm{l}\wedge_2 \qquad \frac{\Gamma \vdash \Delta, P \qquad \Theta \vdash \Xi, Q}{\Gamma, \Theta \vdash \Delta, \Xi, P \wedge Q}\,\mathrm{r}\wedge$$

$$\frac{\Gamma, Q \vdash \Delta \qquad \Theta \vdash \Xi, P}{\Gamma, \Theta, P \to Q \vdash \Delta, \Xi}\,\mathrm{l}\to \qquad \frac{\Gamma, P \vdash \Delta}{\Gamma \vdash \Delta, P \to Q}\,\mathrm{r}\to_1 \qquad \frac{\Gamma \vdash \Delta, Q}{\Gamma \vdash \Delta, P \to Q}\,\mathrm{r}\to_2$$

On dit qu'une proposition P est prouvable si $\vdash P$ est dérivable.

L'objectif de la suite de cette section est de prouver que $\vdash$ définit en fait exactement les tautologies, au sens suivant : pour toutes propositions P et Q, $P \vDash Q$ si et seulement si $P \vdash Q$ (ainsi, en prenant par exemple $P = \top$, on peut montrer que $\vDash P$ si et seulement si $\vdash P$).

2.2.3 Correction du calcul des séquents

La première étape, qui est la plus simple, est de montrer que le calcul des séquents est correct, c'est-à-dire que si l'on arrive à dériver $\Gamma \vdash \Delta$, alors $\bigwedge \Gamma \to \bigvee \Delta$ est une tautologie. La preuve est simplement une induction sur la relation $\vdash$. Comme c'est la première longue preuve par induction, nous allons la rédiger complètement, mais il est d'usage pour prouver une induction dont les cas se ressemblent de ne prouver que quelques cas les plus significatifs.

Pour commencer, introduisons un lemme permettant de plus facilement travailler sur le résultat à prouver.

Lemme 2.2.3.1. *Soient $\Gamma, \Delta \in \mathrm{List}(\mathrm{Prop})$ et ρ un environnement, on a l'équivalence suivante :*

$$\mathrm{Val}_\rho(\bigwedge \Gamma \to \bigvee \Delta) = 1 \iff \min_{P \in \Gamma}(\mathrm{Val}_\rho(P)) \leq \max_{P \in \Delta}(P)$$

De même, les deux conditions équivalentes sont aussi équivalentes au fait qu'il existe $P \in \Gamma$ et $Q \in \Delta$ telles que $\mathrm{Val}_\rho(P) \leq \mathrm{Val}_\rho(Q)$.

Démonstration. Calculons d'abord $\mathrm{Val}_\rho(\bigwedge \Gamma \to \bigvee \Delta)$:

$$\begin{aligned}
\mathrm{Val}_\rho(\bigwedge \Gamma \to \bigvee \Delta) &= \max(1 - \mathrm{Val}_\rho(\bigwedge \Gamma), \mathrm{Val}_\rho(\bigvee \Delta)) \\
&= \max(1 - \min_{P \in \Gamma}(\mathrm{Val}_\rho(P)), \max_{P \in \Delta}(\mathrm{Val}_\rho(P))) \\
&= \max(\max_{P \in \Gamma}(1 - \mathrm{Val}_\rho(P)), \max_{P \in \Delta}(\mathrm{Val}_\rho(P)))
\end{aligned}$$

Ainsi, si $\mathrm{Val}_\rho(\bigwedge \Gamma \to \bigvee \Delta) = 1$, on peut trouver $P \in \Gamma$ tel que $\mathrm{Val}_\rho(P) = 0$ ou $P \in \Delta$ tel que $\mathrm{Val}_\rho(P)$. Dans les deux cas, on en déduit que $\min_{P \in \Gamma}(\mathrm{Val}_\rho(P)) \leq \max_{P \in \Delta}(\mathrm{Val}_\rho(P))$.

Inversement, si $\min_{P \in \Gamma}(\mathrm{Val}_\rho(P)) \leq \max_{P \in \Delta}(\mathrm{Val}_\rho(P))$ alors on peut trouver $P \in \Gamma$ et $Q \in \Delta$ tels que $\mathrm{Val}_\rho(P) \leq \mathrm{Val}_\rho(Q)$. Si $\mathrm{Val}_\rho(P) = 0$ alors par le calcul précédent, on en déduit que $\mathrm{Val}_\rho(\bigwedge \Gamma \to \bigvee \Delta) = 1$. Sinon, alors $\mathrm{Val}_\rho(Q) = 1$ par inégalité, et par le calcul précédent, $\mathrm{Val}_\rho(\bigwedge \Gamma \to \bigvee \Delta) = 1$. □

Exercice 2.2.3.2. Montrer l'équivalence au troisième énoncé.

Théorème 2.2.3.3 (Correction du calcul des séquents). *Soient $\Gamma, \Delta \in \mathrm{List}(\mathrm{Prop})$. Si $\Gamma \vdash \Delta$, alors $\vDash \bigwedge \Gamma \to \bigvee \Delta$.*

Démonstration. Soient $\Gamma, \Delta \in \mathrm{List}(\mathrm{Prop})$ et ρ un environnement, le résultat à montrer est que $\mathrm{Val}_\rho(\bigwedge \Gamma \to \bigvee \Delta) = 1$ sachant que $\Gamma \vdash \Delta$. Pour cela, on procède par induction sur $\vdash$, en utilisant le lemme précédent pour remplacer l'égalité de la valuation par l'existence de témoins d'une inégalité :

- Supposons que $P \in \Gamma$. Alors en choisissant P, on trouve $P \in \Gamma$ et $Q \in \{P\}$ $(= P)$ tels que $\mathrm{Val}_\rho(P) \leq \mathrm{Val}_\rho(Q)$.

- Supposons qu'il existe $Q \in \Gamma$, $R \in \Delta$, $Q' \in \Theta$ et $R' \in \Xi$ tels que

$$\mathrm{Val}_\rho(Q) \leq \mathrm{Val}_\rho(R) \tag{2.2.3.1}$$

$$\mathrm{Val}_\rho(Q') \leq \mathrm{Val}_\rho(R') \tag{2.2.3.2}$$

  Fixons ces Q, R, Q', R'. Montrons qu'alors il existe $Q'' \in \Gamma, \Theta$ et $R'' \in \Delta, \Xi$ tels que $\mathrm{Val}_\rho(Q'') \leq \mathrm{Val}_\rho(R'')$. Pour le prouver, traitons les cas possibles pour la valeur de $\mathrm{Val}_\rho(P)$:
  - si $\mathrm{Val}_\rho(P) = 0$, alors en utilisant l'équation (2.2.3.2), deux cas sont possibles. Dans le cas où $R' = P$, cela signifie que $\mathrm{Val}_\rho(Q') = 0$ et ainsi, en prenant $Q'' = Q'$ et R'' quelconque, l'inégalité est vérifiée. Dans le cas où $R' \neq P$, il suffit de prendre $Q'' = Q'$ et $R'' = R'$.
  - si $\mathrm{Val}_\rho(P) = 1$, alors en utilisant l'équation (2.2.3.1), deux cas sont possibles. Dans le cas où $Q = P$, cela signifie que $\mathrm{Val}_\rho(R) = 1$ et ainsi, en prenant $R'' = R$ et Q'' quelconque, l'inégalité est vérifiée. Dans le cas où $R' \neq P$, il suffit de prendre $Q'' = Q$ et $R'' = R$.
- Supposons qu'il existe $R \in \Gamma, P, Q, \Gamma'$ et $S \in \Delta$ tels que $\mathrm{Val}_\rho(R) \leq \mathrm{Val}_\rho(S)$. Dans ce cas, $R \in \Gamma, Q, P\Gamma'$ et $S \in \Delta$, et R, S respectent l'inégalité voulue.
- On procède comme dans le cas précédent.
- Supposons qu'il existe $Q \in \Gamma, P, P$ et $R \in \Delta$ tels que $\mathrm{Val}_\rho(Q) \leq \mathrm{Val}_\rho(R)$. Dans ce cas, $Q \in \Gamma, P$ et $R \in \Delta$, donnant l'inégalité voulue.
- On procède comme dans le cas précédent.
- Comme précédemment, si l'on trouve $Q \in \Gamma$ alors $Q \in \Gamma, P$, donc on peut réutiliser directement Q pour prouver la propriété sur le séquent du bas.
- On procède comme dans le cas précédent.
- Supposons qu'il existe $Q \in \Gamma$ et $R \in \Delta, P$ tels que $\mathrm{Val}_\rho(Q) \leq \mathrm{Val}_\rho(R)$. Si $R \neq P$, il suffit de reprendre Q et Q déjà définis. Suppsons maintenant que $R = P$. Distinguons les cas des valeurs possibles de $\mathrm{Val}_\rho(P)$:
  - si $\mathrm{Val}_\rho(P) = 0$, alors $\mathrm{Val}_\rho(Q) = 0$, donc on peut prendre n'importe quel autre formule dans Δ pour vérifier l'inégalité.
  - si $\mathrm{Val}_\rho(P) = 1$, alors on trouve $Q' = \neg P \in \Gamma, \neg P$ et $R \in \Delta$ quelconque, on a alors $\mathrm{Val}_\rho(\neg P) = 0 \leq \mathrm{Val}_\rho(R)$.
- On procède comme dans le cas précédent.
- Supposons qu'il existe $R \in \Gamma, P$, $S \in \Delta$, $R' \in \Theta, P$ et $S' \in \Xi$ tels que $\mathrm{Val}_\rho(R) \leq \mathrm{Val}_\rho(S)$ et $\mathrm{Val}_\rho(R') \leq \mathrm{Val}_\rho(S')$. Si R (respectivement R') est choisie comme étant différente de P (respectivement Q), le résultat est direct en reprenant (R, S) (respectivement R', S'). Supposons donc maintenant que $R = P$ et $R' = Q$. Dans ce cas, $\mathrm{Val}_\rho(P \vee Q) = \max(\mathrm{Val}_\rho(P), \mathrm{Val}_\rho(Q)) \leq \max(\mathrm{Val}_\rho(S), \mathrm{Val}_\rho(S'))$ d'où l'inégalité en choisissant S ou S' suivant lequel a la plus grande valuation.
- Les deux cas de $\mathrm{r}\vee_i$ pour $i \in \{1, 2\}$ se traitent de façon analogue, on ne traitera donc que le cas $i = 1$. Supposons donc qu'il existe $Q \in \Gamma$ et $R \in \Delta, P$ tels que $\mathrm{Val}_\rho(Q) \leq \mathrm{Val}_\rho(R)$. Dans le cas où $R \neq P$, le résultat est direct. Si $R = P$, alors comme $\mathrm{Val}_\rho(P \vee Q) \geq \mathrm{Val}_\rho(P)$ et par transitivité de $\leq$, on en déduit l'inégalité pour $Q' = Q$ et $R' = P \vee Q$.

□

On en déduit donc, comme $\bigwedge P = P$ et $\bigvee Q = Q$, que $P \vdash Q \implies P \vDash Q$.

Remarque 2.2.3.4. En utilisant à la fois le théorème de compacité et le théorème de correction, on en déduit que pour $\mathcal{P} \subseteq \mathrm{Prop}$ et $P \in \mathrm{Prop}$, si l'on trouve $\Gamma \in \mathrm{List}(\mathcal{P})$ tel que $\Gamma \vdash P$, alors $\mathcal{P} \vDash P$.

On peut donc définir la relation $\vdash \subseteq \mathcal{P}(\mathrm{Prop}) \times \mathrm{Prop}$ par

$$\mathcal{P} \vdash P \triangleq \exists \Gamma \in \mathrm{List}(\mathcal{P}), \Gamma \vdash P$$

et les résultats précédents prouvent que $\vdash \subseteq \vDash$. Nous allons maintenant montrer le sens réciproque : $\vDash \subseteq \vdash$. Ce résultat s'appelle la complétude.

Théorème 2.2.3.5 (Complétude du calcul des séquents). *Pour tous $\mathcal{P} \subseteq \mathrm{Prop}$ et $P \in \mathrm{Prop}$, si $\mathcal{P} \vDash P$ alors il existe $\Gamma \in \mathrm{List}(\mathcal{P})$ tel que $\Gamma \vdash P$.*

Démonstration. Tout d'abord, par théorème de compacité, il nous suffit de traiter le cas fini : pour $\Gamma \in \mathrm{List}(\mathrm{Prop})$, en prenant F_Γ l'ensemble de formules associé, si $F_\Gamma \vDash P$ alors $F \vdash P$.

On définit la relation $\vdash^*$ de façon analogue à $\vdash$ mais en ajoutant la règle suivante :

$$\frac{}{x_1, \ldots, x_n \vdash^* y_1, \ldots, y_p}\ \mathrm{Ax}^\dagger$$

où $x_1, \ldots, x_n, y_1, \ldots, y_n$ sont des variables propositionnelles toutes deux à deux distinctes.

On prouve maintenant que pour tous $\Gamma, \Delta \in \mathrm{List}(\mathrm{Prop})$, $\Gamma \vdash^* \Delta$. A FAIRE

Maintenant, prouvons que si $\Gamma \vdash^* \Delta$ et $\Gamma \nvDash \Delta$, alors $\Gamma \nvDash \Delta$. On procède par induction sur $\vdash^*$. A FAIRE

Ainsi, comme $\Gamma \vdash^* \Delta$ est toujours vérifié, on en déduit que $\Gamma \nvDash \Delta \implies \Gamma \nvDash \Delta$, d'où par contraposée $\Gamma \vDash \Delta \implies \Gamma \vdash \Delta$. □

Ainsi, nous avons un système syntaxique $\vdash$ permettant d'entièrement caractériser la relation $\vDash$, qui est une relation sémantique (elle relie les propositions par leur sens, là où $\vdash$ n'est qu'un système formel, encodable par exemple sur un ordinateur).

Chapitre 3

Calcul des prédicats

Table des sous-matières

Le premier but de la logique mathématique est de rendre compte du langage mathématique. A ce titre, la logique propositionnelle est clairement insuffisante, et nous l'avons déjà présentée comme une simplification du langage mathématique habituel. Dans ce chapitre, nous allons nous intéresser à la formalisation de ce langage mathématique habituel, qui est le calcul des prédicats de la logique du premier ordre. Le terme prédicat désigne le fait que nos propositions dépendent de termes, qui représentent des objets mathématiques. L'expression « logique du premier ordre » désigne la capacité d'expression de nos propositions : celles-ci ne peuvent parler que des objets mathématiques désignés préalablement par l'univers de discours. Par contraste, la logique du deuxième ordre permet de parler, en plus de ces objets, des propositions elles-mêmes : on peut écrire par exemple $\forall P \in \mathrm{Prop}, P \to P$.

Nous nous attarderons d'abord sur la définition, à partir d'une signature du premier ordre, des termes et des formules, ainsi que les notions syntaxiques de variables libres, liées et de substitution. Ensuite, nous introduirons les notions les plus élémentaires de la théorie des modèles, et la relation de satisfaction $\models$. Enfin, comme dans le chapitre précédent, nous allons définir une syntaxe pour le calcul des prédicats. La différence, cependant, est que nous prouverons le théorème de compacité à partir de la complétude.

Ce chapitre peut être vu comme la base commune de la théorie de la démonstration et de la théorie des modèles. A ce titre, nous donnerons avant tout les définitions des concepts

importants, mais n'allons pas nous attarder sur ceux-ci, puisque nous les reverrons dans des chapitres dédiés. En particulier nous allons donner un formalisme pour la syntaxe du calcul des prédicats et un seul, alors que la partie dédiée à la théorie de la démonstration donnera un résultat plus général sur toute une famille de systèmes de preuves, et présentera plusieurs formalismes.

On se fixe pour tout ce chapitre un ensemble Var dénombrable de variables.

3.1 Signatures, termes et formules

3.1.1 Définition d'une signature

Reprenons l'exemple que nous avions donné au début du chapitre 2 :

$$\forall n \in \mathbb{N}, (\exists m \in \mathbb{N}, n = 2 \times m) \vee (\exists m \in \mathbb{N}, n = 2 \times m + 1)$$

Remarquons tout d'abord que l'on remplace « ou » par le symboles $\vee$, maintenant que nous connaissons le formalisme de la logique propositionnelle. Il nous reste cependant plusieurs points à définir formellement : tout d'abord, la phrase précédente contient des termes, comme 2 ou n. Ceux-ci sont d'une nature différente d'une variable propositionnelle par exemple, puisque dans le premier cas, les formules ne relient pas directement des termes, mais des relations entre termes. Nous devons donc tout d'abord construire un ensemble de termes, qui représenterons les objets dont les formules parleront. Cependant, comme nous cherchons en premier lieu à élaborer des phrases finies, nous cherchons aussi à limiter les symboles que nous utiliserons. Cela s'explique par le fait que pour lire une phrase, il est nécessaire de savoir à l'avance quels sont les symboles constitutifs de ce langage. En particulier, nous devons savoir ce que signifie chaque symbole.

Définition 3.1.1.1 (Signature). Une signature, ou langage, du premier ordre, est un quadruplet $\mathcal{L} = (\mathcal{F}, \mathcal{R}, \alpha, \beta)$ où $\alpha : \mathcal{F} \to \mathbb{N}$ et $\beta : \mathcal{R} \to \mathbb{N}$. On appelle les éléments de $\mathcal{F}$ les symboles de fonction et les éléments de $\mathcal{R}$ les symboles de relation. Pour un symboles de fonction $f \in \mathcal{F}$, $\alpha(f)$ est appelé l'arité de f, et de même $\beta(r)$ est l'arité de r pour $r \in \mathcal{R}$. Si $f \in \mathcal{F}$ est d'arité 0, on dit que c'est une constante.

Exemple. Un premier exemple de langage est le langage des groupes, qui est

$$\mathcal{L}_{\text{Grp}} \triangleq \{e^0, \times^2, ((-)^{-1})^1\}$$

où l'on indique par un exposant l'arité d'un symbole, et où tous les symboles sont des symboles de fonction. De même, comme on préfère la notation additive pour les groupes abéliens, on peut aussi définir

$$\mathcal{L}_{\text{Ab}} \triangleq \{0^0, +^2, -^1\}$$

Exemple. Un autre exemple est le langage des anneaux :

$$\mathcal{L}_{\text{Ring}} \triangleq \{0^0, 1^0, +^2, \times^2, -^1\}$$

Exemple. Un autre exemple classique de langage est celui de l'arithmétique :

$$\mathcal{L}_{\text{Arith}} \triangleq \{0^0, S^1, +^2, \times^2, \leq^2\}$$

où $\leq$ est un symbole de relation, et les autres symboles sont des symboles de fonction.

L'exemple du langage de l'arithmétique permet de voir ce que nous entendons par termes : avec ce langage, nous avons envie de pouvoir écrire 0 (qui est une constante) mais aussi 1 défini par $S\ 0$ ou $S\ S\ 0$. De plus, il doit être possible d'écrire $(S\ S\ 0) + (S\ 0)$ par exemple : l'écriture est donc naturellement donnée comme un ensemble inductif, où les arités des symboles de fonction nous donnent les arités des constructeurs de l'ensemble.

3.1.2 Termes et formules

Définition 3.1.2.1 (Termes). Soit une signature $\Sigma = (\mathcal{F}_\Sigma, \mathcal{R}_\Sigma, \alpha_\Sigma, \beta_\Sigma)$, on définit $\mathrm{Term}(\Sigma)$ comme l'ensemble inductif engendré par $\mathcal{F}_\Sigma \cup \mathrm{Var}$ où l'arité de $f \in \mathcal{F}_\Sigma$ est $\alpha_\Sigma(f)$ et où l'arité de $x \in \mathrm{Var}$ est 0. On peut représenter cet ensemble par la grammaire suivante :

$$t, u ::= x \mid f(t_1, \ldots, t_{\alpha(f)})$$

où $x \in \mathrm{Var}$ et $f \in \mathcal{F}_\Sigma$.

Ainsi, les termes écrits dans notre langage vont représenter les objets mathématiques sur lesquels porteront nos formules. Ces formules sont définies par induction, d'une façon analogue aux propositions de la logique propositionnelle. En l'absence de variables propositionnelles, les éléments atomiques des formules seront construits à partir des termes.

Définition 3.1.2.2 (Proposition atomique). Soit une signature Σ. On définit l'ensemble $\mathrm{Atom}(\Sigma)$ des propositions atomiques par

$$\mathrm{Atom}(\Sigma) \triangleq \{(r, t_1, \ldots, t_k) \mid r \in \mathcal{R}_\Sigma, (t_1, \ldots, t_k) \in (\mathrm{Term}(\Sigma))^k, k = \beta_\Sigma(r)\} \cup \{(\text{“}=\text{”}, t, u) \mid t, u \in \mathrm{Term}(\Sigma)\}$$

où $=$ est un symbole n'appartenant pas à $\mathcal{R}_\Sigma$.

Remarque 3.1.2.3. L'égalité est une relation, mais celle-ci n'appartient pas formellement au langage, car son comportement est donné par les règles logiques, de la même façon que $\vee$ et $\wedge$ ont leur sens imposés. Certains auteurs considèrent au contraire que $=$ doit être ajouté au langage, en tant que symbole de relation binaire, et d'autres font la différence entre un langage égalitaire (incluant le symbole $=$) et un langage non égalitaire. Notre choix est motivé à la fois par la simplicité et par l'expressivité : l'égalité est clairement utile pour formaliser les mathématiques et raisonner dessus, mais chercher à préciser quand nous l'utilisons ne l'est pas, étant donné qu'elle sera toujours présente.

Nous pouvons maintenant définir l'ensemble des formules sur une signature donnée.

Définition 3.1.2.4 (Formules). Soit une signature Σ. On définit l'ensemble $\mathrm{Form}(\Sigma)$ par la grammaire suivante :

$$\varphi, \psi ::= a \mid \top \mid \bot \mid \neg\varphi \mid \varphi \vee \psi \mid \varphi \wedge \psi \mid \varphi \to \psi \mid \forall x, \varphi \mid \exists x, \varphi$$

où $a \in \mathrm{Atom}(\Sigma)$ et $x \in \mathrm{Var}$.

3.1.3 Variables et substitution

Maintenant que les formules sont définies, nous voulons définir les opérations basiques sur celles-ci. Tout d'abord, nous devons introduire les notions élémentaires liées aux variables.

Définition 3.1.3.1 (Variable libre, formule close). On définit la fonction qui étant donné un terme (respectivement une formule), retourne l'ensemble des variables libres y apparaissant. La fonction VL est définie par induction sur $\mathrm{Term}(\Sigma)$ (respectivement $\mathrm{Form}(\Sigma)$) par les équations suivantes :

- si $t = x \in Var$, alors $\mathrm{VL}(t) = \{x\}$.
- si $t = f(t_1, \ldots, t_n)$ où $f \in \mathcal{F}_\Sigma$, $t_1, \ldots, t_n \in \mathrm{Term}(\Sigma)$, alors $\mathrm{VL}(t) = \bigcup_{k=1}^{n} \mathrm{VL}(t_k)$.

- si $\varphi = r(t_1, \ldots, t_n)$ est une proposition atomique où $r \in \mathcal{R}_\Sigma, t_1, \ldots, t_n \in \mathrm{Term}(\Sigma)$, alors $\mathrm{VL}(\varphi) = \bigcup_{k=1}^{n} \mathrm{VL}(t_k)$.
- si $\varphi = \top$, alors $\mathrm{VL}(\varphi) = \varnothing$.
- si $\varphi = \bot$, alors $\mathrm{VL}(\varphi) = \varnothing$.
- si $\varphi = \neg\psi$, alors $\mathrm{VL}(\varphi) = \mathrm{VL}(\psi)$.
- si $\varphi = \psi \vee \chi$, $\varphi = \psi \wedge \chi$ ou $\varphi = \psi \to \chi$, alors $\mathrm{VL}(\varphi) = \mathrm{VL}(\psi) \cup \mathrm{VL}(\chi)$.
- si $\varphi = \forall x, \psi$ ou $\varphi = \exists x, \psi$, alors $\mathrm{VL}(\varphi) = \mathrm{VL}(\psi) \backslash \{x\}$.

On dit que φ est une formule close si $\mathrm{VL}(\varphi) = \varnothing$. On note par $\mathrm{Clos}(\Sigma)$ l'ensemble des formules closes sur la signature Σ. De même, un terme t est appelé un terme clos s'il n'a pas de variables libres (c'est-à-dire s'il n'a pas de variable).

Remarque 3.1.3.2. Une variable non libre est dite liée : les variables liées sont muettes, elles n'ont pas d'importance en elle-même et seulement sur le quantificateur qui les lie. On considère implicitement que si φ et ψ diffèrent seulement en remplaçant la variable liée par un quantificateur et les variables que ce quantificateur lie, alors $\varphi = \psi$. Par exemple, $\forall x, x = x$ et $\forall y, y = y$ sont identifiées.

Comme pour la logique propositionnelle, la valeur de vérité d'une formule va dépendre de la valeur associée aux variables libres. Cependant, pour l'instant, nous n'avons pas de système clair d'évaluation : nous verrons comment évaluer une formule quand nous aborderons la notion de modèle. Au niveau syntaxique, cependant, nous pouvons déjà introduire la substitution, que l'on peut voir comme une évaluation syntaxique : on remplace une variable libre par un terme.

Définition 3.1.3.3 (Substitution). Soient une signature Σ, un terme $t \in \mathrm{Term}(\Sigma)$ et une variable $x \in \mathrm{Var}$. On définit les deux fonctions

$$\begin{array}{rcccc} -[t/x] & : & \mathrm{Term}(\Sigma) & \longrightarrow & \mathrm{Term}(\Sigma) \\ & & u & \longmapsto & u[t/x] \end{array}$$

$$\begin{array}{rcccc} -[t/x] & : & \mathrm{Form}(\Sigma) & \longrightarrow & \mathrm{Form}(\Sigma) \\ & & \varphi & \longmapsto & \varphi[t/x] \end{array}$$

par induction sur la structure de $\mathrm{Term}(\Sigma)$ (respectivement sur la structure de $\mathrm{Form}(\Sigma)$) :

- si $u = x$, alors $u[t/x] = t$.
- si $u = y \in Var$ avec $y \neq x$, alors $u[t/x] = y$.
- si $u = f(u_1, \ldots, u_n)$, alors $u[t/x] = f(u_1[t/x], \ldots, u_n[t/x])$.
- si $\varphi = r(u_1, \ldots, u_n)$, alors $\varphi[t/x] = r(u_1[t/x], \ldots, u_n[t/x])$.
- si $\varphi = \top$ alors $\varphi[t/x] = \top$.
- si $\varphi = \bot$ alors $\varphi[t/x] = \bot$.
- si $\varphi = \neg\psi$ alors $\varphi[t/x] = \neg\psi[t/x]$.
- si $\varphi = \psi \vee \chi$ alors $\varphi[t/x] = \psi[t/x] \vee \chi[t/x]$.
- si $\varphi = \psi \wedge \chi$ alors $\varphi[t/x] = \psi[t/x] \wedge \chi[t/x]$.
- si $\varphi = \psi \to \chi$ alors $\varphi[t/x] = \psi[t/x] \to \chi[t/x]$.
- si $\varphi = \forall z, \psi$ où $z \notin \mathrm{VL}(t)$, alors $\varphi[t/x] = \forall z, \psi[t/x]$.

- si $\varphi = \exists z, \psi$ où $z \notin \mathrm{VL}(t)$, alors $\varphi[t/x] = \exists z, \psi[t/x]$.

Remarque 3.1.3.4. La condition de $z \notin \mathrm{VL}(t)$ dans les derniers cas peut toujours être réalisée quitte à renommer la variable liée z : puisque Var est infini et que $\mathrm{VL}(t)$ est fini, on peut toujours trouver $a \notin \mathrm{VL}(t)$ et remplacer $\forall z, \psi$ par $\forall a, \psi[a/z]$ en utilisant l'identification de la remarque 3.1.3.2.

Si l'on veut être parfaitement formel, il conviendrait de procéder dans l'autre sens : on définit d'abord la substitution comme donnée précédemment, puis on définit la relation $\varphi \equiv \psi$ engendrée par $\forall x, \psi \equiv \forall y, \psi[y/x]$ et $\exists x, \psi \equiv \exists y, \psi[y/x]$ dont on prouve qu'elle est une relation d'équivalence, puis on définit le « vrai » ensemble $\mathrm{Form}(\Sigma)$ par $\mathrm{Form}(\Sigma)/\equiv$ (cela n'est pas nécessaire pour $\mathrm{Term}(\Sigma)$ puisque toute variable est libre, dans un terme), et que la fonction $-[t/x]$ est bien définie sur ce quotient. Ce processus est évidemment plus laborieux et n'apporte rien à la compréhension, c'est pourquoi nous ne le détaillons pas ici.

Exercice 3.1.3.5. Soit $x \in \mathrm{Var}$, $t \in \mathrm{Term}(\Sigma)$ et $\varphi \in \mathrm{Clos}(\Sigma)$ pour une signature Σ quelconque. Montrer que $\varphi[t/x] = \varphi$.

Exercice 3.1.3.6. Soient $t, u, v \in \mathrm{Term}(\Sigma)$ et $x, y \in \mathrm{Var}$, montrer que

$$t[u/x][v/y] = (t[v/y])[u[v/y]/x]$$

3.2 Bases de théorie des modèles

Maintenant que nous avons défini la syntaxe élémentaire, nous allons lui donner un sens : une sémantique. Dans le cas de la logique propositionnelle, le sens d'une proposition était simple à définir, puisqu'il s'agissait d'une valeur de vérité en fonction des variables propositionnelles. Dans le cas de la logique du premier ordre, les propositions parlent d'objets, et il faut donc fixer un univers ambiant sur lequel porte le discours donné par les formules. Par exemple, en prenant le langage de l'arithmétique, le terme $S\,S\,0 + S\,S\,0$ et en considérant 0 comme l'entier naturel 0, S comme la fonction $n \mapsto n+1$ et $+$ comme l'addition usuelle sur les entiers, le terme devient le terme 4, mais on peut imaginer une autre interprétation de ce terme donnant par exemple 5 ou tout autre nombre.

3.2.1 Structure et interprétation

Nous travaillons donc sur l'interprétation d'une formule en deux parties : tout d'abord, nous introduisons la notion de structure, qui offre une interprétation claire du langage dans lequel la formule est écrite, et c'est seulement à partir de cette interprétation que l'on peut évaluer une formule. Cela modifie notre notion de vérité : les formules closes prennent une plus grande importance que le reste des formules, mais il faut quantifier sur des structures en contrepartie.

Définition 3.2.1.1 (Structure). Soit une signature Σ. On appelle Σ-structure (ou simplement structure) $\mathcal{M}$ un triplet $(|\mathcal{M}|, -^{\mathcal{M}}_{\mathcal{F}}, -^{\mathcal{M}}_{\mathcal{R}})$ (on notera les deux $-^{\mathcal{M}}$, sans indice) où :

- $|\mathcal{M}|$ est un ensemble.
- pour chaque $f \in \mathcal{F}_\Sigma$ d'arité n, $f^{\mathcal{M}} : |\mathcal{M}|^n \to |\mathcal{M}|$.
- pour chaque $r \in \mathcal{R}_\Sigma$ d'arité n, $r^{\mathcal{M}} \subseteq |\mathcal{M}|^n$.

On identifie $|\mathcal{M}|^0 \to |\mathcal{M}|$ à $|\mathcal{M}|$: un symbole de constante est associé directement à un élément.

Exemple. En reprenant les différents langages définis précédemments, on peut voir que, par exemple, $(\mathbb{Z}, 0, +, -)$ est une structure sur le langage des groupes. C'est même, en incluant 1 et $\times$, une structure sur le langage des anneaux. De même, $(\mathbb{N}, 0, n \mapsto n+1, +, \times, \leq)$ est une structure sur le langage de l'arithmétique.

Avec ces nouveaux exemples, on voit qu'il devient naturel d'interpréter dans la structure $(\mathbb{N}, 0, n \mapsto n+1, +, \times, \leq)$ le terme $S\ S\ 0 + S\ S\ 0$ par l'élément $4 \in \mathbb{N}$. Nous pouvons donc généraliser ce résultat. Pour cela, on définit d'abord la notion d'environnement, puis de valuation étant donné un environnement.

Définition 3.2.1.2 (Environnement). Soit une signature Σ et une structure $\mathcal{M}$. Un environnement ρ est une fonction partielle $\rho : \mathrm{Var} \rightharpoonup |\mathcal{M}|$. On note $\mathcal{E}$ l'ensemble des environnements. Etant donnés un élément $m \in |\mathcal{M}$, une variable x et un environnement ρ, on note $\rho[x \mapsto m]$ l'environnement coïncidant avec ρ sur $\mathrm{Var} \setminus \{x\}$ et valant m en x.

Définition 3.2.1.3 (Interprétation, valuation). Soit une signature Σ, une structure $\mathcal{M}$ et un environnement ρ. On définit par induction sur t (respectivement φ), où $\mathrm{VL}(t) \subseteq \mathrm{Dom}(\rho)$ (respectivement $\mathrm{VL}(\varphi) \subseteq \mathrm{Dom}(\rho)$) les fonctions suivantes :

$$\begin{array}{rcccc} -_\rho^{\mathcal{M}} & : & \mathrm{Term}(\Sigma) & \longrightarrow & |\mathcal{M}| \\ & & t & \longmapsto & t_\rho^{\mathcal{M}} \end{array}$$

$$\begin{array}{rcccc} \mathrm{Val}_\rho & : & \mathrm{Form}(\Sigma) & \longrightarrow & \{0,1\} \\ & & \varphi & \longmapsto & \mathrm{Val}_\rho(\varphi) \end{array}$$

- si $t = x \in \mathrm{Var}$, alors $t_\rho^{\mathcal{M}} = \rho(x)$.
- si $t = f(t_1, \ldots, t_n)$, alors $t_\rho^{\mathcal{M}} = f^{\mathcal{M}}((t_1)_\rho^{\mathcal{M}}, \ldots, (t_n)_\rho^{\mathcal{M}})$.
- si $\varphi = r(t_1, \ldots, t_n)$ alors $\mathrm{Val}_\rho(\varphi) = \chi_{r^{\mathcal{M}}}((t_1)_\rho^{\mathcal{M}}, \ldots, (t_n)_\rho^{\mathcal{M}})$. On interprète le symbole $=$ par la partie $\{(m, m) \mid m \in |\mathcal{M}|\}$.
- si $\varphi = \top$, alors $\mathrm{Val}_\rho(\varphi) = 1$.
- si $\varphi = \bot$, alors $\mathrm{Val}_\rho(\varphi) = 0$.
- si $\varphi = \neg\psi$, alors $\mathrm{Val}_\rho(\varphi) = 1 - \mathrm{Val}_\rho(\psi)$.
- si $\varphi = \psi \vee \chi$, alors $\mathrm{Val}_\rho(\varphi) = \max(\mathrm{Val}_\rho(\psi), \mathrm{Val}_\rho(\chi))$
- si $\varphi = \psi \wedge \chi$, alors $\mathrm{Val}_\rho(\varphi) = \min(\mathrm{Val}_\rho(\psi), \mathrm{Val}_\rho(\chi))$
- si $\varphi = \psi \to \chi$, alors $\mathrm{Val}_\rho(\varphi) = \max(1 - \mathrm{Val}_\rho(\psi), \mathrm{Val}_\rho(\chi))$
- si $\varphi = \forall x, \psi$, alors $\mathrm{Val}_\rho(\varphi) = \min\limits_{m \in |\mathcal{M}|} (\mathrm{Val}_{\rho[x \mapsto m]}(\psi))$
- si $\varphi = \exists x, \psi$, alors $\mathrm{Val}_\rho(\varphi) = \max\limits_{m \in |\mathcal{M}|} (\mathrm{Val}_{\rho[x \mapsto m]}(\psi))$

Si t est un terme clos, alors $t^{\mathcal{M}}$ est un élément de $|\mathcal{M}|$. Si φ est une formule close, alors $\mathrm{Val}(\varphi)$ est un élément de $\{0, 1\}$. Ces deux éléments sont obtenus en interprétant le terme (respectivement la formule) dans le contexte vide, ou de façon équivalente dans n'importe quel contexte.

Enfin, donnons un résultat statuant que la substitution et l'interprétation commutent.

Proposition 3.2.1.4. *Soit une signature Σ, une structure $\mathcal{M}$, un environnement ρ, une variable x, deux termes t, u et une formule φ. Alors on a les deux égalités suivantes :*

$$(t[u/x])_\rho^{\mathcal{M}} = t_{\rho[x \mapsto u_\rho^{\mathcal{M}}]}^{\mathcal{M}} \qquad \mathrm{Val}_\rho(\varphi[u/x]) = \mathrm{Val}_{\rho[x \mapsto u_\rho^{\mathcal{M}}]}(\varphi)$$

Démonstration. On montre la première égalité par induction sur t, et l'autre par induction sur φ :

- si $t = x$ alors $(t[u/x])^{\mathcal{M}}_{\rho} = u^{\mathcal{M}}_{\rho}$ d'où l'égalité.
- si $t = y$ où $y \in \mathrm{Var}$ et $y \neq x$, alors $(t[u/x])^{\mathcal{M}}_{\rho} = \bot$ si $y \notin \mathrm{Dom}(\rho)$, ce qui est aussi la valeur de $y^{\mathcal{M}}_{\rho[x \mapsto u^{\mathcal{M}}_{\rho}]}$, ou bien $\rho(y)$, qui est là encore la même valeur pour l'autre partie de l'équation.
- si $t = f(t_1, \ldots, t_n)$ où pour tout $i \in \{1, \ldots, n\}$, $(t_i[u/x])^{\mathcal{M}}_{\rho} = (t_i)^{\mathcal{M}}_{\rho[x \mapsto u^{\mathcal{M}}_{\rho}]}$, alors

$$\begin{aligned}(t[u/x])^{\mathcal{M}}_{\rho} &= f((t_1[u/x])^{\mathcal{M}}_{\rho}, \ldots, (t_n[u/x])^{\mathcal{M}}_{\rho}) \\ &= f((t_1)^{\mathcal{M}}_{\rho[x \mapsto u^{\mathcal{M}}_{\rho}]}, \ldots, (t_n)^{\mathcal{M}}_{\rho[x \mapsto u^{\mathcal{M}}_{\rho}]}) \\ &= t^{\mathcal{M}}_{\rho[x \mapsto u^{\mathcal{M}}_{\rho}]}\end{aligned}$$

- si $\varphi = \top$ ou $\varphi = \bot$, l'égalité est directe.
- si $\varphi = r(t_1, \ldots, t_n)$ est une proposition atomique, alors

$$\begin{aligned}\mathrm{Val}_{\rho}(r(t_1, \ldots, t_n)[u/x]) &= \mathrm{Val}_{\rho}(r(t_1[u/x], \ldots, t_n[u/x])) \\ &= \chi_r((t_1[u/x])^{\mathcal{M}}_{\rho}, \ldots, (t_n[u/x])^{\mathcal{M}}_{\rho}) \\ &= \chi_r((t_1)^{\mathcal{M}}_{\rho[x \mapsto u^{\mathcal{M}}_{\rho}]}, \ldots, (t_n)^{\mathcal{M}}_{\rho[x \mapsto u^{\mathcal{M}}_{\rho}]}) \\ &= \mathrm{Val}_{\rho[x \mapsto u^{\mathcal{M}}_{\rho}]}(r(t_1, \ldots, t_n))\end{aligned}$$

- si $\varphi = \neg\psi$, alors

$$\begin{aligned}\mathrm{Val}_{\rho}((\neg\psi)[u/x]) &= \mathrm{Val}_{\rho}(\neg(\psi[u/x])) \\ &= 1 - \mathrm{Val}_{\rho}(\psi[u/x]) \\ &= 1 - \mathrm{Val}_{\rho[x \mapsto u^{\mathcal{M}}_{\rho}]}(\psi) \\ &= \mathrm{Val}_{\rho[x \mapsto u^{\mathcal{M}}_{\rho}]}(\neg\psi)\end{aligned}$$

- si $\varphi = \psi \vee \chi$, alors

$$\begin{aligned}\mathrm{Val}_{\rho}((\psi \vee \chi)[u/x]) &= \mathrm{Val}_{\rho}(\psi[u/x] \vee \chi[u/x]) \\ &= \max(\mathrm{Val}_{\rho}(\psi[u/x]), \mathrm{Val}_{\rho}(\chi[u/x])) \\ &= \max(\mathrm{Val}_{\rho[x \mapsto u^{\mathcal{M}}_{\rho}]}(\psi), \mathrm{Val}_{\rho[x \mapsto u^{\mathcal{M}}_{\rho}]}(\chi)) \\ &= \mathrm{Val}_{\rho[x \mapsto u^{\mathcal{M}}_{\rho}]}(\psi \vee \chi)\end{aligned}$$

- si $\varphi = \psi \wedge \chi$, alors

$$\begin{aligned}\mathrm{Val}_{\rho}((\psi \wedge \chi)[u/x]) &= \mathrm{Val}_{\rho}(\psi[u/x] \wedge \chi[u/x]) \\ &= \min(\mathrm{Val}_{\rho}(\psi[u/x]), \mathrm{Val}_{\rho}(\chi[u/x])) \\ &= \min(\mathrm{Val}_{\rho[x \mapsto u^{\mathcal{M}}_{\rho}]}(\psi), \mathrm{Val}_{\rho[x \mapsto u^{\mathcal{M}}_{\rho}]}(\chi)) \\ &= \mathrm{Val}_{\rho[x \mapsto u^{\mathcal{M}}_{\rho}]}(\psi \wedge \chi)\end{aligned}$$

- si $\varphi = \forall y, \psi$ (sans perte de généralité, $y \notin \mathrm{VL}(u)$), alors

$$\begin{aligned}
\mathrm{Val}_\rho((\forall y, \psi)[u/x]) &= \mathrm{Val}_\rho(\forall y, \psi[u/x]) \\
&= \min_{m \in |\mathcal{M}|} (\mathrm{Val}_{\rho[y \mapsto m]}(\psi[u/x])) \\
&= \min_{m \in |\mathcal{M}|} (\mathrm{Val}_{\rho[y \mapsto m][x \mapsto u^{\mathcal{M}}_{\rho[y \mapsto m]}]}(\psi)) \\
&= \min_{m \in |\mathcal{M}|} (\mathrm{Val}_{\rho[y \mapsto m][x \mapsto u^{\mathcal{M}}_{\rho}]}(\psi)) \\
&= \min_{m \in |\mathcal{M}|} (\mathrm{Val}_{\rho[x \mapsto u^{\mathcal{M}}_{\rho}][y \mapsto m]}(\psi)) \\
&= \mathrm{Val}_{\rho[x \mapsto u^{\mathcal{M}}_{\rho}]}(\forall y, \psi)
\end{aligned}$$

- si $\varphi = \exists y, \psi$ (sans perte de généralité, $y \notin \mathrm{VL}(u)$), alors

$$\begin{aligned}
\mathrm{Val}_\rho((\exists y, \psi)[u/x]) &= \mathrm{Val}_\rho(\exists y, \psi[u/x]) \\
&= \max_{m \in |\mathcal{M}|} (\mathrm{Val}_{\rho[y \mapsto m]}(\psi[u/x])) \\
&= \max_{m \in |\mathcal{M}|} (\mathrm{Val}_{\rho[y \mapsto m][x \mapsto u^{\mathcal{M}}_{\rho[y \mapsto m]}]}(\psi)) \\
&= \max_{m \in |\mathcal{M}|} (\mathrm{Val}_{\rho[y \mapsto m][x \mapsto u^{\mathcal{M}}_{\rho}]}(\psi)) \\
&= \max_{m \in |\mathcal{M}|} (\mathrm{Val}_{\rho[x \mapsto u^{\mathcal{M}}_{\rho}][y \mapsto m]}(\psi)) \\
&= \mathrm{Val}_{\rho[x \mapsto u^{\mathcal{M}}_{\rho}]}(\exists y, \psi)
\end{aligned}$$

D'où le résultat par induction. □

3.2.2 Satisfaction, modèle

La notion de valuation permet de définir la relation de satisfaction, $\models$, d'une façon analogue à ce que nous avons fait pour la logique propositionnelle.

Définition 3.2.2.1 (Satisfaction). Soit une signature Σ, une structure $\mathcal{M}$ et, une formule φ et un environnement ρ tel que $\mathrm{VL}(\varphi) \subseteq \mathrm{Dom}(\rho)$. On définit $\mathcal{M}, \rho \models \varphi$ par

$$\mathcal{M}, \rho \models \varphi \triangleq \mathrm{Val}_\rho(\varphi) = 1$$

Soit un ensemble $\mathcal{F} \subseteq \mathrm{Form}(\Sigma)$ et un environnement ρ tel que $\forall \varphi \in \mathcal{F}, \mathrm{VL}(\varphi) \subseteq \mathrm{Dom}(\rho)$. On dit que $\mathcal{M}, \rho$ satisfont $\mathcal{F}$, ce que l'on écrit $\mathcal{M}, \rho \models \mathcal{F}$, lorsque pour toute formule $\varphi \in \mathcal{F}$, il est vrai que $\mathcal{M}, \rho \models \varphi$.

Dans le cas de formules closes, on écrira directement $\mathcal{M} \models \varphi$ et $\mathcal{M} \models \mathcal{F}$.

Cela permet alors de définir ce qu'est un modèle.

Définition 3.2.2.2 (Modèle). Soit une signature Σ et $\mathcal{C} \subseteq \mathrm{Clos}(\Sigma)$. On dit que $\mathcal{M}$ est un modèle de $\mathcal{C}$ si $\mathcal{M} \models \mathcal{C}$.

Un modèle ne se définit qu'avec un ensemble de formules closes. On pourrait imaginer une définition analogue avec une formule non close, mais faire cela signifie qu'au lieu de donner une structure, il faudrait donner une structure et un environnement en même temps. Le but des modèles est plutôt, ici, de construire une classe particulière de structure vérifiant certaines conditions que l'on peut exprimer au premier ordre.

Un exemple simple est la formule

$$\forall x, \forall y, x + y = y + x$$

Les modèles de cette formule, sur le langage $\{+^2\}$, sont les magmas commutatifs (ensembles munis d'une loi de composition interne commutative) : l'intérêt ici est de pouvoir décrire parmi tous les magmas possibles ceux qui ont une lci commutative, et donc de le faire en quelque sorte uniformément parmi les structures, ce qui n'est pas le cas si les formules n'étaient pas closes.

3.2.3 Théories et conséquence logique

Cela mène naturellement à deux notions connexes : celle de théorie, et celle de conséquence logique. Une théorie permet de décrire des classes de modèles, et la conséquence logique permet de créer des liens entre les formules, de la même façon que nous avions $\vDash$ pour le calcul propositionnel.

Définition 3.2.3.1 (Théorie). Une théorie axiomatique, ou simplement théorie, sur une signature Σ, est une partie $\mathcal{T} \subseteq \mathrm{Clos}(\Sigma)$.

Définition 3.2.3.2 (Conséquence logique, équivalence). Soit un ensemble $\mathcal{F} \subseteq \mathrm{Form}(\Sigma)$ et une formule $F \in \mathrm{Form}(\Sigma)$. On dit que F est conséquence logique de $\mathcal{F}$, ce que l'on écrit $\mathcal{F} \vDash F$, lorsque pour toute structure $\mathcal{M}$, si $\mathcal{M} \models \mathcal{F}$ alors $\mathcal{M} \models F$. Si deux formules F et G sont telles que $F \vDash G$ et $G \vDash F$, alors F et G sont dites logiquement équivalentes, ce que l'on note $F \equiv G$.

On peut relier ces deux notions par celle de théorie saturée.

Définition 3.2.3.3 (Théorie saturée, clôture par conséquence). Soit une théorie $\mathcal{T}$ sur une signature Σ. On dit que $\mathcal{T}$ est saturée si pour toute formule $A \in \mathrm{Form}(\Sigma)$, si $\mathcal{T} \vDash A$ alors $A \in \mathcal{T}$.

Pour une théorie $\mathcal{T}$, on définit sa clôture par conséquence, notée $\overline{\mathcal{T}}^{\vDash}$, par

$$\overline{\mathcal{T}}^{\vDash} \triangleq \{A \in \mathrm{Form}(\Sigma) \mid \mathcal{T} \vDash A\}$$

De plus, la conséquence logique permet directement de décrire une théorie « fausse » : une telle théorie est une théorie dans laquelle la proposition fausse est considérée comme vraie. Au niveau des modèles, cela se traduit par le fait qu'il n'existe pas de modèle de la théorie, puisqu'un tel modèle associerait automatique à $\bot$ la valeur de vérité 0.

Proposition 3.2.3.4. *Soit une signature Σ et une théorie $\mathcal{T}$ sur Σ. Alors $\mathcal{T}$ admet un modèle si et seulement si $\mathcal{T} \nvDash \bot$.*

Démonstration. Supposons que $\mathcal{T}$ admette un modèle $\mathcal{M}$. Alors par définition de Val, $\mathrm{Val}(\bot) = 0$ donc $\mathcal{M} \not\models \bot$. On en déduit que $\mathcal{T} \nvDash \bot$.

Dans le sens réciproque et par contraposée, supposons que $\mathcal{T}$ n'admet pas de modèle. Alors pour tout $\mathcal{M}$ tel que $\mathcal{M} \models \mathcal{T}$, $\mathcal{M} \models \bot$, par vacuité de la condition : on en déduit donc que $\mathcal{T} \vDash \bot$, donc que si $\mathcal{T} \nvDash \bot$, alors $\mathcal{T}$ a un modèle. □

Définition 3.2.3.5 (Théorie cohérente, contradictoire). Soit une signature Σ et une théorie $\mathcal{T}$ sur Σ. On dit que $\mathcal{T}$ est cohérente quand elle admet un modèle, et qu'elle est contradictoire si elle n'admet pas de modèle. De façon équivalente, $\mathcal{T}$ est cohérente si et seulement si elle n'est pas contradictoire, et si et seulement si $\bot \notin \overline{\mathcal{T}}^{\vDash}$.

On voit donc qu'une théorie ne doit pas pouvoir prouver trop de choses. Néanmoins, on peut vouloir une théorie la plus forte possible, qui reste malgré tout cohérente. Une telle théorie est une théorie complète : elle est une théorie dans laquelle si un énoncé est faux, alors sont contraire est vrai. Elle peut donc décider tout énoncé, et puisqu'elle est complète elle ne peut pas décider plus (sinon il serait possible de vérifier A et $\neg A$, ce qui est impossible).

Définition 3.2.3.6 (Théorie complète). Une théorie $\mathcal{T}$ sur une signature Σ est dite complète si pour toute formule $A \in \text{Form}(\Sigma)$, soit $A \in \overline{\mathcal{T}}^{\models}$ soit $\neg A \in \overline{\mathcal{T}}^{\models}$.

Donnons dès maintenant un résultat essentiel, dont on ne donnera la preuve que dans la sous-section 4.3.2, en tant que corrolaire du théorème 4.3.2.4.

Théorème 3.2.3.7 (Extension complète d'une théorie). *Soit une théorie cohérente $\mathcal{T}$ sur une signature Σ. Il existe une théorie $\mathcal{T}'$ complète contenant $\mathcal{T}$.*

3.2.4 Morphismes de modèles

Pour manipuler efficacement les modèles, il est utile de prendre du recul sur ce qui est manipulé, et d'adopter un point de vue plus algébrique. Pour cela, nous allons donner quelques propriétés basiques liées aux morphismes de modèles.

En mathématiques, un morphisme est en toute généralité une application qui préserve la structure. C'est l'idée qui est formalisée dans la notion de morphisme entre structures : un morphisme commute avec les symboles de relation et de fonction.

Définition 3.2.4.1 (Morphisme de structures). Soit une signature Σ et deux structures $\mathcal{M}, \mathcal{N}$, un morphisme φ de $\mathcal{M}$ vers $\mathcal{N}$ est une application

$$\varphi : |\mathcal{M}| \longrightarrow |\mathcal{N}|$$

vérifiant les propositions suivantes :

- pour tout symbole de fonction f d'arité n et tout tuple $(m_1, \ldots, m_n) \in |\mathcal{M}|^n$:
$$\varphi(f^{\mathcal{M}}(m_1, \ldots, m_n)) = f^{\mathcal{N}}(\varphi(m_1), \ldots, \varphi(m_n))$$
- pour tout symbole de relatoin r d'arité n et tout tuple $(m_1, \ldots, m_n) \in |\mathcal{M}|^n$:
$$r^{\mathcal{M}}(m_1, \ldots, m_n) \implies r^{\mathcal{N}}(\varphi(m_1), \ldots, \varphi(m_n))$$

Exercice 3.2.4.2. Montrer que pour toute signature Σ et toute structure $\mathcal{M}$ sur Σ, la fonction $\text{id}_{|\mathcal{M}|}$ induit un morphisme de $\mathcal{M}$ vers elle-même. On notera $\text{id}_{\mathcal{M}}$ ce morphisme.

Exercice 3.2.4.3. Montre que l'opération $\circ$, de composition, s'étend en une opération sur les morphismes de structures, c'est-à-dire que si f et g sont deux morphismes tels que $\text{Dom}(g) = \text{Im}(f)$, alors $g \circ f$ est aussi un morphisme.

Un renforcement de la notion de morphisme est celle de plongement : un plongement est un morphisme dont l'image est isomorphe au domaine, c'est-à-dire que non seulement le morphisme est injectif, mais il le comportement vis à vis des propositions peut s'étudier dans l'image uniquement en étudiant le modèle de départ.

Définition 3.2.4.4 (Plongement). Un plongement d'une structure $\mathcal{M}$ vers une structure $\mathcal{N}$ est un morphisme pour lequel la deuxième condition n'est plus une implication mais une équivalence :

$$\forall (m_1, \ldots, m_n) \in |\mathcal{M}|^n, r^{\mathcal{M}}(m_1, \ldots, m_n) \iff r^{\mathcal{N}}(\varphi(m_1), \ldots, \varphi(m_n))$$

Remarque 3.2.4.5. Puisque nous avons toujours la relation $=$ dont l'interprétation est l'égalité dans son sens naturel, on en déduit qu'un plongement est injectif.

Enfin, la notion d'isomorphisme est celle à laquelle on s'attend.

Définition 3.2.4.6 (Isomorphisme). Un isomorphisme $\varphi : \mathcal{M} \cong \mathcal{N}$ est un morphisme de $\mathcal{M}$ vers $\mathcal{N}$ tel qu'il existe un morphisme ψ de $\mathcal{N}$ vers $\mathcal{M}$ vérifiant

$$\begin{cases} \varphi \circ \psi = \mathrm{id}_{\mathcal{N}} \\ \psi \circ \varphi = \mathrm{id}_{\mathcal{M}} \end{cases}$$

Proposition 3.2.4.7. *Un morphisme $\varphi : \mathcal{M} \to \mathcal{N}$ est un isomorphisme si et seulement si c'est un plongement surjectif.*

Démonstration. Supposons que A FAIRE □

De plus, un ismorphisme préserve entièrement les formules.

Proposition 3.2.4.8. *Si $\varphi : \mathcal{M} \cong \mathcal{N}$ alors pour toute formule close A, $\mathcal{M} \models A$ si et seulement si $\mathcal{N} \models A$.*

3.2.5 Agrandir des modèles

Pour conclure cette partie introductive à propos des modèles, nous allons définir la notion d'enrichissement.

Définition 3.2.5.1 (Enrichissement). Soit deux signatures $\Sigma \subseteq \Sigma'$, c'est-à-dire telles que tout symbole de Σ est un symbole de Σ' de même arité. Soit une structure $\mathcal{M}'$ sur Σ', alors on dit que $\mathcal{M}'$ est un enrichissement d'une structure $\mathcal{M}$ sur Σ si :

- pour tout symbole de fonction $f \in \Sigma$, $f^{\mathcal{M}} = f^{\mathcal{M}'}$
- pour tout symbole de fonction $r \in \Sigma$, $r^{\mathcal{M}} = r^{\mathcal{M}'}$

On définit l'appauvrissement de $\mathcal{M}'$ sur Σ comme l'unique structure sur Σ dont $\mathcal{M}'$ est l'enrichissement.

Remarque 3.2.5.2. On peut toujours appauvrir une strucutre, mais il n'est pas toujours possible d'enrichir (ou du moins naturellement) une structure sur une extension de sa signature.

Par exemple, le langage des anneaux est un enrichissement du langage des groupes. On remarque qu'un anneau est toujours, en particulier, un groupe additif : nous allons généraliser ce résultat pour une structure sur un enrichissement.

Proposition 3.2.5.3. *Soit deux signatures $\Sigma \subseteq \Sigma'$ et $\mathcal{M}'$ une structure sur Σ', alors pour toute formule $A \in \mathrm{Form}(\Sigma)$ et tout environnement ρ, on a*

$$\mathcal{M}, \rho \models A \iff \mathcal{M}', \rho \models A$$

Démonstration. A FAIRE □

Corollaire 3.2.5.4. *Soit deux signatures $\Sigma \subseteq \Sigma'$, une structure $\mathcal{M}'$ et une théorie $\mathcal{T}$ sur Σ. Alors si $\mathcal{M}' \models \mathcal{T}$, alors en prenant l'appauvrissement $\mathcal{M}$ de $\mathcal{M}'$ sur Σ, $\mathcal{M} \models \mathcal{T}$.*

3.3 Système de démonstration du calcul des prédicats

Nous avons défini la sémantique des formules à travers la notion de modèle. Pour suivre ce que nous avons fait dans le chapitre précédent, nous allons maintenant introduire un formalisme pour décrire mathématiquement des preuves en calcul des prédicats.

Nous avons alors plusieurs choix, car plusieurs formalismes existent. Les trois principaux sont les sytèmes à la Hilbert, le calcul des séquents et la déduction naturelle. Nous avons eu un aperçu du calcul des séquents dans le chapitre précédent, et il pourrait être pertinent de continuer à l'utiliser, mais le choix fait ici est d'introduire un nouveau formalisme : celui de la déduction naturelle. Ce choix ce justifier par deux raisons principales :

- Tout d'abord, il permet d'aborder d'autres systèmes de preuves, puisque nous avons déjà exploré le calcul des séquents. Cet argument est en même temps un contre-argument, puisque cela signifie aussi que l'on peut trouver là une occasion de réexplorer le calcul des séquents pour mieux le comprendre, c'est donc seulement une raison mineure qui nous pousse à le choisir. De plus, le formalisme des systèmes à la Hilbert ne sera pas exploré du tout dans cet ouvrage, car s'il est très simple à définir, il possède peu de propriétés intéressantes à explorer contrairement aux deux autres formalismes et n'a rien de naturel à manipuler.
- La deuxième raison, plus importante, est que la déduction naturelle est en quelque sorte le raisonnement le plus primitif d'un mathématicien. Chaque règle exprime une règle loique tout à fait évidente à l'intuition, particulièrement à celle du mathématicien, et on peut trouver une correspondance importante entre une preuve utilisant la déduction naturelle et une preuve en langage naturel (à la différence évidente que la première est illisible pour un profane quand la deuxième est... illisible aussi pour un profane, mais il y a moins de profanes des mathématiques que de profanes de la déduction naturelle).

Il y a de nombreux avantages à définir une syntaxe pour nos preuves en calcul des prédicats. Le premier est évident : avoir, comme dans le chapitre précédent, un système efficace pour prouver des relations entre formules, ne passant pas par une interprétation et une quantification sur toutes les valuations (et ici, en plus : tous les modèles). Le deuxième, moins évident, sera plus lourd de conséquences : une syntaxe nous permet de voir l'ensemble de l'activité mathématique comme un processus finitaire. Nous n'avons le droit que d'employer des phrases finies en des textes finis pour étudier des objets potentiellement infinis. Si ce fait semble en premier lieu purement philosophique (et il est effectivement important, philosophiquement parlant) il mène aussi à une conséquence importante : le théorème de compacité. Ce théorème est un analogue au théorème 2.2.1.7 dans le cas du calcul des prédicats. Simplement, au lieu de quantifier sur les valuations, nous quantifions sur les modèles, et puisque nous nous occupons en priorité des formules closes cela nous donne un énoncé parlant uniquement de modèles.

3.3.1 Déduction naturelle

Commençons par définir la relation $\vdash$ de conséquence syntaxique. Pour cela, comme pour le calcul des séquents, on va définir un système travaillant sur des listes (par nature finies) plutôt que sur des ensembles. Une différence importante : nous n'allons pas relier deux listes, mais une liste avec une proposition. Ainsi un séquent $\Gamma \vdash A$ signifie directement que, sous les hypothèses listées dans Γ, la proposition A est vraie.

Définition 3.3.1.1 (Déduction naturelle). Soit une signature Σ. On définit la relation $\vdash \subseteq \mathrm{List}(\mathrm{Form}(\Sigma)) \times \mathrm{Form}(\Sigma)$ par induction par les règles suivantes :

$$\frac{A \in \Gamma}{\Gamma \vdash A}\,\mathrm{Ax} \qquad \frac{}{\Gamma \vdash \top}\,\top \qquad \frac{\Gamma, \neg A \vdash \bot}{\Gamma \vdash A}\,\bot_{\mathrm{c}}$$

$$\frac{\Gamma, A \vdash \bot}{\Gamma \vdash \neg A}\,\neg_{\mathrm{i}} \qquad \frac{\Gamma \vdash \neg A \qquad \Gamma \vdash A}{\Gamma \vdash \bot}\,\neg_{\mathrm{e}}$$

$$\frac{\Gamma \vdash A}{\Gamma \vdash A \vee B}\,\vee_{\mathrm{i}}^{\mathrm{g}} \qquad \frac{\Gamma \vdash B}{\Gamma \vdash A \vee B}\,\vee_{\mathrm{i}}^{\mathrm{d}} \qquad \frac{\Gamma \vdash A \vee B \qquad \Gamma, A \vdash C \qquad \Gamma, B \vdash C}{\Gamma \vdash C}\,\vee_{\mathrm{e}}$$

$$\frac{\Gamma \vdash A \qquad \Gamma \vdash B}{\Gamma \vdash A \wedge B}\,\wedge_{\mathrm{i}} \qquad \frac{\Gamma \vdash A \wedge B}{\Gamma \vdash A}\,\wedge_{\mathrm{e}}^{\mathrm{g}} \qquad \frac{\Gamma \vdash A \wedge B}{\Gamma \vdash B}\,\wedge_{\mathrm{e}}^{\mathrm{d}}$$

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash A \to B}\,\to_{\mathrm{i}} \qquad \frac{\Gamma \vdash A \to B \qquad \Gamma \vdash A}{\Gamma \vdash B}\,\to_{\mathrm{e}}$$

$$\frac{\Gamma \vdash A[v/x]}{\Gamma \vdash \forall x, A}\,\forall_{\mathrm{i}}^{\dagger} \qquad \frac{\Gamma \vdash \forall x, A}{\Gamma \vdash A[t/x]}\,\forall_{\mathrm{e}}$$

$$\frac{\Gamma \vdash A[t/x]}{\Gamma \vdash \exists x, A}\,\exists_{\mathrm{i}} \qquad \frac{\Gamma \vdash \exists x, A \qquad \Gamma, A[v/x] \vdash B}{\Gamma \vdash B}\,\exists_{\mathrm{e}}^{\dagger}$$

$$\frac{}{\Gamma \vdash t = t}\,=_{\mathrm{i}} \qquad \frac{\Gamma \vdash A[t/x] \qquad \Gamma \vdash t = u}{\Gamma \vdash A[u/x]}\,=_{\mathrm{e}}$$

Où $x, v \in \mathrm{Var}$ et où $t, u \in \mathrm{Term}(\Sigma)$.

Les règles avec † signifient que $v \notin \mathrm{VL}(\Gamma) \cup \mathrm{VL}(B)$.

De plus, on définit la relation $\vdash \subseteq \mathcal{P}(\mathrm{Form}(\Sigma)) \times \mathrm{Form}(\Sigma)$ par $\mathcal{F} \vdash A$ si et seulement s'il existe $\Gamma \in \mathrm{List}(\mathcal{F})$ tel que $\Gamma \vdash A$.

Exercice 3.3.1.2 (Sur la négation). Montrer que l'on peut prouver les deux séquents suivants :

$$\vdash \neg A \to (A \to \bot) \qquad \vdash (A \to \bot) \to \neg A$$

Montrer de plus que $\neg A \vDash A \to \bot$ et que $A \to \bot \vDash \neg A$.

On peut donc, sans perdre d'expressivité, redéfinir $\neg$ comme l'opération $A \mapsto A \to \bot$. Cela nous permet alors de réduire le nombre de règles dans notre système.

Exercice 3.3.1.3 (Sur l'implication). Montrer que l'équivalence suivante est prouvable :

$$A \to B \dashv\vdash \neg A \vee B$$

où $A \dashv\vdash B$ signifie que $A \vdash B$ et $B \vdash A$.

Exercice 3.3.1.4 (De Morgan). Montrer que les lois de De Morgan sont dérivables :

- $\neg(A \vee B) \dashv\vdash \neg A \wedge \neg B$
- $\neg(A \wedge B) \dashv\vdash \neg A \vee \neg B$
- $\neg(\exists x, A) \dashv\vdash \forall x, \neg A$
- $\neg(\forall x, A) \dashv\vdash \exists x, \neg A$
- $\neg\neg A \dashv\vdash A$

Exercice 3.3.1.5 (Tiers exclu et non contradiction). Montrer les deux équivalences suivantes :

- $A \vee \neg A \dashv\vdash \top$
- $A \wedge \neg A \dashv\vdash \bot$

Remarque 3.3.1.6. A partir des exercices précédents, on peut être tenté de réduire l'ensemble des formules et des règles à un fragment tel que les propositions atomiques, $\neg$, $\vee$ et $\exists$. En effet, toute formule est équivalent à une formule écrite avec ce fragment : on peut donc imaginer que toute autre formule est en fait une simple écriture plus lisible de cette constituée uniquement du fragment restreint.

Nous n'emploierons pas cette méthode de restriction car, structurellement, il n'est pas évident par exemple que $\neg(\neg A \vee \neg B)$, qui est code $A \wedge B$, s'utilise de la même façon au niveau des règles. Si l'on sait que l'on peut se ramener en utilisant certaines règles de l'un à l'autre, il faudrait travailler à montrer en plus que les règles à propos de $A \wedge B$ peut se dériver des règles de $\neg$ et $\vee$ sur $\neg(\neg A \vee \neg B)$. Mais cela motive aussi un déroulement plus lent des preuves par induction et des différents cas, qui s'ils sont laborieux peuvent pour autant être instructifs pour une première lecture.

Le cas de $\neg$ que l'on peut remplacer par $\to \bot$ est différent, car les règles à propos de $\neg$ sont exactement celle de $\to \bot$.

A propos de la vacuité Notre formalisme a un problème essentiel : on peut prouver la proposition $\forall x, A \implies \exists x, A$, qui est fausse dans le modèle vide. Deux façons permettent de régler cet écart entre la syntaxe et la sémantique : changer la syntaxe, ou changer la sémantique. Dans notre cas, nous changeons alors la sémantique en considérant que toute structure (et donc tout modèle) est non vide. Cette restriction n'est pas très limitante, puisque le modèle vide est inintéressant en général. Cependant, celle-ci parait particulièrement artificielle. Pour contrer cela, on peut à la place considérer des séquents enrichis de la forme $\Gamma \mid \Theta \vdash \varphi$ où Γ va être un contexte de variables, Θ un contexte logique et φ la conclusion. Dans ce formalisme, les règles avec † ont, plutôt qu'une restriction, une action sur le contexte des variables, avec par exemple

$$\frac{\Gamma, v \mid \Theta \vdash \varphi[v/x]}{\Gamma \mid \Theta \vdash \forall x, \varphi} \forall_{\mathrm{i}}$$

où $\Gamma \mid \Theta \vdash \forall x, \varphi$ doit être une proposition bien typée, imposant ansi que $v \notin \mathrm{VL}(\Theta)$.

Le fait de gérer les variables est bien plus naturel, étant donné qu'une preuve en langage naturel va toujours tenir compte des variables (en particulier, il semblerait incongru de mentionner une variable non déjà introduite dans le contexte), et dans un contexte avec plusieurs sortes, c'est-à-dire où les variables du premier ordre peuvent appartenir à différents ensembles, ce formalisme gagne en utilité. Dans notre cas, il parait trop lourd de devoir gérer les variables pour simplement pouvoir inclure le cas du modèle vide, c'est pourquoi nous préférons simplement modifier notre sémantique.

Plutôt que de prouver directement des résultats, nous allons nous attarder sur le sens de chaque règle, pour montrer en quoi elles sont intuitives (et donc robustes au niveau de l'évidence qu'elles énoncent) et permettent de retranscrire n'importe quelle preuve (en particulier, toute preuve en langage naturel est virtuellement équivalente à un arbre de preuve en déduction naturelle).

- La règle Ax est surement la plus évidente : si A est une hypothèse, alors on peut en déduire A.
- La règle $\top$ dit simplement que $\top$ est toujours prouvable.

- La règle $\bot_c$ dit que pour prouver A, on peut supposer $\neg A$ est aboutir à une contradiction : c'est le raisonnement par l'absurde, d'où l'ince « c » exprimant que cette règle est propre à la logique classique (nous le verrons, remplacer cette règle par une autre plus faible a des conséquences particulièrement intéressantes).
- La règle $\neg_i$ dit que pour prouver $\neg A$, il suffit de prouver que A aboutit à une absurdité.
- La règle $\neg_e$ dit que prouver A et $\neg A$ en même temps est une absurdité.
- Les règles $\vee_i$ montrent que si on prouve A (respectivement B) alors on prouve $A \vee B$.
- La règle $\vee_e$ montre que pour prouver C à partir de $A \vee B$, il suffit de montrer que $A \to C$ et $B \to C$ ou, comme nous l'avons écrit, que l'on peut prouver C à la fois sous l'hypothèse A et sous l'hypothèse B. C'est un raisonnement par disjonction de cas.
- La règle $\wedge_i$ dit que pour prouver $A \wedge B$, il suffit de prouver A d'une part, et B d'autre part.
- Les règles $\wedge_e$ permettent d'affaiblir une preuve de $A \wedge B$ en une preuve de A (respectivement de B).
- La règle $\to_i$ dit que prouver $A \to B$ signifie prouver B en ajoutant l'hypothèse A.
- La règle $\to_e$ dit que si l'on a prouvé $A \to B$ et A, alors on peut en déduire B. C'est la règle du *modus ponens*.
- La règle $\forall_i$ signifie que pour prouver $\forall x, A$, il suffit de prouver A pour une variable v quelconque à la place de x. La nécessité que $v \notin \mathrm{VL}(\Gamma)$ exprime que ce v est quelconque : aucune hypothèse n'est faite sur celui-ci.
- La règle $\forall_e$ signifie qu'à partir d'une preuve de $\forall x, A$ on peut instancier x à un terme t quelconque pour obtenir une preuve de $A[t/x]$.
- La règle $\exists_i$ permet de déduire une preuve de $\exists x, A$ à partir d'une preuve de $A[t/x]$, pour n'importe quel terme t.
- La règle $\exists_e$ dit qu'à partir d'une proposition de la forme $\exists x, A$, on peut déduire une proposition B en ajoutant dans le contexte $A[v/x]$, où v est quelconque (ce qui se traduit par la condition de $v \notin \mathrm{VL}(\Gamma) \cup \mathrm{VL}(B)$).
- La règle $=_i$ est la réflexivité de l'égalité : un terme est égal à lui-même.
- La règle $=_e$, parfois appelée principe de Leibniz, exprime que si deux termes t et u sont égaux, alors ils vérifient les mêmes formules. On appelle aussi ce principe « indiscernabilité des identités ».

3.3.2 Complétude de la déduction naturelle

Nous allons maintenant montrer que $\vDash$ et $\vdash$ coïncident. Un sens est élémentaire : montrer que $\vdash \subseteq \vDash$, nous allons donc le traiter en premier. C'est la propriété qu'on appelle correction. Elle énonce que ce que notre syntaxe dérive est valide. L'autre direction, disans que tout ce qui est valide est dérivable, est très souvent plus techniques : ce fut le cas pour la logique propositionnelle, c'est encore le cas pour le calcul des prédicats.

On appelle en général, par abus de langage, complétude du système la propriété que $\vdash = \vDash$, plutôt que simplement la propriété $\vDash \subseteq \vdash$. Il est en effet peu pratique de devoir citer deux théorèmes différents lorsque l'on parle de la correspondance des deux relations, c'est pourquoi on préfère englober les deux en un résultat, et comme le sens le plus technique est celui de complétude, c'est celui qu'on utilise pour nommer le théorème.

Comme $\vdash$ est définie par induction, la preuve de correction est directement une induction sur sa structure. Remarquons cependant qu'il est nécessaire de pouvoir prendre en compte

les variables, et donc d'introduire des environnements en plus. Nous allons donc utiliser la proposition 3.2.1.4.

Théorème 3.3.2.1 (Correction). *Soit une signature Σ, une liste $\Gamma \in \mathrm{List}(\mathrm{Form}(\Sigma))$ et une formule $A \in \mathrm{Form}(\Sigma)$. Si $\Gamma \vdash A$ alors, en notant X_Γ l'ensemble des formules dans Γ, pour toute structure $\mathcal{M}$, tout environnement ρ sur $\mathcal{M}$, si $\mathcal{M}, \rho \models X_\Gamma$ alors $\mathcal{M}, \rho \models A$.*

Démonstration. Nous allons prouver ce résultat par induction sur $\Gamma \vdash A$, en supposant introduits $\mathcal{M}$ et ρ :

- Si $A \in \Gamma$, alors il est évident que $\mathcal{M}, \rho \models A$.
- Peu importe les prémisses, $\mathcal{M}, \rho \models \top$.
- Supposons qu'aucun modèle $\mathcal{M}$ et aucun environnement ρ ne sont tels que $\mathcal{M}, \rho \models X_\Gamma \cup \{\neg A\}$. Alors si $\mathcal{M}, \rho \models X_\gamma$, on ne peut pas avoir $\mathcal{M}, \rho \models \neg A$, donc $\mathrm{Val}_\rho(\neg A) = 0$, donc $\mathrm{Val}_\rho(A) = 1$, d'où $\mathcal{M}, \rho \models A$.
- Supposons qu'aucun $(\mathcal{M}, \rho)$ ne vérifie que $\mathcal{M}, \rho \models X_\Gamma \cup \{A\}$, alors si $\mathcal{M}, \rho \models X_\Gamma$, $\mathrm{Val}_\rho(A) = 0$ donc $\mathrm{Val}_\rho(\neg A) = 1$, d'où $\mathcal{M}, \rho \models \neg A$.
- Supposons qu'un modèle de Γ est un modèle à la fois de A et de $\neg A$. Alors $\mathrm{Val}_\rho(A) = 1$ et $\mathrm{Val}_\rho(\neg A) = 1$, mais $\mathrm{Val}_\rho(\neg A) = 1 - \mathrm{Val}_\rho(A)$, donc $0 = 1$: c'est absurde, donc il n'y a pas de modèle de Γ.
- Supposons que si $\mathcal{M}, \rho \models X_\Gamma$ alors $\mathcal{M}, \rho \models A$. Soit $\mathcal{M}, \rho \models X_\Gamma$, par hypothèse $\mathrm{Val}_\rho(A) = 1$, donc $\mathrm{Val}_\rho(A \vee B) = \max(1, \mathrm{Val}_\rho(B))$, donc $\mathcal{M}, \gamma \models A \vee B$.
- L'argument précédent fonctionne exactement de la même manière.
- A partir de maintenant, nous n'expliciterons plus les hypothèses d'induction, ni ce que l'on cherche à prouver, pour gagner de la place. Supposons que $\mathcal{M}, \rho \models X_\Gamma$. Alors par hypothèse d'induction, $\mathcal{M}, \rho \models A \vee B$, donc $\max(\mathrm{Val}_\rho(A), \mathrm{Val}_\rho(B)) = 1$: on en déduit qu'au moins l'un des deux entre A et B est tel que $\mathrm{Val}_\rho = 1$. Sans perte de généralité, supposons que $\mathrm{Val}_\rho(A) = 1$. On sait donc que $\mathcal{M}, \rho \models X_\Gamma \cup \{A\}$, donc par hypothèse d'induction $\mathcal{M}, \rho \models C$.
- Supposons que $\mathcal{M}, \rho \models X_\Gamma$, alors par hypothèse d'induction $\mathrm{Val}_\rho(A) = 1$ et $\mathrm{Val}_\rho(B) = 1$, donc $\mathrm{Val}_\rho(A \wedge B) = \min(1, 1) = 1$, donc $\mathcal{M}, \rho \models A \wedge B$.
- Supposons que $\mathcal{M}, \rho \models X_\Gamma$, alors par hypothèse d'induction $\mathrm{Val}_\rho(A \wedge B) = 1$, donc $\mathrm{Val}_\rho(A) = 1$ car cette valeur est supérieure à $\mathrm{Val}_\rho(A \wedge B)$: donc $\mathcal{M}, \rho \models A$.
- L'argument précédent fonctionne de la même manière.
- Supposons que $\mathcal{M}, \rho \models X_\Gamma$. On travaille par disjonction de cas sur la valeur de $\mathrm{Val}_\rho(A)$:
  - si $\mathrm{Val}_\rho(A) = 0$ alors $\mathrm{Val}_\rho(A \to B) = 1$ donc $\mathcal{M}, \rho \models A \to B$.
  - si $\mathrm{Val}_\rho(A) = 1$ alors $\mathcal{M}, \rho \models X_\Gamma \cup \{A\}$, donc par hypothèse d'induction $\mathcal{M}, \rho \models B$, donc $\mathrm{Val}_\rho(A \to B) = 1$, donc $\mathcal{M}, \rho \models A \to B$.
- Supposons que $\mathcal{M}, \rho \models X_\Gamma$. On sait donc que $\mathrm{Val}_\rho(A \to B) = 1$ et $\mathrm{Val}_\rho(A) = 1$. Ainsi, $\min(0, \mathrm{Val}_\rho(B)) = 1$: on en déduit que $\mathrm{Val}_\rho(B) = 1$, c'est-à-dire que $\mathcal{M}, \rho \models B$.
- Supposons que $\mathcal{M}, \rho \models X_\Gamma$. On veut montrer que $\mathrm{Val}_\rho(\forall x, A) = 1$. Pour cela, soit $m \in |\mathcal{M}|$: on remarque que, par hypothèse d'induction, $\mathcal{M}, \rho \models A[v/x]$ (ici on remarque un point important pour la question de la vacuité : on suppose que ρ est bien définie en v, ce qui peut se faire sans perte de généralité si on a bien un élément dans le modèle), mais à ce moment-là comme v est libre dans Γ, on remarque que $\mathcal{M}, \rho[v \mapsto m] \models X_\Gamma$, soit $\mathcal{M}, \rho[v \mapsto m] \models A[v/x]$, d'où $\mathrm{Val}_{\rho[v \mapsto m]}(A[v/x]) = 1$, mais l'expression de gauche vaut $\mathrm{Val}_{\rho[x \mapsto m]}(A)$. Comme cela fonctionne pour tout m, on en déduit que $\mathrm{Val}_\rho(\forall x, A) = 1$.

- Supposons que $\mathcal{M}, \rho \models X_\Gamma$, on sait donc que $\mathcal{M}, \rho \models \forall x, A$, ce qui signifie en particulier que pour tous, $m \in \mathcal{M}$, $\mathcal{M}, \rho[x \mapsto m] \models A$. Mais alors, en considérant $m = t_\rho^{\mathcal{M}}$, on trouve que $\mathrm{Val}_{\rho[x \mapsto t_\rho^{\mathcal{M}}]}(A) = 1$, mais l'expression de gauche correpsond exactement à $\mathrm{Val}_\rho(A[t/x])$, donc $\mathcal{M}, \rho \models A[t/x]$.
- Supposons que $\mathcal{M}, \rho \models X_\Gamma$, alors $\mathcal{M}, \rho \models A[t/x]$, mais comme $\mathrm{Val}_\rho(A[t/x]) \leq \mathrm{Val}_\rho(\exists x, A)$ on en déduit que $\mathrm{Val}_\rho(\exists x, A) = 1$, donc $\mathcal{M}, \rho \models \exists x, A$.
- Supposons que $\mathcal{M}, \rho \models X_\Gamma$, alors $\mathcal{M}, \rho \models \exists x, A$, donc on peut trouver un élément $m \in |\mathcal{M}|$ tel que $\mathrm{Val}_{\rho[x \mapsto m]}(A) = 1$. Alors, en considérant $\rho[v \mapsto m]$, on a $\mathcal{M}, \rho \models X_\Gamma$, donc $Val_{\rho[x \mapsto m][v \mapsto m]}(A) = 1$, ce qui revient à dire que $\mathrm{Val}_{\rho[x \mapsto m]}(A[v/x]) = 1$, donc $\mathrm{Val}_{\rho[x \mapsto m]}(B) = 1$. Mais on sait que $x \notin \mathrm{VL}(B)$, donc $\mathrm{Val}_\rho(B) = 1$: ainsi, $\mathcal{M}, \rho \models B$.
- Pour tout modèle $\mathcal{M}$ et environnement ρ, on a forcément $\mathrm{Val}_\rho(t = t) = 1$ puisque $(t_\rho^{\mathcal{M}}, t_\rho^{\mathcal{M}}) \in \{(m, m) \mid m \in |\mathcal{M}|\}$.
- Supposons que $\mathcal{M}, \rho \models X_\Gamma$, alors $\mathcal{M}, \rho \models A[t/x]$ et $\mathcal{M}, \rho \models t = u$. On en déduit que $\mathrm{Val}_\rho(t = u) = 1$, c'est-à-dire que $t_\rho^{\mathcal{M}} = u_\rho^{\mathcal{M}}$, donc $\mathrm{Val}_{\rho[x \mapsto t_\rho^{\mathcal{M}}]}(A) = \mathrm{Val}_{\rho[x \mapsto u_\rho^{\mathcal{M}}]}(A)$, et à partir du fait que $\mathcal{M}, \rho \models A[t/x]$ on en déduit donc que $\mathcal{M}, \rho \models A[u/x]$.

Ainsi, par induction, si $\mathcal{M}, \rho \models X_\Gamma$, alors $\mathcal{M}, \rho \models A$. En particulier, si $\mathcal{M} \models X_\Gamma$, alors $\mathcal{M} \models A$, et si pour $\mathcal{F} \subseteq \mathrm{Form}(\Sigma)$, on a $\mathcal{M} \models \mathcal{F}$ et $\mathcal{F} \vdash A$, cela implique donc que $\mathcal{M} \models A$. Ainsi $\vdash \subseteq \models$. □

Il nous reste à prouver le sens réciproque. En réalité, le point critique pour la démonstration est celui de l'existence d'un modèle. Plutôt que de montrer réellement que $\models \subseteq \vdash$, nous allons montrer qu'une théorie consistante, c'est-à-dire une théorie $\mathcal{T}$ telle que $\mathcal{T} \nvdash \bot$, possède un modèle. On va donc commencer par montrer que notre résultat suffira à prouver $\models \subseteq \vdash$.

Lemme 3.3.2.2. *Supposons que pour toute théorie $\mathcal{T}$ telle que $\mathcal{T} \nvdash \bot$, il existe une structure $\mathcal{M}$ telle que $\mathcal{M} \models \mathcal{T}$. Alors $\models \subseteq \vdash$.*

Démonstration. Supposons que $\mathcal{F} \models A$ pour $A \in \mathrm{Form}(\Sigma)$ et $\mathcal{F} \subseteq \mathrm{Form}(\Sigma)$. On voit donc que $\mathcal{F} \cup \{\neg A\}$ n'a pas de modèle : par contraposée de notre hypothèse, cela signifie que $\mathcal{F}, \neg A \vdash \bot$. En appliquant simplement $\bot_c$, on en déduit que $\mathcal{F} \vdash A$. Ainsi $\models \subseteq \vdash$. □

Remarque 3.3.2.3. Nous avons défini une théorie comme un ensemble de formules closes, donc la démonstration précédente n'est pas tout à fait exacte. Une façon de corriger cela est d'enrichir le langage : pour chaque variable x libre dans $\mathcal{F}$ ou A, on ajoute un symbole de constante c_x, et on considère ensuite la théorie où x est remplacé par c_x. On a une correspondance entre l'existence d'un modèle avec x ou avec c_x, puisque notre théorie sans c_x pourra s'évaluer dans le modèle avec $x \mapsto c_x$.

Il nous reste donc à démontrer qu'une théorie consistante admet bien un modèle. Cette construction est technique, c'est pourquoi on va commencer par donner l'idée de la preuve.

L'idée principale est de construire un modèle syntaxique, c'est-à-dire un modèle dont les éléments sont exactement les termes du langage. Ceux-ci seront quotientés par l'égalité, c'est-à-dire que si $\mathcal{T} \vdash t = u$, alors les termes t et u seront identifiés. En construisant un tel modèle $\mathcal{M}$, on va chercher à ce que $\mathcal{T} \vdash \varphi$ si et seulement si $\mathcal{M} \models \varphi$ puisque ce que vérifie $\mathcal{M}$ est exactement ce que $\mathcal{T}$ peut prouver. On a alors besoin de deux caractéristiques essentielles :

- tout d'abord, $\mathcal{T}$ doit être complète. L'ensemble des énoncés vrais dans un modèle est une théorie complète, puisqu'un énoncé est vrai dans un modèle si et seulement si sa négation est fausse. Ainsi, si $\mathcal{T}$ n'est pas complète, ça ne peut pas être la théorie

d'un modèle. Moralement, on peut voir ça comme le fait qu'au moment de construire un modèle, il faut faire des choix parmi les propositions, car certains modèles de $\mathcal{T}$ peuvent vérifier telle proposition ou telle autre, si $\mathcal{T}$ n'est pas complète. Nous avons donc besoin d'étendre notre théorie en une théorie complète.

- ensuite, il reste un problème au niveau des quantifications. Supposons que $\exists x, P$ soit vraie : on veut pouvoir exhiber un élément m tel que $P[m/x]$ est vrai, et nos éléments sont des termes. On en déduit donc qu'il faut pour chaque proposition P avoir un terme t_P correspondant tel que $\exists x, P \implies P[t_P/x]$. Cela n'est pas vrai *a priori*, on va donc chercher à élargir notre langage pour ajouter à chaque fois des constantes témoignant pour une proposition $\exists x, P$ vraie qu'un terme correspond : c'est ce que l'on appelle la méthode des témoins de Henkin.
- une fois cela fait, il ne nous restera plus qu'à appliquer la complétion de la théorie enrichie par témoins de Henkin pour obtenir une théorie $\overline{\mathcal{T}}$ contenant $\mathcal{T}$ et qui nous permettra de construire un modèle.

On va donc définir la complétion par témoins de Henkin.

Définition 3.3.2.4 (Propriété de Henkin). On dit qu'une théorie $\mathcal{T}$ sur une signature Σ a la propriété de Henkin si pour toute formule telle que $\mathcal{T} \vdash \exists x, P$, il existe un terme $c_P \in \mathrm{Term}(\Sigma)$ tel que $\mathcal{T} \vdash P[c_P/x]$.

Le point important de cette propriété est qu'elle peut être vérifiée, au prix d'une augmentation de la théorie et du langage. Il faut cependant vérifier, alors, que la nouvelle théorie ne peut toujours pas prouver $\bot$.

Définition 3.3.2.5 (Clôture par témoins). Soit Σ une signature, et $\mathcal{T}$ une théorie sur Σ. On construit de façon itérative la suite Σ_n et $\mathcal{T}_n$ de signatures et de théories, où $\mathcal{T}_n$ est une théorie sur Σ_n :

- $\Sigma_0 = \Sigma$ et $\mathcal{T}_0 = \mathcal{T}$.
- Si Σ_n et $\mathcal{T}_n$ sont construits, on définit la signature Σ_{n+1} en ajoutant, pour chaque formule $F \in \mathrm{Form}(\Sigma_n)$, une constante c_F. On définit $\mathcal{T}_{n+1}$ en ajoutant à la théorie $\mathcal{T}$ la famille de formules $\{(\exists x, F) \to F[c_F/x]\}_{F \in \mathrm{Form}(\Sigma_n)}$.

On définit alors

$$\overline{\Sigma}^{\mathrm{H}} \triangleq \bigcup_{n \in \mathbb{N}} \Sigma_n \qquad \overline{\mathcal{T}}^{\mathrm{H}} \triangleq \bigcup_{n \in \mathbb{N}} \mathcal{T}_n$$

et $\overline{\mathcal{T}}^{\mathrm{H}}$ est une théorie sur $\overline{\Sigma}^{\mathrm{H}}$.

Propriété 3.3.2.6. *La théorie $\overline{\mathcal{T}}^{\mathrm{H}}$ a la propriété de Henkin.*

Démonstration. Supposons que $\overline{\mathcal{T}}^{\mathrm{H}} \vdash \exists x, F$ où $F \in \overline{\Sigma}^{\mathrm{H}}$. Remarquons que F ne peut contenir qu'un nombre fini de symboles de fonctions et de relations : on en déduit qu'il existe $n \in \mathbb{N}$ tel que $F \in \mathrm{Form}(\Sigma_n)$. Cela signifie donc que $(\exists x, F) \to F[c_F/x] \in \mathcal{T}_{n+1}$. On en déduit

$$\dfrac{\dfrac{}{\overline{\mathcal{T}}^{\mathrm{H}} \vdash (\exists x, F) \to F[c_F/x]}\,\mathrm{Ax} \qquad \overline{\mathcal{T}}^{\mathrm{H}} \vdash \exists x, F}{\overline{\mathcal{T}}^{\mathrm{H}} \vdash F[c_F/x]}\,\to_{\mathrm{e}}$$

□

De plus, cette construction est stable par extension (ce qui nous servira lorsque nous complèterons notre théorie).

Propriété 3.3.2.7. *Si $\overline{\mathcal{T}}^{\mathrm{H}} \subseteq \mathcal{S}$ pour une certaine théorie $\mathcal{S}$, alors $\mathcal{S}$ a la propriété de Henkin.*

Démonstration. Comme $(\exists x, F) \to F[c_F/x]$ appartient à $\mathcal{S}$, l'arbre de preuve précédent fonctionne encore en remplaçant $\overline{\mathcal{T}}^{\mathrm{H}}$ par $\mathcal{S}$. □

Enfin, la clôture par témoins conserve la cohérence.

Propriété 3.3.2.8. *Si $\mathcal{T} \nvdash \bot$, alors $\overline{\mathcal{T}}^{\mathrm{H}} \nvdash \bot$.*

Démonstration. On suppose que $\nvdash \bot$, on montre alors par récurrence que $\mathcal{T}_n \nvdash \bot$:

- Comme $\mathcal{T}_0 = \mathcal{T}$, le résultat est direct.
- Supposons que $\mathcal{T}_n \nvdash \bot$, alors A FAIRE

Ainsi, $\mathcal{T}_n \nvdash \bot$ pour tout $n \in \mathbb{N}$. Ceci suffit à notre preuve, car si $\overline{\mathcal{T}}^{\mathrm{H}} \vdash \bot$, alors il existe une liste (finie) $\Gamma \in \mathrm{List}(\overline{\mathcal{T}}^{\mathrm{H}})$ telle que $\Gamma \vdash \bot$, mais $\Gamma \in \mathrm{List}(\mathcal{T}_n)$ pour un certain n, puisque cette liste est finie. Comme $\mathcal{T}_n \nvdash \bot$, cela est absurde. Ainsi $\overline{\mathcal{T}}^{\mathrm{H}} \nvdash \bot$. □

En combinant les deux dernières propriété avec le théorème 3.2.3.7 nous obtenons une extension de $\mathcal{T}$ complète et vérifiant la propriété de Henkin.

Lemme 3.3.2.9. *Si $\mathcal{T}$ est une théorie sur Σ telle que $\mathcal{T} \nvdash \bot$, alors il existe $\Sigma \subseteq \overline{\Sigma}^{\mathrm{H}}$ et une théorie $\mathcal{T} \subseteq \mathcal{T}'$ complète et possédant la propriété de Henkin.*

Démonstration. On applique le théorème 3.2.3.7 à $\overline{\mathcal{T}}^{\mathrm{H}}$: comme on sait que $\overline{\mathcal{T}}^{\mathrm{H}} \nvdash \bot$, on peut effectivement compléter cette théorie. Avec la propriété 3.3.2.7, on sait que cette complétion vérifie aussi la propriété de Henkin. □

On fixe maintenant la théorie $\mathcal{T}'$ construite à partir de $\mathcal{T}$. On fixe aussi la signature Σ' comme étant $\overline{\Sigma}^{\mathrm{H}}$.

Définition 3.3.2.10 (Modèle syntaxique). On définit le modèle syntaxique $\mathcal{M}_{\mathcal{T}'}$ par :

- $|\mathcal{M}_{\mathcal{T}'}|$ défini comme $\mathrm{Term}(\Sigma')/\equiv$ où $\equiv$ est défini par

$$t \equiv u \triangleq \mathcal{T}' \vdash t = u$$

- pour chaque symbole de fonction $f \in \Sigma'$ d'arité n, on associe la fonction

$$\begin{array}{rccc} f^{\mathcal{T}'} & : & |\mathcal{M}_{\mathcal{T}'}|^n & \longrightarrow & |\mathcal{M}_{\mathcal{T}'}| \\ & & (\overline{t_1}, \ldots, \overline{t_n}) & \longmapsto & \overline{f(t_1, \ldots, t_n)} \end{array}$$

- pour chaque symbole de relation $r \in \Sigma'$ d'arité n, on définit la relation $r^{\mathcal{T}'}$ par

$$r^{\mathcal{T}'}(\overline{t_1}, \ldots, \overline{t_n}) \triangleq \mathcal{T}' \vdash r(t_1, \ldots, t_n)$$

Démonstration. Pour que cette définition ait du sens, il convient de montrer que $\equiv$ est une relation d'équivalence, et qu'elle est compatible avec les symboles de fonction et d'équivalence (c'est-à-dire que nos définitions de $f^{\mathcal{T}'}$ et $r^{\mathcal{T}'}$ ne dépendent pas du représentant choisi).

Montrons d'abord que $\equiv$ est une relation d'équivalence :

- Grâce à $=_{\mathrm{i}}$, on sait que $\mathcal{T}' \vdash t = t$ pour tout $t \in \mathrm{Term}(\Sigma')$, donc $\equiv$ est réflexive.
- Supposons que $t \equiv u$, c'est-à-dire que $\mathcal{T}' \vdash t = u$, on peut alors construire l'arbre de preuve suivant :

$$\dfrac{\dfrac{}{\mathcal{T}' \vdash t = t}\,{=_\mathrm{i}} \qquad \mathcal{T}' \vdash t = u}{\mathcal{T}' \vdash u = t}\,{=_\mathrm{e}}$$

donc $u \equiv t$ ($t = t$ peut se lire comme $(x = t)[t/x]$).

- Supposons que $t \equiv u$ et $u \equiv v$, montrons alors que $t \equiv v$:

$$\dfrac{\mathcal{T}' \vdash u = v \qquad \dfrac{\dfrac{}{\mathcal{T}' \vdash t = t}\,{=_\mathrm{i}} \qquad \mathcal{T}' \vdash t = u}{\mathcal{T}' \vdash u = t}\,{=_\mathrm{e}}}{\mathcal{T}' \vdash t = v}\,{=_\mathrm{e}}$$

Ainsi $\equiv$ est bien une relation d'équivalence.

Soit f un symbole de fonction d'arité n. Pour simplifier la preuve, on suppose que f est d'arité 1 (il suffit ensuite de faire une récurrence sur n pour généraliser la preuve que nous allons faire). Pour que $\equiv$ soit compatible avec f, il faut que l'image de f ne dépende pas du choix du représentant, c'est-à-dire que si $t \equiv u$ alors $f^{\mathcal{T}'}(\bar{t}) = f^{\mathcal{T}'}(\overline{u})$, c'est-à-dire que $\mathcal{T}' \vdash f(t) = f(u)$, ce que l'on peut prouver par

$$\dfrac{\mathcal{T}' \vdash t = u \qquad \dfrac{}{\mathcal{T}' \vdash f(t) = f(t)}\,{=_\mathrm{i}}}{\mathcal{T}' \vdash f(t) = f(u)}\,{=_\mathrm{e}}$$

donc $f^{\mathcal{T}'}$ est bien définie.

De même, pour une relation r prise pour simplifier d'arité 1, il convient de montrer que si $t \equiv u$ alors $r^{\mathcal{T}'}(\bar{t}) \to r^{\mathcal{T}'}(\overline{u})$ (il faudrait une équivalence, mais il suffit en fait de montrer l'implication puisque $\equiv$ a été montrée symétrique). L'arbre de preuve suivant nous le montre :

$$\dfrac{\mathcal{T}' \vdash t = u \qquad \mathcal{T}' \vdash r(t)}{\mathcal{T}' \vdash r(u)}\,{=_\mathrm{e}}$$

donc $r^{\mathcal{T}'}$ est bien définie. □

On peut alors démontrer le lemme qui permettra de prouver le théorème de complétude.

Lemme 3.3.2.11. *Pour toute proposition $\varphi \in (\Sigma')$, on a l'équivalence suivante :*

$$\mathcal{T}' \vdash \varphi \iff \mathcal{M}_{\mathcal{T}'} \models \varphi$$

Démonstration. On démontre ce résultat par induction sur la structure de φ :

- si φ est $\top$ ou $\bot$, le résultat est évident (en particulier $\mathcal{T}' \nvdash \bot$).
- si $\varphi = r(t_1, \ldots, t_n)$ alors $\mathcal{T}' \vdash r(t_1, \ldots, t_n)$ signifie exactement que $r^{\mathcal{T}'}(\overline{t_1}, \ldots, \overline{t_n})$, donc $\mathcal{M}_{\mathcal{T}'} \models r(t_1, \ldots, t_n)$. Ce fait est une équivalence puisqu'il provient de la définition même de $r^{\mathcal{T}'}$.
- si $\varphi = \neg\psi$, montrons que $\mathcal{T}' \vdash \varphi \implies \mathcal{M}_{\mathcal{T}'} \models \varphi$. Comme $\mathcal{T}' \vdash \neg\psi$, on en déduit que $\mathcal{T}' \nvdash \psi$ (car $\mathcal{T}' \nvdash \bot$), mais puisque $\mathcal{T}'$ est une théorie complète, on en déduit que A FAIRE

□

On peut maintenant prouver le théorème de complétude.

Théorème 3.3.2.12 (Complétude de la déduction naturelle). *Soit une signature Σ et une théorie $\mathcal{T}$ sur Σ telle que $\mathcal{T} \nvdash \bot$. Alors il existe un modèle $\mathcal{M} \models \mathcal{T}$.*

Démonstration. En reprenant notre modèle $\mathcal{M}_{\mathcal{T}'}$, on sait à partir du lemme précédent que $\mathcal{M}_{\mathcal{T}'} \models \mathcal{T}'$ grâce à la règle d'axiome. Comme $\mathcal{T} \subseteq \mathcal{T}'$, on en déduit donc que $\mathcal{M}_{\mathcal{T}'} \models \mathcal{T}$. □

Théorème 3.3.2.13 (Complétude, deuxième version). *Soit Σ une signature. Alors les deux relations $\vdash, \models \subseteq \mathcal{P}(\mathrm{Form}(\Sigma)) \times \mathrm{Form}(\Sigma)$ coïncident.*

On récupère en tant que conséquence un théorème essentiel de la théorie des modèles : le théorème de compacité.

Théorème 3.3.2.14 (Théorème de compacité). *Soit Σ une signature, $\mathcal{F} \subseteq \mathrm{Form}(\Sigma)$ et $F \in \mathrm{Form}(\Sigma)$, alors $\mathcal{F} \models F$ si et seulement s'il existe $\mathcal{A} \subseteq_{\mathrm{fin}} \mathcal{F}$ tel que $\mathcal{A} \models F$.*

Démonstration. En effet, si $\mathcal{F} \models F$ alors $\mathcal{F} \vdash F$, d'où par définition l'existence de $\Gamma \in \mathrm{List}(\mathcal{F})$ tel que $\Gamma \vdash A$. Il n'y a qu'un nombre fini de propositions dans Γ, donc on peut trouver $\mathcal{A} \subseteq_{\mathrm{fin}} \mathcal{F}$ tel que $\Gamma \in \mathrm{List}(\mathcal{A})$: on en déduit donc que $\mathcal{A} \vdash F$, d'où par correction que $\mathcal{A} \models F$. □

Chapitre 4

Théorie des ensembles ordonnés

Table des sous-matières

Dans ce chapitre, nous aborderons la théorie des ensembles ordonnés. Celle-ci est fortement liée à la logique, et permet entre autre de formaliser des structures logiques comme les algèbres de Boole ou les algèbres de Heyting.

Les ensembles ordonnés peuvent être vus comme des cas dégénérés de catégories. Si nous ne donnons pas dans cet ouvrage de rudiments de théorie des catégories, l'étudiant catégoricien pourra voir dans plusieurs définitions des éléments similaires à ce qu'il connait en théorie des catégories.

Nous commencerons par définir les idées générales sur les ensembles ordonnés. Cette première partie introduira les notions élémentaires d'ordre, de borne supérieure, inférieure, ainsi que les ordres bien fondés se reposant sur les éléments minimaux. Nous aborderons ensuite la théorie des treillis, en allant des demi-treillis aux algèbres de Boole et de Heyting. Cela motivera notre étude des filtres et ultrafiltres, dont le comportement dans une algèbre de Boole est fortement simplifiée, et nous aurons alors l'occasion de démontrer le théorème d'extension d'une théorie cohérente en une théorie complète.

4.1 Ensembles ordonnés

Commençons par donner les définitions les plus élémentaires.

4.1.1 Définitions

Définition 4.1.1.1 (Ensemble ordonné). Un ensemble ordonné est un couple $(X, \leq)$ où $\leq$ est une relation binaire sur X vérifiant :

- $\leq$ est reflexive : pour tout $x \in X$, $x \leq x$.
- $\leq$ est antisymétrique : si deux éléments $x, y \in X$ sont tels que $x \leq y$ et $y \leq x$ alors $x = y$.
- $\leq$ est transitive : si trois éléments $x, y, z \in X$ sont tels que $x \leq y$ et $y \leq z$ alors $x \leq z$.

Un couple $(X, \leq)$ tel que $\leq$ est seulement réflexive et transitive est appelé un ensemble pré-ordonné, et $\leq$ est appelé un pré-ordre sur X.

Exemple (Ensemble des parties). Soit X un ensemble, l'ensemble $(\mathcal{P}(X), \subseteq)$ est un ensemble ordonné.

La plupart des résultats que nous énoncerons pourront aussi bien se traduire sur des pré-ordres, en remplaçant en général $x = y$ par $x \leq y \wedge y \leq x$. En fait, un pré-ordre peut se ramener à un ordre par un quotient.

Propriété 4.1.1.2. *Soit $(X, \leq)$ un ensemble pré-ordonné, alors la relation $\sim \subseteq X \times X$ définie par*

$$x \sim y \triangleq x \leq y \wedge y \leq x$$

est une relation d'équivalence, et $\leq$ induit un ordre sur $X/\sim$.

Démonstration. Vérifions que $\sim$ est une relation d'équivalence :

- $\sim$ est transitive : $x \leq x$ et $x \leq x$, donc $x \sim x$.
- $\sim$ est symétrique : supposons que $x \sim y$, alors $x \leq y$ et $y \leq x$, donc $y \leq x$ et $x \leq y$, donc $y \sim x$.
- $\sim$ est réflexive : supposons que $x \sim y$ et $y \sim z$, alors $x \leq y$ et $y \leq z$, donc $x \leq z$, et de même $z \leq y$ et $y \leq x$ donc $z \leq z$, ainsi $x \sim z$.

Pour vérifier que $\leq$ induit un ordre sur $X/\sim$, il suffit de vérifier que si $x \sim x'$ et $y \sim y'$, alors $x \leq y$ si et seulement si $y \leq x$.

Si $x \leq y$ alors comme $x' \leq x$ et $y \leq y'$, on en déduit que $x' \leq y'$ et récuproquement, comme $x \leq x'$ et $y' \leq y$, il vient que $x' \sim y' \implies x \sim y$. □

Exercice 4.1.1.3. Soit une signature Σ, montrer que la relation $\vdash \subseteq \mathrm{Form}(\Sigma) \times \mathrm{Form}(\Sigma)$ définie comme restriction de la relation $\vdash$ que nous avons vue précédemment, mais où la liste de gauche ne contient qu'une proposition, est un pré-ordre.

Remarque 4.1.1.4. La relation $\dashv\vdash$ s'écrit aussi $\equiv$, ce qui est cohérent avec notre définition sémantique de $\equiv$, étant donné que l'on sait que $\vdash$ et $\vDash$ coïncident.

L'ensemble ordonné construit par $\mathrm{Form}(\Sigma)/\dashv\vdash$ est appelé l'algèbre de Lindenbaum-Tarski. Nous la verrons plus en détail plus tard.

Une autre définition possible d'un ensemble ordonné est ce que l'on appelle habituellement un ensemble strictement ordonné. Nous en donnons la définition et montrons que l'on peut faire correspondre à chaque ordre un ordre strict (et réciproquement).

Définition 4.1.1.5 (Ordre strict). Un ensemble strictement ordonné est une paire $(X, <)$ où $<$ est une relation binaire sur X vérifiant :

- $<$ est antiréflexive : pour tout $x \in X, x \not< x$.
- $<$ est transitive.

Proposition 4.1.1.6. *On a une bijection entre les ensembles ordonnés et les ensembles strictement ordonnés en associant à $(X, \leq)$ l'ensemble strictement ordonné $(X, <)$ défini par*

$$x < y \triangleq x \leq y \wedge x \neq y$$

Démonstration. On définit, dans l'autre sens, la relation $\leq$ à partir de $<$ par

$$x \leq y \triangleq x < y \vee x = y$$

Le fait que $((x \leq y) \wedge x \neq y) \vee x = y$ est équivalent à $x \leq y$ et que $(x < y \vee x = y) \wedge x \neq y$ est équivalent à $x < y$ se vérifie directement.

Montrons que $<$ défini à partir de $\leq$ est un ordre strict :

- par définition, $x < y \implies x \neq y$, donc $x \not< x$.
- si $x < y$ et $y < z$, alors $x \leq z$ par transitivité de $\leq$. Si $x = z$ alors $z < y$ par substitution de x par z, donc par transitivité $y < y$, ce qui est absurde. Donc $x \neq z$.

Montrons que $\leq$ défini à partir de $<$ est un ordre :

- par définition, si $x = y$ alors $x \leq y$.
- si $x \leq y$ et $y \leq x$, alors soit $x = y$, soit $x < y$. Dans le premier cas, on a le résultat voulu. Dans le deuxième cas, on utilise le fait que $y \leq x$ pour en déduire que soit $y = x$, soit $y < x$. Encore une fois, le cas $y = x$ est directement ce qu'il faut démontrer. Si $y < x$, alors par transitivité avec $x < y$ on en déduit que $x < x$, ce qui est absurde.
- si $x \leq y$ et $y \leq z$, alors soit $x = y$, auquel cas on voit directement que $x \leq z$, soit $x < y$. En utilisant le fait que $y \leq z$, on en déduit que soit $y = z$, soit $y < z$. Dans le premier cas, il est évident que $x \leq z$. Dans le deuxième cas, par transitivité, on en déduit que $x < z$.

□

On définira donc des propriétés indifféremment sur un ordre ou sur un ordre strict (si on dit qu'une propriété P définie sur les ordres strictes est vérifiée pour un ordre $\leq$, cela signifie que $<$ vérifie P).

Intéressons-nous d'abord aux éléments : comme un ordre est (sauf indication du contraire) partiel, deux éléments peuvent ne pas être mis en relation. Dans ces cas-là, deux notions distinctes apparaissent : la comparabilité et la compatibilité. Pour deux éléments, être comparable est une condition forte, mais parfois justement trop forte : la compatibilité signifie juste que deux éléments peuvent être rejoints.

Définition 4.1.1.7 (Comparabilité, compatibilité). Soit $(X, \leq)$ un ensemble ordonné. On dit que deux éléments x et y sont comparables si $x \leq y$ ou $y \leq x$. Deux éléments x et y sont dits compatibles s'il existe $z \in X$ tel que $z \leq x$ et $z \leq x$. Deux éléments sont dits incompatibles s'ils ne sont pas compatibles.

Si tous les éléments de X sont compatibles, on dit que X est un ensemble totalement ordonné. Si tous les éléments de X sont compatibles, on dit que X est un ensemble ordonné filtrant vers le bas.

Exemple. Deux éléments de $\mathcal{P}(X)$ sont toujours compatibles en prenant comme élément z leur intersection. Pourtant, $\{x\}$ et $\{y\}$ pour $x \neq y$ ne sont pas comparables.

4.1.2 Dualité

Relevons dès maintenant un phénomène important en théorie des ordres, appelé la dualité : lorsque l'on a un ordre $\leq$, on peut définir la relation $\geq$ par $y \geq x \iff x \leq y$, et cette relation est aussi une relation d'ordre, inversant quel élément est le plus grand et quel élément est le plus petit. Ainsi lorsque l'on définit une notion sur un ensemble ordonné $(X, \leq)$, une notion duale est directement définie en considérant l'objet dans $(X, \geq)$.

Ce phénomène permet de simplifier beaucoup de preuves : si l'on veut prouver deux propositions P et Q où Q est obtenue en remplaçant $\leq$ dans P par $\geq$ (et toutes les définitions par leurs définitions duales), si les propriétés portent sur tous les ensembles ordonnés, alors prouver P suffit à déduire Q, qui est une conséquence de P pour l'ordre dual.

Nous allons donc souvent écrire deux propositions pour une seule preuve, car la proposition duale est souvent aussi importante que la proposition de base.

4.1.3 Bornes et majorations

On peut voir la théorie des ordres sous deux primes : le prisme algébrique et le prisme analytique.

Suivant le prisme algébrique, nous étudions des structures munies d'une relation (on peut voir cela comme de la théorie des modèles), et l'accent est alors mis sur les propositions du premier ordre. Par exemple, la comparabilité ou la compatibilité sont des notions algébriques. De même on peut parler de la borne supérieure de deux éléments, ou d'être le maximum entre x et y, voire d'un ensemble fini d'éléments.

Suivant le prisme analytique, nous nous intéressons aux parties de l'ensemble ordonné (potentiellement infinie). Nous parlons donc par exemple de la borne supérieure d'un ensemble quelconque. Les structures pour lesquelles on peut appliquer le point de vue analytique sont donc moins nombreuses, mais nous allons étudier les définitions basiques de ces principes de bornes supérieures, de majorant ou autre.

Définition 4.1.3.1 (Majorant, minorant). Soit $(X, \leq)$ un ensemble ordonné et $F \subseteq X$ une partie de X. On dit que x est un majorant de F si

$$\forall y \in F, y \leq x$$

et que x est un minorant de F si

$$\forall y \in F, x \leq y$$

Par abus de notation, on écrira $F \leq x$ (respectivmeent $x \leq F$) pour dire que x est un majorant (respectivement un minorant) de F.

Définition 4.1.3.2 (Borne supérieure, borne inférieure). Soit $(X, \leq)$ un ensemble ordonné et $F \subseteq X$ une partie de X. On dit que x est la borne supérieure de F si

$$F \leq x \wedge (\forall y, F \leq y \implies x \leq y)$$

et que x est la borne inférieure de F si

$$x \leq F \wedge (\forall y, x \leq F \implies y \leq x)$$

Démonstration. Puisque nous avons dit « la » il est nécessaire de montrer l'unicité de la borne supérieure (l'unicité de la borne inférieure se prouver par dualité).

Supposons que x, y sont des bornes supérieures de F. Comme $F \leq x$ et y est une borne supérieure, $y \leq x$. De même, comme $F \leq y$ et x est une borne supérieure, $x \leq y$. Donc $x = y$ par antisymétrie. □

Notation 4.1.3.3. Si la borne supérieure de F existe, on la notera $\bigvee F$. Si la borne inférieure de F existe, on la notera $\bigwedge F$.

Exercice 4.1.3.4. Soit $(X, \leq)$ un ensemble ordonné et $A \subseteq \mathcal{P}X$ une partie telle que $\bigcup A$ admet une borne supérieure et tout élément de A admet une borne supérieure, montrer alors que

$$\bigvee(\bigcup A) = \bigvee\left(\left\{\bigvee a \;\middle|\; a \in A\right\}\right)$$

Définition 4.1.3.5 (Maximum,minimum). Soit $(X, \leq)$ un ensemble ordonné et $F \subseteq X$. On dit que x est le maximum de F si $F \leq x$ et $x \in F$. On dit que x est le minimum de F si $x \leq F$ et $x \in F$.

Notation 4.1.3.6. On notera $\max F$ et $\min F$ le maximum (respectivement le minimum) de F.

Propriété 4.1.3.7. *S'il existe, le maximum (respectivement le minimum) d'une partie est sa borne supérieure (respectivement inférieure).*

Démonstration. Si x est un majorant de F, alors comme $\max F \in F$, on en déduit que $\max F \leq x$, d'où le résultat. □

Définition 4.1.3.8 (Élément maximal, minimal). Soit un ensemble ordonné $(X, \leq)$ et une partie $F \subseteq X$. On dit que $x \in F$ est un élément maximal (respectivement minimal) dans F s'il n'existe pas $y \in F$ tel que $x \leq y$ (respectivement tel que $y \leq x$).

Exercice 4.1.3.9. Montrer que le maximum d'une partie en est un élément maximal de cette partie.

Une autre partie intéressant est l'étude de parties sur lesquelles l'ordre induit se comporte particulièrement bien. C'est ce que nous allons voir avec les chaînes, les antichaînes et les parties filtrantes.

Définition 4.1.3.10 (Chaîne, antichaîne). Soit $(X, \leq)$ un ensemble ordonné. On dit que $C \subseteq X$ est une chaîne de X si l'ensemble $(C, \leq)$ (avec l'ordre induit) est un ensemble totalement ordonné. On dit que $A \subseteq X$ est une antichaîne de X si l'ensemble $(A, \leq)$ (avec l'ordre induit) ne possède pas deux éléments comparables, c'est-à-dire si $(A, \leq)$ est l'ordre discret. On dit que $A \subseteq X$ est une antichaîne forte de X si pour tous $x, y \in A$, x et y ne sont pas compatibles.

Cela nous mène à la notion d'ensemble inductif, importante pour établir le lemme de Zorn, qui est une version de l'axiome du choix plus maniable. Nous prouverons ce résultat lorsque nous étudierons la théorie des ensembles, mais pouvons déjà en donner le résultat.

Définition 4.1.3.11 (Ensemble inductif). Un ensemble ordonné $(X, \leq)$ est dit inductif si toute chaîne admet un majorant.

Théorème 4.1.3.12 (Lemme de Zorn). *Si $(X, \leq)$ est un ensemble inductif et $x \in X$, alors il existe un élément maximal $y \in X$ qui est supérieur à x.*

4.1.4 Ordre bien fondé et bon ordre

Nous donnons maintenant les notions d'ordre bien fondé et de bon ordre. Les bons ordres auront de l'importance pour la théorie des ensembles, et les ordres bien fondés sont utiles pour étendre le théorème d'induction.

Définition 4.1.4.1 (Ordre bien fondé). Soit $(X, \leq)$ un ensemble ordonné. On dit que $\leq$ est bien fondé si toute partie $F \subseteq X$ non vide possède un élément minimal, c'est-à-dire si

$$\forall F \subseteq X, F \neq \varnothing \implies \exists x \in F, \forall y \in F, y \leq x \implies x = y$$

Remarque 4.1.4.2. Une définition équivalente, sur les ordres stricts, est que $<$ est bien fondé si :

$$\forall F \subseteq X, F \neq \varnothing \implies \exists x \in F, \forall y \in F, y \not< x$$

Définition 4.1.4.3 (Bon ordre). Un ensemble ordonné $(X, \leq)$ est un bon ordre lorsque toute partie $F \subseteq X$ possède un minimum.

Une première conséquence évidente est qu'un bon ordre est un ordre bien fondé. De plus, un bon ordre est total, puisque pour $x, y \in X$ il existe un minimum à $\{x, y\}$. Ceci caractérise en fait les bons ordres parmi les ordres bien fondés.

Proposition 4.1.4.4. *Si $(X, \leq)$ est un ordre bien fondé total, alors c'est un bon ordre.*

Démonstration. Supposons que $(X, \leq)$ est un ordre bien fondé total. Soit $F \subseteq X$, comme $\leq$ est un ordre bien fondé, on trouve m un élément minimal de F. Soit $x \in F$, alors soit $x \leq m$ soit $m \leq x$. Dans le premier cas, par minimalité de m, on en déduit que $m = x$, donc que $m \leq x$. Ainsi, dans tous les cas, m est un minorant de F : F admet donc un minimum.

Donc $(X, \leq)$ est un bon ordre. □

Nous avons vu que la notion de structure inductive nous offre le principe d'induction, qui dans le cas particulier de $\mathbb{N}$ correspond au principe de récurrence. Le principe de récurrence forte, lui, est de nature légèrement différente : il se base sur la notion d'ordre. Il dit que si pour tout $k \in \mathbb{N}$, on peut prouver

$$(\forall i < k, P(i)) \implies P(k)$$

alors on peut en déduire que $\forall n, P(n)$.

Ce procédé peut se généraliser à un ensemble bien fondé : si l'on définit l'ensemble $x^{-1} = \{y \in X \mid y < x\}$, on veut montrer que $P(x)$ est vrai pour tout x à partir du fait que $(\forall y \in x^{-1}, P(y)) \implies P(x)$.

L'idée derrière cette démonstration est que pour un ordre bien fondé, on a un moyen d'ordonner les éléments de sorte qu'on peut prouver P élément par élément en s'appuyant sur le fait que P est vérifié aux étapes précédentes.

Théorème 4.1.4.5 (Induction bien fondée). *Soit $(X, \leq)$ un ensemble ordonné bien fondé. Soit $P \subseteq X$ une partie vérifiant*

$$\forall x, (\forall y < x, y \in P) \implies x \in P$$

est égale à P.

Démonstration. Supposons que $P \neq X$. On trouve alors $x \in X \setminus P$ minimal pour $<$. Comme $<$ est minimal dans $X \setminus P$, on en déduit que pour tout $y < x$, $y \notin X \setminus P$, c'est-à-dire $y \in X$. Par hypothèse, puisque $\forall y < x, y \in P$, on en déduit que $x \in P$: c'est absurde. Ainsi, par l'absurde, on a prouvé que $X \setminus P = \varnothing$, donc que $P = X$. □

Remarque 4.1.4.6. En construisant un ensemble inductif à partir d'une signature (C, α) en prenant une suite d'ensembles $(X_i)_{i\in\mathbb{N}}$, on peut définir un ordre canonique bien fondé, l'ordre de sous-terme, par récurrence sur i :

- pour X_0, on a facilement un ordre bien fondé sur $\varnothing$.
- supposons qu'on possède un ordre bien fondé sur X_n, alors on définit notre ordre sur X_{n+1} par $x_i < c(x_1, \ldots, x_n)$ pour tout $i \in \{1, \ldots, n\}$, et en en prenant la clôture transitive.

Cet ordre, en particulier, permet de généraliser l'hypothèse d'induction de notre définition d'ensemble inductif en une induction forte.

Exercice 4.1.4.7. Montrer que l'ordre défini plus tôt est bien un ordre bien fondé.

4.2 Treillis

Nous allons maintenant étudier le cas des treillis, qui sont des ensembles ordonnés munis de propriétés de clôture. Pour être exhaustif et situer exactement quelles conditions mènent à quels résultats, nous allons définir nos structures de la plus faible à la plus forte, en montrant tous les résultats que nous souhaitons avoir sur une structure avant de passer à la suivante.

4.2.1 Demi-treillis

Commençons par étudier les demi-treillis : ceux-ci sont des ensembles ordonnés avec une direction privilégiée, selon laquelle il existe toujours une borne supérieure pour un ensemble fini.

Définition 4.2.1.1 (Demi-treillis). Un ensemble $(X, \leq)$ est un sup demi-treillis si toute partie $F \subseteq_{\text{fin}} X$ admet une borne supérieure $\bigvee F$. C'est un inf demi-treillis si toute partie $F \subseteq_{\text{fin}} X$ admet une borne inférieure $\bigwedge F$.

Propriété 4.2.1.2. *De manière équivalent, un ensemble $(X, \leq)$ est un sup demi-treillis si et seulement s'il admet un élément $\bot$ et une opération $\vee : X \times X \to X$ tels que $\bot$ est un minorant de X et $x \vee y$ est la borne supérieure de x et y.*

C'est un inf demi-treillis si et seulement s'il admet un majorant $\top$ et une opération $\wedge : X \times X \to X$ prenant la borne inférieure de deux éléments.

Démonstration. Si $(X, \leq)$ est un sup demi-treillis, on trouve $\bot$ en prenant $\bigvee \varnothing$ et $x \vee y$ en prenant $\bigvee\{x, y\}$.

Réciproquement, montrons par récurrence sur le cardinal de $F \subseteq_{\text{fin}} X$ qu'il existe une borne supérieure à F :

- si $F = \varnothing$, alors $\bigvee F = \bot$.
- si $F = F' \cup \{x\}$ et $\bigvee F'$ existe, alors $\bigvee F = \bigvee\{\bigvee F', \bigvee\{x\}\} = \bigvee F \vee x$ donc $\bigvee F$ existe.

□

Cela nous donne une caractérisation plus algébrique d'un demi-treillis : c'est un ensemble muni d'opérations. Cependant, ces opérations ont encore des définitions trop proche de la théorie des ordres. Nous allons voir que l'on peut transformer cette définition en une définition purement algébrique.

Définition 4.2.1.3 (Monoïde commutatif idempotent). Un ensemble X muni d'une opération $\cdot$ et d'un élément e est un monoïde idempotent commutatif si les propriétés suivantes sont vérifiées :

- $\forall x \in X, x \cdot e = e \cdot x = x$
- $\forall x, y, z \in X, x \cdot (y \cdot z) = (x \cdot y) \cdot z$
- $\forall x, y \in X, x \cdot y = y \cdot x$
- $\forall x \in X, x \cdot x = x$

On définit sur un monoïde commutatif idempotent $(X, \cdot, e)$ les relations $\leq_\vee$ et $\leq_\wedge$ par

$$x \leq_\vee y \triangleq x \cdot y = y$$
$$x \leq_\wedge y \triangleq x \cdot y = x$$

Proposition 4.2.1.4. *L'ensemble $(X, \bot, \vee)$ est un sup demi-treillis pour $\leq_\vee$ si et seulement si $(X, \bot, \vee)$ est un monoïde commutatif idempotent.*

L'ensemble $(X, \top, \wedge)$ est inf demi-treillis pour $\leq_\wedge$ si et seulement si $(X, \top, \wedge)$ est un monoïde commutatif idempotent.

Démonstration. Montrons d'abord que $\leq_\vee$ est bien un ordre pour un monoïde commutatif idempotent :

- par idempotence, $x \vee x = x$ donc $x \leq_\vee x$.
- si $x \leq_\vee y$ et $y \leq_\vee x$, alors $x \vee y = y$ et $y \vee x = x$ donc

$$\begin{aligned} x &= y \vee x \\ &= x \vee y \qquad \text{par commutativité} \\ &= y \end{aligned}$$

- si $x \leq_\vee y$ et $y \leq_\vee z$, alors $x \vee y = y$ et $y \vee z = z$ donc

$$\begin{aligned} x \vee z &= x \vee (y \vee z) \\ &= (x \vee y) \vee z \\ &= y \vee z \\ &= z \end{aligned}$$

Supposons maintenant que $(X, \bot, \vee)$ est un sup demi-treillis, alors tout d'abord $x \leq y$ si et seulement si $y = \bigvee\{x, y\}$, si et seulement si $y = x \vee y$, d'où le fait que $\leq_\vee$ coïncide avec $\leq$. Montrons maintenant que $(X, \bot, \vee)$ est un monoïde commutatif idempotent :

- $\bot$ est un élément neutre pour $\vee$: pour tout $x \in X$, $\bot \leq x$ donc $\bot \vee x = x$, et de même pour $x \vee \bot$.
- $\vee$ est associatif : $x \vee (y \vee z) = \bigvee\{x, y, z\} = (x \vee y) \vee z$.
- $\vee$ est commutatif : $x \vee y = \bigvee\{x, y\} = y \vee x$.
- $\vee$ est idempotent : $x \vee x = \bigvee\{x\} = x$.

Réciproquement, montrons que so $(X, \bot, \vee)$ est un monoïde commutatif idempotent, alors $\bot$ est un minorant de X pour $\leq_\vee$ et $\vee$ est la borne supérieure de deux éléments pour $\leq_\vee$:

- pour tout x, $\bot \vee x = x$ donc $\bot \leq_\vee x$.
- pour tous x, y, $x \vee (x \vee y) = (x \vee x) \vee y = x \vee y$ donc $x \vee y$ est un majorant de x (de même on prouve que $x \vee y$ est un majorant de y). Supposons que z est un majorant de $\{x, y\}$ pour $\leq_\vee$. Cela signifie que $x \vee z = z$ et $x \vee y = z$. Alors $(x \vee y) \vee z = x \vee (y \vee z) = x \vee z = z$ donc $x \vee y \leq_\vee z$, donc $x \vee y$ est inférieur à tout majorant de $\{x, y\}$: c'est la borne supérieure de $\{x, y\}$.

□

Ainsi on parlera de sup demi-treillis (respectivement inf demi-treillis) en utilisant $\leq$ et les opérations $\bot, \vee$ (respectivement $\top, \wedge$) sans souci, puisque toutes les définitions nous donnent l'ensemble de ces notions.

Exercice 4.2.1.5. Soit Σ une signature, on rappelle que l'algèbre de Lindenbaum-Tarski sur Σ, notée $\mathcal{L}(\Sigma)$, est l'ensemble quotient $\mathrm{Form}(\Sigma)/\dashv\vdash$ muni de l'ordre $\vdash$. Montrer que cette structure est un inf demi-treillis et un sup demi-treillis où le majorant est $\top$, le minorant $\bot$, la borne supérieure $\vee$ et la borne inférieure $\wedge$.

Exercice 4.2.1.6. Soit un ensemble X, montrer que $(\mathcal{P}(X), \subseteq)$ est à la fois un sup et un inf demi-treillis, où X, $\varnothing$, $\cup$ et $\cap$ sont respectivement le majorant, le minorant, la borne supérieure et la borne inférieure.

On peut aussi définir les morphismes de demi-treillis, qui sont des morphismes entre structures au sens attendu.

Définition 4.2.1.7 (Morphisme de demi-treillis). Soient $(X, \leq)$ et $(Y, \leq)$ deux sup demi-treillis (respectivement deux inf demi-treillis). Un morphisme de sup demi-treillis (respectivement d'inf demi-treillis) est une fonction $f : X \to Y$ croissante telle que $f(\bot_X) = \bot_Y$ (respectivement $f(\top_X) = \top_Y$) et telle que pour tous $x, y \in X$, $f(x \vee y) = f(x) \vee f(y)$ (respectivement telle que pour tous $x, y \in X$, $f(x \wedge y) = f(x) \wedge f(y)$).

4.2.2 Treillis

Un demi-treillis ne concerne qu'une direction entre le sup et l'inf. Un treillis, comme son nom l'indique, est une structure qui possède les propriétés des deux demi-treillis à la fois.

Définition 4.2.2.1 (Treillis). Un ensemble ordonné $(X, \leq)$ est un treillis si toute partie finie $F \subseteq_{\text{fin}} X$ possède à la fois une borne supérieure $\bigvee F$ et une borne inférieure $\bigwedge F$.

En théorie des ordres, il est direct de voir qu'être un treillis est équivalent à être à la fois un inf demi-treillis et un sup demi-treillis, mais cela n'est pas aussi évident algébriquement, car un inf demi-treillis est défini algébriquement par la relation $\leq_\wedge$ et un sup demi-treillis par la relation $\leq_\vee$: il faut vérifier que ces deux relations coïncident.

Proposition 4.2.2.2. *Un bimonoïde $(X, \bot, \vee, \top, \wedge)$ est un treillis pour $\leq_\vee$ (respectivement pour $\vee_\wedge$) si et seulement si $(X, \bot, \vee)$ est un inf demi-treillis, $(X, \top, \wedge)$ est un inf demi-treillis, $\leq_\vee$ et $\leq_\wedge$ coïncident, cette dernière condition étant équivalente aux deux identités d'absorption suivantes :*

- $\forall x, y \in X, (x \vee y) \wedge x = x$

- $\forall x, y \in X, (x \wedge y) \vee x = x$

Démonstration. Soit $(X, \bot, \vee, \top, \wedge)$ un bimonoïde commutatif idempotent. On suppose de plus que $\leq_\vee = \leq_\wedge$, montrons les deux identités :

- Soient $x, y \in X$, comme $x \leq_\vee x \vee y$, on en déduit que $x \leq_\wedge x \vee y$, c'est-à-dire que $x \wedge (x \vee y) = x$ d'où l'égalité par commutativité de $\wedge$.
- Soient $x, y \in X$, comme $x \wedge y \leq_\wedge x$, on en déduit que $x \wedge y \leq_\vee x$, c'est-à-dire que $(x \wedge y) \vee x = x$.

Réciproquement, supposons vraies les identités. Montrons par double inclusion que $\leq_\vee = \leq_\wedge$:

- soient x, y tels que $x \leq_\vee y$, alors $x \vee y = y$, montrons que $x \wedge y = x$:

$$\begin{aligned} x \wedge y &= x \wedge (x \vee y) \\ &= x \end{aligned}$$

  par la première identité. Ainsi $x \leq_\wedge y$, donc $\leq_\vee \subseteq \leq_\wedge$.
- soient x, y tels que $x \leq_\wedge y$, alors $x \wedge y = x$, montrons que $x \vee y = y$:

$$\begin{aligned} x \vee y &= (x \wedge y) \vee y \\ &= y \end{aligned}$$

  par la deuxième identité. Ainsi $x \leq_\vee y$, donc $\leq_\wedge \subseteq \leq_\vee$.

Donc $\leq_\vee = \leq_\wedge$. □

Cela nous donne donc une définition purement algébrique d'un treillis (en particulier, un treillis est un modèle d'une certaine théorie).

Définition 4.2.2.3 (Morphisme de treillis). Soient $(X, \leq)$ et $(Y, \leq)$ deux treillis, un morphisme de treillis entre ces deux treillis est une fonction $f : X \to Y$ qui est à la fois un morphisme de sup demi-treillis et un morphisme d'inf demi-treillis.

Exercice 4.2.2.4. Soit une signature Σ, montrer que $\mathcal{L}(\Sigma)$ est un treillis.

Exercice 4.2.2.5. Soit un ensemble X, montrer que $\mathcal{P}(X)$ est un treillis.

Exercice 4.2.2.6. Soit une signature Σ. Soit $\mathcal{M}$ une structure sur Σ. Montrer que la fonction

$$\begin{array}{rccc} \Phi & : & \mathcal{L}(\Sigma_{ZF}) & \longrightarrow & \mathcal{P}(\mathcal{M}) \\ & & \varphi(x) & \longmapsto & \{x \in |\mathcal{M}| : \mathcal{M} \models \varphi(x)\} \end{array}$$

est un morphisme de treillis.

Cependant, la partie sup et la partie inf d'un treillis n'interagissent pas suffisamment bien en général, c'est pourquoi il est nécessaire d'ajouter la distributivité de $\vee$ sur $\wedge$ et de $\wedge$ sur $\vee$, si l'on veut pouvoir l'utiliser.

Définition 4.2.2.7 (Treillis distributif). Soit $(X, \leq)$ un treillis. On dit que c'est un treillis distributif si de plus l'une des deux propositions équivalentes est vraie :

(i) $\forall x, y, z \in X, x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$

(ii) $\forall x, y, z \in X, x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$

Démonstration. Vérifions que ces deux propositions sont équivalentes. Par dualité, il nous suffit de prouver que $(i) \implies (ii)$. On suppose donc (i). Soient $x, y, z \in X$, par un calcul :

A FAIRE

D'où l'équivalence. □

Exercice 4.2.2.8. Montrer que $\mathcal{L}(\Sigma)$ et $\mathcal{P}(X)$ sont des treillis distributifs.

4.2.3 Algèbre de Boole

Nous voyons maintenant notre premier cas de structure logique décrite par un ensemble ordonné. Avec la structure de treillis, nous avons naturellement une structure permettant d'encoder la conjonction et la disjonction. Pour travailler sur la logique propositionnelle, il ne nous manque que la négation, l'implication pouvant se coder grâce à l'équivalence logique $a \to b \equiv \neg a \vee b$.

Nous souhaitons donc avoir, en plus de ce que nous avons déjà défini, une fonction $\neg : X \to X$ permettant de jouer le rôle de la négation. En logique classique, qui est notre cadre de travail pour l'instant, deux propriétés en particulier décrivent le comportement de $\neg$: le tiers exclu et le principe de non contradiction, disant que $x \vee \neg x$ est vraie et que $x \wedge \neg x$ est fausse. La notion de « vrai » et de « faux » sont assez naturellement encodées par $\top$ et $\bot$, respectivement, d'où la définition suivante.

Définition 4.2.3.1 (Complément). Soit un treillis $(X, \leq)$. On dit que y est un complément de x si

$$x \vee y = \top \qquad x \wedge y = \bot$$

En fait, ceci suffit à caractériser la négation. En particulier, il n'y a qu'un complément par élément, pour un treillis distributif.

Proposition 4.2.3.2. *Soit $(X, \leq)$ un treillis distributif et $x \in X$. Alors s'il existe un complément à x, celui-ci est unique.*

Démonstration. Supposons que y, z soient deux compléments de x. Alors

$$\begin{aligned} y \vee z &= y \vee z \vee (z \wedge x) \\ &= z \vee ((y \vee z) \wedge (y \vee x)) \\ &= z \vee ((y \vee z) \wedge \top) \\ &= z \vee (y \vee z) \\ &= z \end{aligned}$$

d'où $y \leq z$. En inversant le rôle de y et z, on trouve aussi que $y \leq z$, d'où $y = z$. □

Une algèbre de Boole peut ainsi se décrire grâce à cette notion de complément.

Définition 4.2.3.3 (Algèbre de Boole). Une algèbre de Boole $(B, \leq)$ est un treillis distributif complémenté. Pour un élément $x \in B$, on notera $\neg x$ sont complément.

Exercice 4.2.3.4. Soit une signature Σ et un ensemble X. Montrer que $\mathcal{L}(\Sigma)$ et $\mathcal{P}(X)$ sont des algèbres de Boole.

Remarque 4.2.3.5. En particulier, pour $X = \{\varnothing\}$, on obtient une algèbre de Boole de cardinal 2, qui correspond en fait au corps $\mathbb{Z}/2\mathbb{Z}$. On notera cette algèbre de Boole B_2, celle-ci a la particularité d'être la plus petite algèbre de Boole non dégénérée (telle que $\bot \neq \top$) et d'être la seule algèbre de Boole qui a une structure de corps.

Exercice 4.2.3.6. Montrer que les identités de De Morgan sont vérifiées :

$$\neg\neg x = x \qquad \neg(x \vee y) = \neg x \wedge \neg y \qquad \neg(x \wedge y) = \neg x \vee \neg y$$

Exercice 4.2.3.7. Montrer que si B et B' sont deux algèbres de Boole, alors un morphisme f de treillis entre B et B' est en particulier un morphisme d'algèbre de Boole, au sens où la propriété suivante est vérifiée :

$$\forall x \in B, f(\neg x) = \neg f(x)$$

L'exercice suivant donne une caractérisation purement algébrique d'une algèbre de Boole : on peut le voir comme un anneau idempotent.

Exercice 4.2.3.8. Soit $(B, \leq)$ une algèbre de Boole. On définit $\oplus$ sur B par

$$x \oplus y \triangleq (x \vee y) \wedge (\neg x \vee \neg y)$$

Montrer que $(B, \bot, \top, \oplus, \wedge)$ est un anneau (où $\bot$ est le neutre additif et $\top$ le neutre multiplicatif).

Réciproquement, soit $(A, 0, 1, +, \times)$ un anneau idempotent, c'est-à-dire vérifiant

$$\forall x \in A, a \times a = a$$

montrer les propositions suivantes :

(i) A est un anneau de caractéristique 2.

(ii) la relation $\leq_\wedge$ définie par

$$x \leq_\wedge y \triangleq x \times y = x$$

est une relation d'ordre.

(iii) $(A, \leq_\wedge)$ est une algèbre de Boole, de majorant 1, de minorant 0, où $\times$ est la borne inférieure et où la borne supérieure est

$$x \vee y \triangleq x + y + x \times y$$

4.2.4 Algèbre de Heyting

Pour nous intéresser aux algèbres de Heyting, il nous faut d'abord parler de la logique intuitionniste. Pour ne pas nous attarder sur les questions philosophiques qui sont à l'origine de l'essor des mathématiques intuitionnistes et constructives, nous nous contenterons de remarquer que c'est une logique plus faible, mais donnant de meilleurs résultats (les résultats sont plus durs à montrer mais apportent plus d'information).

La logique intuitionniste naît d'une volonté de travailler sans le tiers exclu. D'un point de vue algorithmique, par exemple, ce principe peut être remis en question : comment pouvons-nous effectivement décider toute proposition ? Il est en effet difficile d'imaginer une machinerie finie qui pourrait décider, par exemple, si tel objet est infini. Ainsi une part des mathématiques n'a ni besoin, ni envie du principe du tiers exclu, qui est équivalent au raisonnement par l'absurde. Nous étudierons en détail cette logique lors de notre étude de la théorie de la démonstration, mais nous donnons ici la définition de la logique intuitionniste.

Définition 4.2.4.1. Soit une signature Σ. On définit la relation de déduction syntaxique $\vdash_{\mathrm{NJ}} \subseteq \mathrm{List}(\mathrm{Form}(\Sigma)) \times \mathrm{Form}(\Sigma)$ par les mêmes règles que pour $\vdash$, à l'exception de $\bot_{\mathrm{c}}$ que l'on remplace par la règle suivante :

$$\frac{\Gamma \vdash_{\mathrm{NJ}} \bot}{\Gamma \vdash_{\mathrm{NJ}} A} \bot_{\mathrm{i}}$$

appelée communément principe d'explosion, ou bien par sa dénomination latine *ex falso quodlibet* (version raccourcie de *ex falso sequitur quodlibet*).

Exercice 4.2.4.2. Soit la règle

$$\frac{}{\Gamma \vdash A \vee \neg A} \mathrm{EDN}$$

Montrer que celle-ci est équivalente à $\bot_c$ modulo $\vdash_{NJ}$, c'est-à-dire que si l'on considère $\vdash_{NJ}$ et qu'on rajoute l'une des deux règles, alors la seconde est dérivable.

Exercice 4.2.4.3. Vérifier que dans le cas intuitionniste aussi, la relation $\vdash_{NJ}$ est une relation de pré-ordre. On notera $\mathcal{L}_i(\Sigma)$ l'algèbre de Lindenbaum-Tarski intuitionniste associée à Σ.

Exercice 4.2.4.4. Montrer que $\mathcal{L}_i(\Sigma)$ est un treillis distributif.

On peut alors se demander quelle est la différence entre ce que nous pouvons prouver en logique intuitionniste et ce que nous pouvons prouver en logique classique. La réponse est difficile à donner directement, mais un point important est que ce que l'on peut prouver en logique intuitionniste est constructif : par exemple, prouver qu'un objet existe signifie qu'on a effectivement exhibé un objet. En comparaison, on peut prouver grâce à l'absurde qu'un objet existe en prouvant simplement qu'il ne peut pas ne pas existe.

Une conséquence de cet effet est que certaines tautologies ne sont plus valides, en particulier la tautologie

$$\neg(\forall x, \neg A) \implies \exists x, A$$

exprimant qu'à partir de l'impossibilité que tous les x rejettent A, on peut extraire un certain x acceptant A.

De même, la tautologie

$$\neg(A \wedge B) \implies \neg A \vee \neg B$$

car pour la prouver, il faudrait de façon uniforme décider si $\neg A$ est vraie ou si $\neg B$ est vraie. Comme cela dépend directement que la valeur de vérité de A et de celle de B, il est impossible de conclure en logique intuitionniste. En échange, si l'on prouve $\vdash_{NJ} A \vee B$, cela signifie qu'on peut soit prouver $\vdash_{NJ} A$, soit prouver $\vdash_{NJ} B$.

La question, maintenant, est de savoir s'il existe une structure similaire aux algèbre de Boole, mais dans laquelle la logique appliquée est la logique intuitionniste. C'est le cas des algèbres de Heyting, dont la définition repose avant tout sur celle de l'opération $\to$, d'où l'on dérive le sens de $\neg$ par $\neg x = x \to \bot$, comme nous l'avons déjà fait.

Définition 4.2.4.5 (Pseudo-complément relatif). Soit un treillis distributif $(X, \leq)$, et deux éléments $a, b \in X$. On dit que x est le pseudo-complément de a relatif à b si x est la borne supérieure de l'ensemble $\{x \in X \mid a \wedge x \leq b\}$. Si le pseudo-complément de a relatif à b existe, on le note $a \to b$. On appelle pseudo-complément de a le pseudo complément de a relatif à $\bot$, s'il existe, et on le note alors $\neg x$.

Définition 4.2.4.6 (Algèbre de Heyting). Un treillis distributif $(X, \leq)$ est une algèbre de Heyting si tout élément admet un pseudo-complément relatif à tout autre élément, c'est-à-dire si pour tous $a, b \in X$, l'ensemble $\{x \in X \mid a \wedge x \leq b\}$ admet une borne supérieure.

Remarque 4.2.4.7. Une inégalité par rapport à un pseudo-complément peut se réécrire par une conjonction : $a \leq b \to c$ revient directement à $a \wedge b \leq c$ étant donné que $b \to c$ est la borne supérieure des éléments x tels que $x \wedge b \leq c$. On peut donc formaliser le modus ponens : si l'on a $a \leq b \to c$ et $a \leq b$, alors $a \wedge b \leq c$ et $a \wedge b = a$, donc $a \leq c$.

Exercice 4.2.4.8. Soit une signature Σ. Montrer que $\mathcal{L}_i(\Sigma)$ est bien une algèbre de Heyting.

Exercice 4.2.4.9. Montrer qu'une algèbre de Boole est en particulier une algèbre de Heyting, et que le complément d'un élément est en particulier son pseudo-complément.

4.2.5 Treillis complet

Notre première définition d'un treillis peut paraître légèrement plus artificelle que sa caractérisation algébrique. En effet, dans la version algébrique, on peut voir un treillis comme un certain bimonoïde, que l'on peut ensuite étendre en l'existence d'une borne supérieure et d'une borne inférieure pour toute partie finie (en raisonnant juste par récurrence). A l'inverse, lorsque l'on part de la définition de treillis comme ensemble ordonné possédant toutes les bornes sur des parties finies, une question naît immédiatement : pourquoi ne pas autoriser des parties infinies ? C'est pour intégrer cette possibilité que nous introduisons la notion de treillis complet.

Définition 4.2.5.1 (Treillis complet). Un treillis complet est un ensemble ordonné $(X, \leq)$ tel que toute partie $Y \subseteq X$ possède à la fois une borne supérieure et une borne inférieure.

Un treillis complet est, par définition, plus fort qu'un treillis. Dans la pratique, on trouve effectivement plusieurs treillis qui ne sont pas complets (le fait d'avoir deux notions n'est donc pas une distinction simplement formelle mais revêt une vraie différence conceptuelle). Remarquons aussi qu'un treillis complet peut se caractériser par l'existence d'une seule borne.

Propriété 4.2.5.2. *L'ensemble ordonné $(X, \leq)$ est un treillis complet si et seulement si toute partie possède une borne supérieure.*

Démonstration. Un sens est évident. Pour l'autre, supposons que toute partie de X possède une borne supérieure, et montrons alors qu'elle possède une borne inférieure.

Soit $Y \subseteq X$, on définit $Y\downarrow$ par

$$Y\downarrow \triangleq \{x \in X \mid \forall y \in Y, x \leq y\}$$

et soit y_0 la borne supérieure de $Y\downarrow$. Montrons que y_0 est la borne inférieure de Y. Tout d'abord, si $y \in Y$, alors par définition, y est un majorant de $Y\downarrow$, donc puisque y_0 est le plus petit majorant, y_0 est un minorant de Y. De plus, si y est un minorant de Y, cela signifie que $y \in Y\downarrow$, donc comme y_0 est un majorant de $Y\downarrow$, on en déduit que $y \leq y_0$. Ainsi y_0 est le plus grand minorant de Y : c'est la borne inférieure de Y. □

On peut maintenant prouver l'un des théorèmes importants que nous avons présenté dans le chapitre 1 : le théorème 1.2.1.3, mais dans une version bien plus forte et générale.

Théorème 4.2.5.3 (Knaster-Tarski). *Soit $(X, \leq)$ un treillis complet et $f : X \to X$ une fonction croissante. Alors l'ensemble*

$$\mathrm{fix}(f) \triangleq \{x \in X \mid f(x) = x\}$$

est un treillis complet pour l'ordre induit par $\leq$.

Démonstration. On définit tout d'abord l'ensemble des pré points fixes de f :

$$\mathrm{prefix}(f) \triangleq \{x \in X \mid f(x) \leq x\}$$

et on montre qu'il est un treillis complet pour $\leq$. Soit une partie $Y \subseteq \mathrm{prefix}(f)$, montrons que Y possède une borne inférieure. Nous prenons pour candidat de la borne inférieure de Y dans $\mathrm{prefix}(f)$, la borne inférieure $\bigwedge Y$ dans X : montrons donc que $\bigwedge Y$ est bien un pré point fixe de f (cela suffira à montrer qu'on a une borne inférieure, puisque la borne

inférieure dans Y est aussi la borne inférieure dans $\text{prefix}(f)$, à partir du moment où la borne appartient effectivement à $\text{prefix}(f)$).

On montre que $f(\bigwedge Y) \leq \bigwedge Y$. Pour cela, il nous suffit de montrer que $f(\bigwedge Y)$ est un minorant de Y, puisque l'on sait que $\bigwedge Y$ est le plus grand minorant. Soit $y \in Y$, on remarque que $\bigwedge Y \leq y$, donc $f(\bigwedge Y) \leq f(y)$, mais comme par définition $f(y) \leq y$ et par transitivité, on en déduit que $f(\bigwedge Y) \leq y$. Ainsi $f(\bigwedge Y) \leq \bigwedge Y$.

On sait donc que toute partie de $\text{prefix}(f)$ admet une borne inférieure. Par la propriété précédente (son dual, plus précisément), $\text{prefix}(f)$ est donc un treillis complet pour $\leq$. On a donc un treillis complet $(\text{prefix}(f), \leq)$ et une fonction $f : \text{prefix}(f) \to \text{prefix}(f)$. En prenant le dual, on obtient donc que $\text{postfix}(\text{prefix}(f))$ est un treillis complet.

Montrons maintenant que $\text{postfix}(\text{prefix}(f)) = \text{fix}(f)$. Par définition,

$$\text{postfix}(\text{prefix}(f)) = \{x \in \{x \in X \mid f(x) \leq x\} \mid x \leq f(x)\}$$

ce qui correspond exactement à l'ensemble $\{x \in X \mid (f(x) \leq x) \wedge (x \leq f(x))\}$. Comme $\leq$ est transitif, cet ensemble est $\text{fix}(f)$. Ainsi $\text{fix}(f)$ est un treillis complet. □

Remarque 4.2.5.4. On en déduit donc, en particulier, qu'il existe un plus petit et un plus grand point fixe pour f. En rentrant un peu plus en détail dans la démonstration, on peut vérifier que le plus petit point fixe est le plus petit pré point fixe, et que le plus grand point fixe est le plus grand post point fixe.

Exercice 4.2.5.5. Montrer que $(\mathcal{P}(X), \subseteq)$ est un treillis complet, pour tout ensemble X. En déduire le théorème 1.2.1.3.

Enfin, donnons deux mots à propos des algèbre de Heyting complètes. On pourrait s'attendre à ce qu'un treillis complet X suffise à modéliser ce qu'on attend d'une algèbre de Heyting complète, étant donné que pour tous a, b, l'ensemble

$$\{x \in X \mid a \wedge x \leq b\}$$

possède une borne inférieure. Ce qui bloque, cependant, est le caractère distributif : on ne sait pas si le treillis complet est distributif. A la place, une algèbre de Heyting complète va posséder une version complète de la distributivité.

Définition 4.2.5.6 (Algèbre de Heyting complète). Un ensemble ordonné $(H, \leq)$ est une algèbre de Heyting complète si $(H, \leq)$ est un treillis complet et s'il vérifie l'identité suivante, pour toute partie $X \subseteq H$ et tout élément $y \in H$:

$$\left(\bigvee_{x \in X} x\right) \wedge y = \bigvee_{x \in X} (x \wedge y)$$

4.3 Filtres

Enfin, présentons la notion de filtre, et la notion duale d'idéal. Nous allons détailler comment se comportent les filtres dans les structure de plus en plus fortes : demi-treillis, treillis et algèbre de Boole, puis nous verrons les ultrafiltres et le lemme de l'ultrafiltre.

4.3.1 Définitions et caractérisations

Commençons par définir ce qu'est un filtre.

Définition 4.3.1.1 (Filtre). Soit $(X, \leq)$ un ensemble ordonné. Un filtre est une partie $\mathcal{F} \subseteq X$ vérifiant :

- $\mathcal{F}$ est non vide : $\mathcal{F} \neq \varnothing$.
- $\mathcal{F}$ est clos par le haut : pour tout $x \in \mathcal{F}$ et $y \in X$, si $x \leq y$ alors $y \in \mathcal{F}$.
- tous les éléments de $\mathcal{F}$ sont compatibles : pour tous $x, y \in \mathcal{F}$, il existe $z \in \mathcal{F}$ tel que $z \leq x$ et $z \leq y$.

Un filtre $\mathcal{F}$ est dit propre si $\mathcal{F} \neq X$.

Un filtre peut donc être vu comme une description d'une notion de grandeur : un élément d'un filtre est vu comme un « grand » élément. En plus de cela, cette notion de grandeur doit avoir une notion de compatibilité : deux grands éléments doivent avoir en commun un grand élément. La notion duale, qui peut donc être vue comme la spécification d'une notion de petitesse, est celle d'idéal.

Définition 4.3.1.2 (Idéal). Soit $(X, \leq)$ un ensemble ordonné. Un idéal est une partie $\mathcal{I} \subseteq X$ vérifiant :

- $\mathcal{I}$ est non vide : $\mathcal{I} \neq \varnothing$.
- $\mathcal{I}$ est clos par le bas : pour tout $x \in \mathcal{I}$ et $y \in X$, si $y \leq x$ alors $y \in \mathcal{I}$.
- tous les éléments de $\mathcal{I}$ sont compatibles par le haut : pour tous $x, y \in \mathcal{I}$, il existe $z \in \mathcal{I}$ tel que $x \leq z$ et $y \leq z$.

Un idéal $\mathcal{I}$ est dit propre si $\mathcal{I} \neq X$.

Nous ne travaillerons par sur les idéaux, mais donnons la définition car celle-ci reste importante.

Exemple. Soit un ensemble X et un élément $x \in X$. On peut définir le filtre principal sur $\mathcal{P}(X)$ en x, $\mathcal{F}_x$, par

$$\mathcal{F}_x \triangleq \{A \subseteq X \mid x \in A\}$$

On voit que $X \in \mathcal{F}_x$, que si $A \subseteq B$ et $x \in A$ alors $x \in B$, et que si $x \in A$ et $x \in B$ alors $x \in A \cap B$. Ainsi l'ensemble ordonné $\mathcal{P}(X)$ contient au moins autant de filtres que d'éléments dans X.

On peut donner une première caractérisation des filtres, dans le cas d'un inf demi-treillis.

Propriété 4.3.1.3. *Soit $(X, \leq)$ un inf demi-treillis de majorant $\top$ et de borne inf $\wedge$. Alors $\mathcal{F} \subseteq X$ est un filtre sur X si et seulement si les conditions suivantes sont vérifients :*

- $\top \in X$.
- *$\mathcal{F}$ est clos par le haut.*
- *si $x, y \in \mathcal{F}$, alors $x \wedge y \in \mathcal{F}$.*

Démonstration. Supposons que $\mathcal{F}$ vérifie les propriétés listées. Le fait d'être clos par le haut est par définition. Comme $\top \in \mathcal{F}$, $\mathcal{F}$ est non vide. Si $x, y \in \mathcal{F}$, alors il existe effectivement $z \in \mathcal{F}$ tel que $z \leq x$ et $z \leq y$, en prenant $z = x \wedge y$.

Réciproquement, supposons que $\mathcal{F}$ est un filtre. Alors, puisqu'il possède un élément, disons x, et que $x \leq \top$, on en déduit par clôture par le haut que $\top \in \mathcal{F}$. Le fait que $\mathcal{F}$ est clos par le haut est par hypothèse. Soient $x, y \in \mathcal{F}$, alors par hypothèse il existe $z \in \mathcal{F}$ tel que $z \leq x$ et $z \leq y$. Par définition de la borne supérieure, on a donc que $z \leq x \wedge y$, et par clôture par le haut on en déduit que $x \wedge y \in \mathcal{F}$. □

Les inf demi-treillis permettent donc de mieux caractériser les treillis. En quelque sorte, puisqu'ils permettent de délimiter le fait de devenir plus petit, il est plus aisé d'y définir ce qui est grand (c'est-à-dire ce qui est arrive à rester grand quand on le rend plus petit). Les sup demi-treillis, eux, ne permettent pas de mieux définir un treillis, mais permettent de décrire ses propriétés.

Propriété 4.3.1.4. *Soit $(X, \leq)$ un sup demi-treillis et $\mathcal{F} \subseteq X$ un filtre. Alors $\mathcal{F}$ est propre si et seulement si $\bot \notin \mathcal{F}$.*

Démonstration. Puisque $\mathcal{F}$ est clos par le haut, si $\bot \in \mathcal{F}$ alors tout élément est dans $\mathcal{F}$, donc $\mathcal{F}$ n'est pas propre (et si $\mathcal{F}$ n'est pas propre, alors il contient $\bot$). □

Les sup demi-treillis permettent de définir la propriété d'être premier :

Définition 4.3.1.5 (Filtre premier). Soit $(X, \leq)$ un sup demi-treillis et $\mathcal{F}$ un filtre sur X. On dit que $\mathcal{F}$ est premier si pour tous $x, y \in X$, si $x \vee y \in \mathcal{F}$ alors $x \in F$ ou $y \in F$.

On peut voir une analogie ici avec un idéal dans un anneau : un idéal premier est un idéal tel que, lorsqu'on a un produit $a \times b$ dedans, celui-ci doit nécessairement provenir du produit par un élément de l'idéal. De la même façon, ici, on sait naturellement que si $x \in \mathcal{F}$ alors $x \vee y \in \mathcal{F}$, et la notion de primalité signifie que tout élément de la forme $x \vee y$ doit provenir d'un cas de ce type, où x (ou y) était déjà dans $\mathcal{F}$.

Pour synthétiser les différentes propriétés dans le cas d'un treillis : les filtres propres (qui sont les filres intéressants) sont les parties closes par le haut qui possède $\top$ mais pas $\bot$, et qui sont closes par intersection finie.

Supposons maintenant qu'on possède une partie X (non vide) qui n'est pas un filtre. Comment trouver un filtre correspondant naturellement à X ? Premièrement, il nous faut d'abord clore X par le haut : on prend l'ensemble des éléments supérieurs à au moins un élément de X. Il nous faut aussi nous assurer que pour tout x, y dans notre filtre, un élément inférieur à x et y soit dans le filtre. Le plus simple ici est de considérer un inf demi-treillis : on veut ajouter tous les $x \wedge y$, mais comme il est possible d'itérer ces bornes inférieures, il nous faut ajouter tous les $\bigwedge F$ pour $F \subseteq_{\text{fin}} X$. On obtient ainsi la construction suivante :

Définition 4.3.1.6 (Filtre engendré par une partie). Soit $(X, \leq)$ un inf demi-treillis et $Y \subseteq X$ une partie non vide de X. On définit alors

$$\overline{Y}^{\mathcal{F}} \triangleq \{x \in X \mid \exists F \subseteq_{\text{fin}} Y, \bigwedge F \leq x\}$$

Proposition 4.3.1.7. *Soit $(X, \leq)$ un inf demi-treillis et $Y \subseteq X$ une partie non cide. Alors $\overline{Y}^{\mathcal{F}}$ est un filtre sur X.*

Démonstration. Tout d'abord, on remarque que $Y \subseteq \overline{Y}^{\mathcal{F}}$, donc $\overline{Y}^{\mathcal{F}}$ est non vide. Soit $x \in \overline{Y}^{\mathcal{F}}$ et $y \in X$ tel que $x \leq y$. On trouve $F \subseteq_{\text{fin}} Y$ tel que $\bigwedge F \leq x$, alors $\bigwedge F \leq y$ par transitivité, donc $y \in \overline{F}^{\mathcal{F}}$. Soient x et y des éléments de $\overline{Y}^{\mathcal{F}}$, on trouve F et F' des parties finies de Y telles que $\bigwedge F \leq x$ et $\bigwedge F' \leq y$. Alors

$$\bigwedge(F \cup F') = \left(\bigwedge F\right) \wedge \left(\bigwedge F\right) \leq x \wedge y$$

donc $x \wedge y \in \overline{Y}^{\mathcal{F}}$. Ainsi $\overline{Y}^{\mathcal{F}}$ est un filtre. □

Par défaut, le filtre engendré par une partie peut tout à fait contenir tous les éléments. Prenons par exemple un ensemble X, une partie Y et le filte engendré par $\{Y, X \setminus Y\}$: il est évident que ce filtre n'est pas propre. Cependant, il existe une propriété assurant qu'un filtre engendré est propre, est qui est tout à fait naturelle : on veut que les $\bigwedge F$ soient tous différents de l'ensemble vide. Evidemment, cette propriété n'a d'intérêt que dans le cas d'un treillis.

Définition 4.3.1.8 (Propriété de l'intersection finie). Soit $(X, \leq)$ un treillis et $Y \subseteq X$. On dit que Y a la propriété de l'intersection finie si

$$\forall F \subseteq_{\text{fin}} Y, \bigwedge F \neq \bot$$

Proposition 4.3.1.9. *Soit $(X, \leq)$ un treillis et $Y \subseteq X$ une partie non vide de X. Alors $\overline{Y}^{\mathcal{F}}$ est un filtre propre si et seulement si Y a la propriété de l'intersection finie.*

Démonstration. Par définition, $\bot \in \overline{Y}^{\mathcal{F}}$ si et seulement s'il existe $F \subseteq_{\text{fin}} Y$ tel que $\bigwedge F \leq \bot$, mais la condition que $\bigwedge F \leq \bot$ est équivalente à ce que $\bigwedge F = \bot$, étant donné que l'autre inégalité est toujours vérifiée. Ainsi $\bot \in \overline{Y}^{\mathcal{F}}$ si et seulement s'il existe $F \subseteq_{\text{fin}} Y$ tel que $\bigwedge F = \bot$, donc $\overline{Y}^{\mathcal{F}}$ est propre si et seulement si Y possède la propriété de l'intersection finie. □

Exercice 4.3.1.10. Soit Σ une signature. On considère $\mathcal{T} \subseteq \text{Clos}(\Sigma)$ une théorie. Montrer que $\mathcal{T}$ est cohérente (c'est-à-dire que $\mathcal{T} \nvdash \bot$) si et seulement si $\mathcal{T}'$, la projection de $\mathcal{T}$ dans $\mathcal{L}(\Sigma)$, a la propriété de l'intersection finie.

Exercice 4.3.1.11. Soit Σ une signature. Soit $\mathcal{T} \subseteq \text{Clos}(\Sigma)$ une théorie. Montrer que la projection de $\overline{\mathcal{T}}^{\models}$ dans $\mathcal{L}(\Sigma)$ est le filtre engendré par la projection de $\mathcal{T}$ dans $\mathcal{L}(\Sigma)$.

4.3.2 Ultrafiltre

On définit maintenant la notion d'ultrafiltre, qui est la notion duale de celle de filtre maximale, souvent rencontrée en algèbre.

Définition 4.3.2.1 (Ultrafiltre). Soit $(X, \leq)$ un ensemble ordonné. On dit qu'un filtre $\mathcal{F}$ est un ultrafiltre s'il est un filtre propre maximal pour l'inclusion, c'est-à-dire si pour tout autre filtre propre $\mathcal{H}$ de X, si $\mathcal{F} \subseteq \mathcal{H}$ alors $\mathcal{F} = \mathcal{H}$.

Dans le cas d'un treillis distributif, on a une coïncidence entre les ultrafiltres et les filtres premiers.

Proposition 4.3.2.2. *Soit $(X, \leq)$ un treillis distributif. Alors un filtre $\mathcal{F}$ est un ultrafiltre si et seulement s'il est un filtre premier.*

Démonstration. A FAIRE □

On a, de plus, une caractérisation encore meilleure dans le cas d'une algèbre de Boole.

Proposition 4.3.2.3. *Soit $(B, \leq)$ une algèbre de Boole. Un filtre $\mathcal{F}$ est un ultrafiltre si et seulement si on a la propriété suivante, pour tout $x \in B$:*

$$x \in \mathcal{F} \iff \neg x \notin \mathcal{F}$$

Démonstration. Si $\mathcal{F}$ est un filtre premier, alors comme $\top \in \mathcal{F}$ et $\top = x \vee \neg x$, on sait que soit x soit $\neg x$ est dans $\mathcal{F}$. Comme par définition $\bot \notin \mathcal{F}$ et que $x \wedge \neg x = \bot$, on en déduit qu'un seul des deux est dans $\mathcal{F}$, ce qui est le résultat qu'on veut démontrer.

Supposons maintenant que pour tout $x \in B$, $x \in \mathcal{F} \iff \neg x \notin \mathcal{F}$. Alors si $x \vee y \in \mathcal{F}$, supposons que $x \notin \mathcal{F}$ et $y \notin \mathcal{F}$. Alors $\neg x \in \mathcal{F}$ et $\neg y \in \mathcal{F}$ grâce à notre hypothèse (et au fait que $\neg\neg x = x$). Mais alors $\neg x \wedge \neg y \in \mathcal{F}$, et $x \vee y \in \mathcal{F}$, donc $\neg(x \vee y)$ et $x \vee y$ sont dans $\mathcal{F}$: $\mathcal{F}$ n'est donc par propre. Par l'absurde, on en déduit que soit $x \in \mathcal{F}$, soit $y \in \mathcal{F}$. □

Pour finir, nous allons démontrer le lemme de l'ultrafiltre, qui est une conséquence du lemme de Zorn. Il dit que tout filtre propre peut être étendu en un ultrafiltre.

Théorème 4.3.2.4 (Lemme de l'ultrafiltre). *Soit $(X, \leq)$ un treillis. Soit $\mathcal{F}$ un filtre sur X. Alors il existe un ultrafiltre $\mathcal{U}$ contenant $\mathcal{F}$.*

Démonstration. On définit l'ensemble ordonné

$$\mathcal{X} \triangleq \{\mathcal{H} \subseteq X \mid \mathcal{H} \text{ est un filtre propre}\}$$

muni de l'inclusion ensembliste.

On veut appliquer le lemme de Zorn sur cet ensemble. Pour cela, on vérifie que $\mathcal{X}$ est inductif.

Soit $\{\mathcal{F}_i\}_{i \in I}$ une chaîne d'éléments de $\mathcal{X}$, montrons que $\mathcal{F}_0 = \bigcup_{i \in I} \mathcal{F}_i$ est un majorant de cette chaîne. Il nous suffit pour cela de vérifier que $\mathcal{F}_0$ est bien un filtre propre :

- soit $i \in I$, comme $\top \in \mathcal{F}_i$, on en déduit que $\top \in \mathcal{F}_0$.
- soit $x \in \mathcal{F}_0$ et $y \in X$ tel que $x \leq y$. Par définition de $\mathcal{F}_0$, on trouve $i \in I$ tel que $x \in \mathcal{F}_i$, et comme $\mathcal{F}_i$ est un filtre on en déduit que $y \in \mathcal{F}_i$, d'où $y \in \mathcal{F}_0$.
- soient $x, y \in \mathcal{F}_0$. On trouve $i, j \in I$ tels que $x \in \mathcal{F}_i$ et $y \in \mathcal{F}_j$. Comme $\{\mathcal{F}_i\}$ est une chaîne pour l'inclusion, alors soit $\mathcal{F}_i \subseteq \mathcal{F}_j$, soit $\mathcal{F}_j \subseteq \mathcal{F}_i$. Sans perte de généralité, on suppose donc que $\mathcal{F}_i \subseteq \mathcal{F}_j$, donc que $x \in \mathcal{F}_j$. Comme $\mathcal{F}_j$ est un filtre, on en déduit que $x \wedge y \in \mathcal{F}_j$, donc que $x \wedge y \in \mathcal{F}_0$.
- supposons que $\bot \in \mathcal{F}_0$. Alors on trouve $i \in I$ tel que $\bot \in \mathcal{F}_i$, mais comme tous les $\mathcal{F}_i$ sont supposés propres, $\bot \notin \mathcal{F}_i$. Par contradiction, on en déduit que $\mathcal{F}_0$ est propre.

On applique donc le théorème 4.1.3.12 : on trouve un élément maximal $\mathcal{H}$ de $\mathcal{X}$ qui est supérieur à $\mathcal{F}$, c'est-à-dire qui contient $\mathcal{F}$. Il ne nous reste qu'à vérifier que $\mathcal{H}$ est bien un ultrafiltre : si $\mathcal{G}$ est un filtre propre contenant $\mathcal{H}$, alors il contient $\mathcal{F}$, donc il appartient à $\mathcal{X}$. Par maximalité de $\mathcal{H}$ dans $\mathcal{X}$, on en déduit que $\mathcal{H} = \mathcal{G}$: $\mathcal{H}$ est un ultrafiltre sur X contenant $\mathcal{F}$. □

On peut maintenant montrer le théorème 3.2.3.7 :

Démonstration. Soit une signature Σ et une théorie $\mathcal{T}$ sur Σ supposée cohérente. On peut définir l'algèbre de Lindenbaum-Tarski pour la sémantique :

$$\mathcal{L}(\Sigma) \triangleq \mathrm{Form}(\Sigma)/\equiv$$

(comme nous prouvons un résultat permettant la démonstration du théorème de complétude, nous ne pouvons pas utiliser l'équivalence entre $\vdash$ et $\models$, donc nous construisons cette fois-ci l'algèbre en utilisant la relation $\models$ dès le début).

On laisse au lecteur le soin de vérifier que cet ensemble ordonné est bien une algèbre de Boole avec les constructeurs induits par $\top, \bot, \wedge, \vee, \neg$.

La théorie $\mathcal{T}$ induit sur $\mathcal{L}(\Sigma)$ une partie vérifiant la propriété d'intersection finie, et la théorie $\overline{\mathcal{T}}^{\models}$ est ainsi un filtre propre : on peut donc grâce au théorème 4.3.2.4 trouver un ultrafiltre $\mathcal{S}$ contenant $\overline{\mathcal{T}}^{\models}$.

Soit alors la théorie $\mathcal{T}' = \bigcup \mathcal{S}$, qui contient donc toutes les formules appartenant à une classe d'équivalence de $\mathcal{S}$. On remarque que $\mathcal{T}'$ est close par $\models$, donc puisque $\overline{\bot} \notin \mathcal{S}$, cela signifie que $\bot \notin \mathcal{T}'$. Il nous reste à vérifier que $\mathcal{T}'$ est complète : pour toute formule φ, on sait que $\varphi \vee \neg\varphi \in \mathcal{T}'$, donc $\overline{\varphi} \vee \overline{\neg\varphi} \in \mathcal{S}$, mais comme $\mathcal{S}$ est un ultrafiltre, il est en particulier premier. On en déduit que soit $\overline{\varphi} \in \mathcal{S}$, soit $\overline{\neg\varphi} \in \mathcal{S}$: dans le premier cas, $\varphi \in \mathcal{T}'$ et dans le deuxième cas, $\neg\varphi \in \mathcal{T}'$. Ainsi pour toute formule φ, soit $\varphi \in \mathcal{T}'$ soit $\neg\varphi \in \mathcal{T}'$ (c'est donc en particulier encore le cas pour $\overline{\mathcal{T}'}^{\models} = \mathcal{T}'$).

Donc il existe une théorie complète $\mathcal{T}'$ qui contient la théorie $\mathcal{T}$. □

Deuxième partie

Théorie des ensembles

Chapitre 5

Formaliser les mathématiques

Pour commencer notre étude, nous allons étudier l'une des propriétés les plus connues de la théorie des ensembles : la capacité d'expression suffisante pour formaliser toutes les mathématiques usuelles. Ce chapitre se concentrera donc sur l'étude des axiomes de ZFC et de leur utilisation pour construire les objets mathémtiques que nous connaissons habituellement.

Nous devons ici faire quelques éclaircissements d'ordre philosophique. Tout d'abord, sur la méta-théorie : nous considérons que l'univers ambiant dans lequel nous pratiquons les mathématiques est lui-même un modèle de la théorie ZFC. Plus précisément, en notant $\mathcal{U}$ l'objet mathématique qui sera la collection de tous les objets mathématiques, ZFC peut être vu comme une théorie donnant une approximation du comportement de $\mathcal{U}$.

En particulier, $\mathcal{U}$ contient des collections, que nous appelons d'habitude ensembles, et si un objet X apparait dans une collection C, on écrit $X \in C$. Cependant, les objets que nous manipulerons ne seront pas ces collections, et la relation $\in$ utilisée dans ZFC ne sera pas « être un objet mathématique apparaissant dans cette collection intuitive » : nous utiliserons des outils purement formels pour étudier ZFC, et devons donc les distinguer des objets intuitifs associés. Aussi nous appellerons « ensemble » un objet de la théorie ZFC, et « collection » un ensemble au sens intuitif, dans $\mathcal{U}$. De même, les objets formels seront associés à la notion d'appartenance, et les collections à la notion d'occurrence.

Remarquons cependant qu'un ensemble donne de fait naissance à une collection : l'ensemble X définit la collection $\{x \in \mathcal{U} \mid x \in X\}$.

Enfin, nous traiterons régulièrement des classes. S'il est possible de ne se restreindre qu'aux ensembles pour l'étude de ZFC, nous verrons qu'il est bien plus pratique d'énoncer certaines constructions en terme de classes. Une classe est une collection qui ne peut pas être représentée par un ensemble (par exemple la classe des ensembles, *cf.* le paradoxe de Russell), mais que l'on peut décrire par un prédicat. Ainsi, une classe peut se représenter par le prédicat lui correspondant. Si une classe C est décrite par un prédicat $\varphi(x)$, alors $x \in C$ doit se lire comme une reformulation plus lisible de $\varphi(x)$. Un léger défaut vient avec cette approche : certains théorèmes, se référant aux classes, doivent se lire comme des schémas de théorèmes, c'est-à-dire des théorèmes paramétrés par le prédicat φ définissant une classe.

5.1 Axiomes de ZFC

5.1.1 Extensionalité, paire, union, ensemble des parties

Commençons par étudier la théorie ZFC en en listant les axiomes. Un premier point important est de définir le langage sur lequel nous travaillerons : nous utiliserons $\mathcal{L}_{\text{ZF}} = \{\in^2\}$. En effet, tous les énoncés seront écrits seulement à l'aide du symbole de relation binaire $\in$ et de la relation $=$, symbole logique déjà inclus dans tout langage.

Remarque 5.1.1.1. En réalité, nous le verrons, notre langage sera bien plus riche : nous ajouterons des symboles de fonctions, de constantes, et diverses relations. Cependant, tous ces ajouts doivent être considérés comme des aides à la lecture : toute proposition formulée dans l'enrichissement que nous donnerons au fur et à mesure doit pouvoir se formuler dans L_{ZF}, mais avec un nombre possiblement bien plus grand de symboles.

Axiome 5.1.1.2 (Extensionalité). *L'axiome d'extensionalité exprime que deux ensembles sont égaux exactement lorsqu'ils possèdent les mêmes éléments :*

$$\forall x \forall y, \qquad x = y \iff (\forall z, z \in x \iff z \in y)$$

Cet axiome définit ce que signifie l'égalité dans le monde des ensembles. Cela permet directement de voir que l'ordre ou le nombre d'occurrences n'importent pas dans un ensemble, contrairement par exemple au cas des listes. C'est un axiome au statut particulier car il ne permet pas de construire de nouvel ensemble. Les autres axiomes, eux, donneront principalement de nouvelles méthodes pour, à partir d'ensembles déjà construits, définir de nouveaux ensembles.

Axiome 5.1.1.3 (Paire). *L'axiome de la pair exprime que si deux ensembles x et y ont été construits, alors l'ensemble $\{x, y\}$ peut être construit à partir d'eux :*

$$\forall x \forall y \exists z, \qquad \forall a, a \in z \iff a = x \vee a = y$$

Exercice 5.1.1.4. Montrer que pour tous ensembles x, y, il existe un unique ensemble z tel que décrit dans l'axiome 5.1.1.3. Cela justifie donc la notation $\{x, y\}$ pour cet ensemble, puisqu'il n'y a pas d'ambigüité sur lequel il est.

En utilisant l'axiome de la paire, on peut ainsi construire des collections finies telles que $\{x, \{y, z\}\}$, mais on ne peut pas par exemple définir $\{x, y, z\}$: il nous faut un axiome permettant d'accéder à l'intérieur d'un ensemble pour en construire un nouveau.

Axiome 5.1.1.5 (Union). *Pour tout ensemble x, on peut construire l'ensemble $\bigcup x$ contenant les éléments des éléments de x :*

$$\forall x \exists y, \qquad \forall z, z \in y \iff (\exists a, z \in a \wedge a \in x)$$

Notation 5.1.1.6. De la même façon qu'avec l'axiome de la paire, l'ensemble y défini par l'axiome est unique, on l'écrira donc $\bigcup x$.

Exercice 5.1.1.7. Soient $x_1, \ldots, x_n$ des ensembles, montrer par récurrence qu'il existe l'ensemble $\{x_1, \ldots, x_n\}$.

Notation 5.1.1.8. Etant donnée une famille $x_1, \ldots, x_n$ d'ensembles, on définit $\bigcup_{i=1}^{n} x_i$ comme $\bigcup\{x_1, \ldots, x_n\}$. En particulier, $x \cup y$ est $\bigcup\{x, y\}$.

Enfin, l'axiome de l'ensemble des parties permet de considérer comme un ensemble la collection des parties d'un ensemble. En un sens, il permet de faire croître considérablement la taille de ce que l'on peut construire.

Notation 5.1.1.9. On définit le prédicat binaire $\subseteq$, appelé l'inclusion, par

$$x \subseteq y \triangleq \forall z, z \in x \implies z \in y$$

Axiome 5.1.1.10 (Ensemble des parties). *Pour tout ensemble x, il existe l'ensemble $\mathcal{P}(x)$ dont les éléments sont exactement les ensembles inclus dans x :*

$$\forall x \exists y, \qquad \forall z, z \in y \iff z \subseteq x$$

Notation 5.1.1.11. Comme le y défini plus haut est unique, on peut là encore lui donner une notation, qui est bien sûr $\mathcal{P}(x)$.

5.1.2 Les schémas d'axiomes

Les axiomes précédemment donnés constituent les briques de base pour construire des ensembles, mais sont en général trop grossières. L'intérêt de la théorie des ensembles est de pouvoir construire des ensembles correspondant à des collections, ce qui manque pour l'instant à notre système. C'est pour cela que nous ajoutons l'axiome de compréhension : étant donnée une formule φ et un ensemble X, on peut construire l'ensemble $\{x \in X \mid \varphi(x)\}$ des éléments de x vérifiant φ.

La restriction de la compréhension à des parties d'un ensemble s'explique par le paradoxe de Russell : si l'on pouvait construire un ensemble pour chaque formule, on pourrait construire $\{x \mid x \notin x\}$, qui appartient à lui-même si et seulement s'il n'appartient pas à lui-même.

Un autre problème doit être contourné, et il est la raison pour laquelle on parle de schéma d'axiomes et non d'axiome. Si l'on voulait définir le schéma avec ce qui a été dit, celui-ci commencerait moralement par $\forall \varphi$: cela n'est pas une quantification du premier ordre, et il n'est donc pas possible de donner un axiome pour toutes les formules. A la place, on introduit un schéma d'axiomes, c'est-à-dire une infinité d'axiomes ayant tous la même forme et dépendant d'un paramètre que l'on quantifie sur les formules.

Axiome 5.1.2.1 (Schéma de compréhension). *Soit $\varphi(x_0, \ldots, x_n)$ une phrase mathématique à n variables libres. Alors pour tous ensembles X et $a_1, \ldots, a_n$, il existe l'ensemble $\{x \in X \mid \varphi(x, a_1, \ldots, a_n)\}$, ce qui s'écrit formellement*

$$\forall X \forall a_1 \cdots a_n \exists y, \qquad \forall x, x \in y \iff (x \in X \wedge \varphi(x, a_1, \ldots, a_n))$$

Notation 5.1.2.2. On définit donc la notation $\{x \in X \mid \varphi(x, a_1, \ldots, a_n)\}$ pour l'ensemble précédent.

Exercice 5.1.2.3. Soit un ensemble x non vide, montrer qu'il existe l'ensemble $\bigcap x$ des éléments qui sont dans tous les éléments de x.

Notation 5.1.2.4. On définit des notations pour l'intersection analogues à celles pour l'union : $\bigcap_{i=1}^{n} x_i$ et $x \cap y$.

Exercice 5.1.2.5. Soient x et y deux ensembles. On définit le couple (x, y) par

$$(x, y) \triangleq \{\{x, y\}, \{x\}\}$$

Montrer que pour tous x, y, x', y', $(x, y) = (x', y')$ si et seulement si $x = x'$ et $y = y'$.

Construire un prédicat $\varphi(x, y, z)$ tel que $\varphi(x, y, z)$ est vrai si et seulement si $z = (x, y)$. En déduire en considérant une partie bien choisie de $\mathcal{P}(\mathcal{P}(x \cup y))$ que

$$x \times y \triangleq \{(a, b) \mid a \in x, b \in y\}$$

est un ensemble bien défini.

Le second axiome, le schéma d'axiomes de remplacement, peut être vu comme une version plus forte de la compréhension. Plutôt que de s'intéresser à filtrer des éléments dans un ensemble plus gros, le but de ce schéma d'axiomes est de construire un ensemble par une fonction. Comme la notion de fonction n'est pas encore définie, nous utilisons à la place la notion de relation fonctionnelle.

Définition 5.1.2.6 (Relation fonctionnelle). Une relation binaire est ici une formule à deux variables libres. On dit qu'une relation $R(x, y)$ est fonctionnelle si pour chaque x, il existe au plus un y tel que $R(x, y)$. On écrira pour raccourcir

$$\text{Funct}(R) \triangleq \forall x \forall y \forall z, R(x, y) \wedge R(x, z) \implies y = z$$

Pour une relation fonctionnelle $R(x, y)$, on définit la collection du domaine de R :

$$\text{Dom}(R)(x) \triangleq \exists y. R(x, y)$$

et la collection de l'image de R :

$$\text{Im}(R)(y) \triangleq \exists x. R(x, y)$$

Axiome 5.1.2.7 (Schéma de remplacement). *Pour toute formule $R(x_0, \ldots, x_{n+1})$, pour tous ensembles $X, a_1, \ldots, a_n$, l'ensemble image de $R(a_1, \ldots, a_n)$ par X est aussi un ensemble :*

$$\text{Funct}(R) \implies \forall X \forall a_1 \cdots a_n \exists y, \qquad \forall x, x \in y \iff (\exists z, z \in X \wedge R(a_1, \ldots, a_n, z, x))$$

Exercice 5.1.2.8. En remarquant qu'une relation fonctionnelle peut représenter une fonction partielle, montrer que le schéma d'axiomes de compréhension peut se déduire du schéma d'axiomes de remplacement.

Exercice 5.1.2.9. En réutilisant le prédicat φ de l'exercice 5.1.2.5, montrer grâce au schéma d'axiomes de remplacement que pour tous ensembles x, y, l'ensemble $x \times y$ est bien défini même sans l'axiome de l'ensemble des parties.

Donnons aussi l'axiome le plus évident.

Axiome 5.1.2.10 (Univers non vide). *Il existe un ensemble.*

Exercice 5.1.2.11. Montrer que l'axiome précédent est équivalent à l'existe de l'ensemble vide $\varnothing$ défini par

$$\forall z, z \notin \varnothing$$

(où $x \notin y$ signifie $\neg(x \in y)$)

5.1.3 L'axiome de l'infini et les entiers

Pour introduire l'axiome suivant, il nous faut d'abord motiver l'utilisation de ses éléments constitutifs. L'ensemble mathématique le plus élémentaire que l'on est amené à étudier est certainement $\mathbb{N}$, l'ensemble des entiers naturels. Une formalisation habituelle de cet ensemble demande en général les trois constituants suivant:

- l'élément 0
- la fonction unaire S, correspondant à $n \mapsto n+1$
- le principe de récurrence, que l'on peut encoder dans les ensembles par le fait que si $F \subseteq \mathbb{N}$, $0 \in F$ et $\forall n, n \in F \implies S\ n \in F$ alors $F = \mathbb{N}$.

Chercher à définir $\mathbb{N}$ dans ZFC nous demande donc de définir ces éléments. Un candidat naturel à 0 est $\varnothing$: les deux sont les objets nuls par excellence, et $\varnothing$ est le premier ensemble que l'on peut construire. La question de savoir ce qu'est $S\ x$ pour un ensemble x est alors naturelle : pour cela, nous utilisons le codage de Von Neumann consistant à coder l'entier naturel n par $\{0, \dots, n-1\}$. En effet, cela nous offre une définition naturelle à la fonction S :

$$S\ x \triangleq x \cup \{x\}$$

Remarquons que l'on peut déjà construire tous les entiers que l'on souhaite : on peut construire 0 et itérer la fonction S. Malheureusement, rien ne nous dit que la collection $\{S^n\ 0 \mid n \in \mathbb{N}\}$ est bien elle-même un ensemble (où le $\mathbb{N}$ apparaissant dans la définition est l'ensemble des entiers naturels de notre méta-théorie). Pour palier ce problème, et pour éviter d'utiliser notre méta-théorie explicitement, on va à la place définir $\mathbb{N}$ comme le plus petit ensemble contenant 0 et stable par la fonction S. Il nous reste à savoir qu'un ensemble contenant 0 et stable par S existe bien.

Axiome 5.1.3.1 (Infini). *Il existe un ensemble contenant $\varnothing$ et stable par S :*

$$\exists x, \qquad \varnothing \in x \wedge \forall y, y \in x \implies S\ y \in x$$

Remarque 5.1.3.2. On peut en fait se passer de l'axiome de l'ensemble vide en prenant l'axiome de l'infini (et en adaptant sa définition pour ne pas appeler explicitement l'ensemble vide).

Définition 5.1.3.3 (Entiers naturels). Soit X l'ensemble défini par l'axiome de l'infini. On définit alors $\mathbb{N}$ comme

$$\mathbb{N} \triangleq \{x \in X \mid \forall Y, Y \subseteq X \wedge \varnothing \in Y \wedge (\forall a, a \in Y \implies S\ a \in Y) \implies x \in Y\}$$

On vérifie alors le principe de récurrence.

Théorème 5.1.3.4 (Récurrence). *Soit F une partie de $\mathbb{N}$ telle que $\varnothing \in F$ et $\forall n, n \in F \implies S\ n \in F$, alors $F = \mathbb{N}$.*

Démonstration. Comme F est une partie de $\mathbb{N}$, il nous suffit de montrer que $\mathbb{N} \subseteq F$. Par transitivité de l'inclusion, $F \subseteq X$ pour X l'ensemble à partir duquel $\mathbb{N}$ a été défini. On sait de plus que

$$x \in \mathbb{N} \implies \forall Y, Y \subseteq X \wedge \varnothing \in Y \wedge (\forall a, a \in Y \implies S\ a \in Y) \implies x \in Y$$

d'où, en spécialisant Y en F, et sachant que $\varnothing \in F$ et $\forall n, n \in F \implies S\ n \in F$, il vient que

$$x \in \mathbb{N} \implies x \in F$$

ce qui est exactement $\mathbb{N} \subseteq F$, d'où le résultat. □

Notation 5.1.3.5. A partir de maintenant, pour fluidifier l'écriture, on adoptera un style plus laxiste sur l'écriture des propositions. Par exemple on se permettra d'écrire $\forall x \in X, \psi$ pour $\forall x, x \in X \implies \psi$ et tous les légers abus de notations du même genre.

5.1.4 Axiome du choix et fonctions

Pour introduire l'axiome du choix, le mieux est de parler de ce qu'on voudrait faire mais ne peut pas faire sans lui. Prenons $X_1, \dots, X_n$ des ensembles finis non vides. On peut prouver qu'il existe $(x_1, \dots, x_n) \in X_1 \times \cdots \times X_n$ (en codant le produit itéré et les tuples de taille arbitraire). Pour cela, il suffit de raisonner par récurrence sur n : pour passer de l'étape n à l'étape $n+1$ il suffit de remarquer que X_{n+1} est non vide, et prendre alors un élément $x \in X_{n+1}$ pour l'ajouter au tuple déjà construit.

Le problème ici est que, si l'on peut obtenir ce résultat sur un nombre fini d'ensembles non vides, ce processus ne fonctionne plus lorsque l'on passe au cas infini. On ne peut pas faire une infinité de choix, puisque l'on n'utilise que des méthodes finies. C'est pourquoi l'axiome du choix existe : il permet de choisir dans des familles infinies d'ensembles non vides. Donnons d'abord la définition de cet axiome.

Axiome 5.1.4.1 (Choix). *Soit X un ensemble dont les éléments sont non vides et deux à deux disjoints. Alors il existe un ensemble C tel que pour tout $x \in X$, $C \cap x$ est un singleton. Écrit formellement :*

$$\forall X, \varnothing \notin X \wedge (\forall x, y \in X, x \cap y \neq \varnothing \implies x = y) \implies \exists C, \forall x \in X, \exists y, C \cap x = \{y\}$$

Il existe plusieurs reformulation naturelles de l'axiome du choix. Pour celles-ci, nous allons introduire le formalisme pour traiter des fonctions.

Définition 5.1.4.2 (Fonction). Soient X, Y deux ensembles. Une fonction $f : X \to Y$ est une partie $\Gamma_f \subseteq X \times X$ (appelée graphe de f) telle que le prédicat $(x, y) \in \Gamma$, défini pour $x \in X$ et $y \in Y$, est une relation fonctionnelle. On désigne par $\mathrm{Dom}(f)$ et $\mathrm{Im}(f)$ les ensembles associés à la relation fonctionnelle. On dit que f est partielle si $\mathrm{Dom}(f) \neq X$, et totale si $\mathrm{Dom}(f) = X$.

Pour deux ensembles X, Y, on note par $\mathrm{Funct}(X, Y)$ l'ensemble des fonctions de X dans Y :

$$\mathrm{Funct}(X, Y) \triangleq \{f : X \to Y\}$$

Exercice 5.1.4.3. Soient X et Y deux ensembles. Montrer que $\mathrm{Funct}(X, Y)$ est bien un ensemble.

Remarque 5.1.4.4. Dire que f est partielle, dans ce cadre, n'est pas bien défini. En effet, on peut toujours prendre une fonction $f : X \to Y$ et l'étendre en $f : S\ X \to Y$: la fonction est la même mais cette propriété change. Pour éviter cela, une fonction consistera souvent plutôt en un triplet (X, Y, Γ_f), mais nous nous concentrerons simplement sur Γ_f. Il y a donc un abus de notation ici lorsque nous construirons par exemple une fonction en construisant juste son graphe, mais celui-ci reste léger car on connait toujours grâce au contexte les deux autres composantes de la fonction.

Notation 5.1.4.5. Soit une fonction $f : X \to Y$. On s'autorise à ajouter f comme symbole de fonction dans notre langage, dont l'argument est un élément de X. Cela n'est pas autorisé par défaut, et nous devons donc le coder dans notre langage élémentaire (de la même manière que l'on convient qu'il est possible de réécrire $x \cup y$ uniquement grâce à $\in$ et $=$). Pour cela, on définit la traduction d'une phrase φ faisant potentiellement intervenir l'expression $f(x)$ en une phrase φ' sans $f(x)$, par induction sur la structure de φ :

- si $\varphi = R(f(x))$ où R est une proposition atomique (une relation avec potentiellement d'autres termes que $f(x)$), alors

$$\varphi' = \forall y \in Y, (x,y) \in \Gamma_f \implies R(y)$$

- si $\varphi = \neg\psi$, alors $\varphi' = \neg\psi'$.
- si $\varphi = \psi \vee \chi$, alors $\varphi' = \psi' \vee \chi'$.
- si $\varphi = \psi \wedge \chi$, alors $\varphi' = \psi' \wedge \chi'$.
- si $\varphi = \exists a, \psi$, alors $\varphi' = \exists a, \psi'$.
- si $\varphi = \forall a, \psi$, alors $\varphi' = \forall a, \psi'$.

On considère donc qu'introduire la notation $f(x)$ est une facilité d'écriture.

Remarque 5.1.4.6. Cet abus de notation est aussi possible dans le cas d'une relation fonctionnelle qui n'est pas donnée par une fonction (au sens ensembliste).

On définit de plus les notions d'injectivité, de surjectivité et de bijectivité.

Définition 5.1.4.7 (Injectivité, surjectivité, bijectivité). Une fonction $f : A \to B$ est dite injective si chaque élément $b \in B$ a au plus un antécédent, c'est-à-dire si la proposition suivante est vérifiée :

$$\forall x, y \in A, f(x) = f(y) \implies x = y$$

Une fonction $f : A \to B$ est dite surjective si tout élément $b \in B$ a au moins un antécédent, c'est-à-dire si la proposition suivante est vérifiée :

$$\forall y \in B, \exists x \in A, f(x) = y$$

Une fonction $f : A \to B$ est dite bijective si elle est à la fois injective et surjective.

Donnons un exercice pour manipuler la notion de fonction et celle de récurrence.

Exercice 5.1.4.8. Montrer que l'on peut définir des fonctions par récursion sur $\mathbb{N}$, c'est-à-dire montrer que si l'on a un ensemble X, un élément $x_0 \in X$ et une fonction $f_S : \mathbb{N} \times X \to X$ alors il existe une unique fonction $f : \mathbb{N} \to X$ telle que

$$\begin{cases} f(0) = x_0 \\ \forall n \in \mathbb{N}, f(S\ n) = f_S(n, f(n)) \end{cases}$$

L'exercice suivant permet de définir des fonctions depuis un produit cartésien en considérant seulement ses projections.

Exercice 5.1.4.9. Soient X, Y, Z des ensembles. Montrer qu'il existe une bijection entre $\mathrm{Funct}(X \times Y, Z)$ et $\mathrm{Funct}(X, \mathrm{Funct}(Y, Z))$ donnée par $f \mapsto (x \mapsto y \mapsto f(x,y))$.

Donnons maintenant la deuxième formulation de l'axiome du choix.

Proposition 5.1.4.10. *L'axiome du choix est équivalent à la proposition suivante :*

$$\forall x, \exists f : \mathcal{P}(x)\backslash\{\varnothing\} \to x, \forall y \in \mathcal{P}(x)\backslash\{\varnothing\}, f(y) \in y$$

Démonstration. A faire. □

Une autre formulation importante nécessite la notion de produit, que l'on peut simplement considérer comme une généralisation de l'opération $\times$.

Définition 5.1.4.11 (Produit cartésien quelconque). Soit X un ensemble, on définit $\prod X$ par

$$\prod X \triangleq \left\{ f : X \to \bigcup X \;\middle|\; \forall x \in X, f(x) \in x \right\}$$

Si $\{X_i\}_{i \in I}$ est une famille de fonctions (c'est-à-dire une fonction $g : i \mapsto X_i$) alors on considère que

$$\prod_{i \in I} X_i \triangleq \left\{ f : I \to \bigcup_{i \in I} X_i \;\middle|\; \forall i \in I, f(i) \in X_i \right\}$$

Proposition 5.1.4.12. *L'axiome du choix est équivalent à la proposition suivante :*

$$\forall x, (\forall y \in x, y \neq \varnothing) \implies \prod x \neq \varnothing$$

Démonstration. A FAIRE □

5.1.5 Axiome de fondation

Le dernier axiome de ZFC est l'axiome de fondation. Selon les auteurs, il est ou non ajouté à ZFC : ici, nous considérerons que ZFC le contient. Cet axiome empêche l'existence de cycles d'appartenance, par exemple d'ensemble x tel que $x \in x$. Formellement, l'axiome énonce que la relation $\in$ sur l'univers ensembliste $\mathcal{U}$ est bien fondée.

Axiome 5.1.5.1 (Fondation). *Pour tout ensemble x non vide, il existe $y \in x$ tel que $x \cap y = \varnothing$. Formellement, cela s'écrit :*

$$\forall x, x \neq \varnothing \implies \exists y \in x, x \cap y = \varnothing$$

Si l'on voit un ensemble x comme un arbre, où chaque nœud est un ensemble et où l'on place une arête entre x et y si $x \in y$, alors l'axiome de fondation stipule que pour chaque ensemble, l'arbre ainsi construit n'a pas de branche infinie (ni de cycle).

Remarquons que l'arbre que nous avons construit contient non seulement les éléments de x mais aussi les éléments des éléments de x, et ainsi de suite. Cette idée nous permet d'introduire la notion d'ensemble transitif, qui est un ensemble tel que l'arbre ainsi généré ne contient que des éléments de l'ensemble (c'est-à-dire que tous les éléments des éléments de x sont aussi éléments de x).

Définition 5.1.5.2 (Ensemble transitif). Soit un ensemble x. On dit que x est transitif si tout élément d'un élément de x est aussi élément de x. Formellement, cela s'écrit

$$\text{trans}(x) \triangleq \forall y \in x, \forall z \in y, z \in x$$

Remarquons qu'étant donné un ensemble, on peut toujours ajouter ses éléments, les éléments de ses éléments et ainsi de suite pour obtenir à la fin un ensemble transitif.

Proposition 5.1.5.3. *Soit x un ensemble. Alors il existe un plus petit ensemble* $\text{trcl}(x)$ *transitif contenant x.*

Démonstration. Pour construire $\text{trcl}(x)$, on construit la suite d'ensembles $(x_i)_{i \in \mathbb{N}}$ suivante :

- $x_0 = x$
- $x_{i+1} = \bigcup x_i$

L'ensemble $\mathrm{trcl}(x)$ est alors $\bigcup_{n\in\mathbb{N}} x_n$.

Vérifions que $\mathrm{trcl}(x)$ est bien transitif et contient x. Comme $x_O \subseteq \mathrm{trcl}(x)$ par définition de l'union, $\mathrm{trcl}(x)$ contient x. Soient $y \in \mathrm{trcl}(x)$ et $z \in y$. Par définition, il existe $n \in \mathbb{N}$ tel que $y \in x_n$. Alors $z \in x_{n+1}$ puisque $z \in y \in x_n$ et $x_{n+1} = \bigcup x_n$. Ainsi $z \in \mathrm{trcl}(x)$.

Soit un ensemble transitif y contenant x. Alors par récurrence sur n, $x_n \subseteq y$:

- par hypothèse, $x \subseteq y$.
- si $x_n \subseteq y$, alors soit $z \in x_{n+1}$, par définition on trouve $a \in x_n$ tel que $z \in a \in x_n$, et comme $x_n \subseteq y$, $a \in y$. Mais y est transitif, donc $z \in y$: on en déduit que $x_{n+1} \subseteq y$.

Par récurrence, on en déduit que $\forall n \in \mathbb{N}, x_n \subseteq y$, d'où $\mathrm{trcl}(x) \subseteq y$. □

Exercice 5.1.5.4. Montrer qu'un ensemble x est transitif si et seulement si la propriété suivante est vérifiée :

$$\forall y \in x, y \subseteq x$$

Les deux propriétés de bonne fondation de $\in$ et de transitivité mènent au lemme d'effondrement de Mostowski. Celui-ci montre que toute classe munie d'une relation se comportant suffisamment comme la relation $\in$ peut en fait se simuler par une unique classe transitive bien fondée où la relation est directement $\in$.

Lemme 5.1.5.5 (Effondrement de Mostowski). *Soit $\mathcal{M}$ une classe et R une relation vérifiant :*

- *R définit des ensembles, c'est-à-dire que pour tout $x \in \mathcal{M}$, la collection définie par $R^{-1}[x] = \{y \mid yRx\}$ est un ensemble.*
- *R est bien fondée, c'est-à-dire :*

$$\forall X \subseteq \mathcal{M}, \exists x \in X, R^{-1}[x] \cap X = \varnothing$$

- *R est extensionnelle, c'est-à-dire :*

$$\forall x, y \in \mathcal{M}, R^{-1}[x] = R^{-1}[y] \implies x = y$$

Alors il existe une unique classe $\mathcal{N}$ et un unique isomorphisme $\phi : \mathcal{M} \to \mathcal{N}$ tels que $(\mathcal{M}, R) \overset{\phi}{\cong} (\mathcal{N}, \in)$.

Remarque 5.1.5.6. Nous n'avons pas formellement défini ce qu'est un isomorphisme entre classes. Cette définition est assez naturelle : étant données deux classes $\mathcal{M}$ et $\mathcal{N}$ avec chacune une relation R (respectivement R'), un isomorphisme $\phi : (\mathcal{M}, R) \cong (\mathcal{N}, R')$ est une proposition ϕ à deux variables libres telle que, pour une proposition ψ définissant $\mathcal{M}$ et une proposition χ définissant $\mathcal{N}$, les propositions suivantes peuvent être prouvées :

- $\forall x, \forall y, \forall z, \psi(x) \wedge \chi(y) \wedge \chi(z) \wedge \phi(x, y) \wedge \phi(x, z) \implies y = z$
- $\forall x, \psi(x) \implies \exists y, \chi(y) \wedge \phi(x, y)$
- $\forall x, \forall y, \forall z, \psi(x) \wedge \psi(y) \wedge \chi(z) \wedge \phi(x, z) \wedge \phi(y, z) \implies x = y$
- $\forall y, \chi(y) \implies \exists x, \psi(x) \wedge \phi(x, y)$
- $\forall x, \forall x', \forall y, \forall y', \psi(x) \wedge \psi(x') \wedge \chi(y) \wedge \chi(y') \wedge \phi(x, y) \wedge \phi(x', y') \implies (xRx' \iff yR'y')$

Cependant, nous allons simplement travailler sur les classes comme sur les ensembles, puisque les manipulations syntaxiques correspondant à nous arguments ensemblistes habituels sont les mêmes que celles que nous allons faire sur nos classes. Remarquons simplement que ce lemme est, puisqu'il parle de classe, un schéma de lemmes paramétré par les propositions définissant $\mathcal{M}$ et R.

Démonstration. Montrons d'abord que si M_1 et M_2 sont deux classes transitives et que $\theta : M_1 \to M_2$ est un isomorphisme pour la rleation $\in$, alors θ est l'identité.

Par l'absurde, supposons que θ n'est pas l'identité. L'ensemble $C = \{x \mid \theta(x) \neq x\}$ n'est donc pas vide. Par l'axiome de fondation, on trouve un élément minimal $x \in C$, et soit $y = \theta(x)$. On peut alors remarquer que par minimalité de x, pour tous $z \in x$, $\theta(z) = z$, et comme θ est un isomorphisme on en déduit que $z \in y$. Ainsi $x \subseteq y$. Il nous reste à montrer que $y \subseteq x$: si $z \in y$, alors $z \in M_2$ comme M_2 est une classe transitive, donc on trouve $a \in M_1$ tel que $\theta(a) = z$, mais comme $\theta(a) \in \theta(x)$, on en déduit que $a \in x$, et puisque x est minimal dans C, $\theta(a) = a$, c'est-à-dire $z = a$. Donc $y \subseteq x$. On a donc une absurdité étant donné que $x = \theta(x)$: ainsi C est vide, donc θ est l'identité.

Reprenons maintenant notre classe $\mathcal{M}$. On définit alors la classe $\mathcal{N}$ par l'image de la fonction θ définie comme suit :

$$\begin{array}{rccl} \theta & : & \mathcal{M} & \longrightarrow & \mathcal{U} \\ & & x & \longmapsto & \{\theta(y) \mid y \in R^{-1}[x]\} \end{array}$$

c'est-à-dire que la relation θ est définie par la proposition

$$\theta(x, y) \triangleq y = \{\theta(z) \mid z \in R^{-1}[x]\}$$

où l'on prouve par l'absurde que cette proposition vérifie les propriétés d'un isomorphisme entre $(\mathcal{M}, R)$ et $(\mathcal{N}, \in)$, en utilisant un principe analogue à la preuve précédente. La classe $\mathcal{N}$ est définie par la proposition

$$y \in \mathcal{N} \triangleq \exists x \in \mathcal{M}, \theta(x, y)$$

L'unicité découle alors du fait que s'il existe une autre telle classe $\mathcal{N}'$, alors l'isomorphisme entre elle et $\mathcal{N}$ vérifie les hypothèses du résultat intermédiaire, et est donc l'identité. D'où $\mathcal{N} = \mathcal{N}'$. □

5.2 Construction des autres ensembles de nombres

Nous avons construit $\mathbb{N}$ directement grâce à l'axiome de l'infini, mais celui-ci n'est pas le seul ensemble de nombres important à construire. Dans cette section, nous allons vois la construction de $\mathbb{Z}, \mathbb{Q}$ et $\mathbb{R}$, ainsi que leurs propriétés essentielles. Nous allons commencer notre étude par celle de $\mathbb{N}$, puisque nous avons simplement défini cet ensemble et le principe de récurrence.

5.2.1 Les entiers naturels

Rappelons les points essentiels à propos de $\mathbb{N}$: cet ensemble contient 0, une fonction S unaire, et le principe de récurrence. De plus, on peut définir une fonction par récurrence. Nous pouvons donc définir les fonctions $+$ et $\times$.

Définition 5.2.1.1 (Addition, multiplication). On définit la fonction $+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ et la fonction $\times : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ par récurrence sur leur premier argument :

- $\forall n \in \mathbb{N}, n + 0 = n$
- $\forall n, m \in \mathbb{N}, n + (S\ m) = S(n + m)$
- $\forall n \in \mathbb{N}, n \times 0 = 0$
- $\forall n, m \in \mathbb{N}, n \times (S\ m) = n \times m + n$

Notation 5.2.1.2. On définit $1 = S\ 0$, $2 = S\ S\ 0$ et ainsi de suite. Plutôt que $S\ n$, nous utiliserons $n + 1$.

Propriété 5.2.1.3. *Les propriétés suivantes sont vérifiées :*

- 0 *est un élément neutre pour* $+$ *: pour tout* $n \in \mathbb{N}$, $0 + n = n$ *et* $n + 0 = n$.
- $+$ *est commutatif : pour tous* $n, m \in \mathbb{N}$, $n + m = m + n$.
- $+$ *est associatif : pour tous* $n, m, p \in \mathbb{N}$, $n + (m + p) = (n + m) + p$.
- $+$ *est régulier : pour tous* $n, p, m \in \mathbb{N}$, *si* $n + m = n + p$ *alors* $m = p$.
- *pour tous* $n, m \in \mathbb{N}$, *si* $n + m = 0$ *alors* $n = 0$ *et* $m = 0$.
- $+$ *distribue sur* $\times$ *: pour tous* $n, m, p \in \mathbb{N}$, $n \times (m + p) = n \times m + n \times p$ *et* $(n + m) \times p = n \times p + m \times p$.
- 1 *est un élément neutre pour* $\times$.
- 0 *est absorbant pour* $\times$ *: pour tout* $n \in \mathbb{N}$, $0 \times n = 0$.
- $\times$ *est commutatif.*
- $\times$ *est associatif.*
- *La structure est intègre : pour tous* $n, m \in \mathbb{N}$, *si* $n \times m = 0$ *alors* $n = 0$ *ou* $m = 0$.

Démonstration. A FAIRE □

La relation $\leq$ est aussi définissable grâce à notre langage ensembliste.

Définition 5.2.1.4 (Inégalité). On définit la relation $\leq \subseteq \mathbb{N} \times \mathbb{N}$ par :

$$n \leq m \triangleq \exists k \in \mathbb{N}, m = n + k$$

Propriété 5.2.1.5. *La relation* $\leq$ *est une relation d'ordre.*

Démonstration. Prouvons que la relation est réfléxive, antisymétrique et transitive :

- soit $n \in \mathbb{N}$, alors $n = n + 0$, d'où $n \leq n$.
- soient $n, m \in \mathbb{N}$, supposons que $n \leq m$ et $m \leq n$. On trouve alors k, k' tels que $n = m + k$ et $m = n + k'$. En substituant la deuxième inégalté dans la première, on obtient alors

$$\begin{aligned} n &= (n + k') + k \\ &= n + (k' + k) \\ n + 0 &= n + (k' + k) \\ 0 &= k + k' \qquad \text{par régularité de } + \end{aligned}$$

  d'où $0 = k$, donc $n = m$.
- soient $n, m, p \in \mathbb{N}$ tels que $n \leq m$ et $m \leq p$. On trouve donc k, k' tels que $m = n + k$ et $p = m + k'$. Par substitution, on en déduit que $p = (n + k) + k'$, et par transitivité $p = n + (k + k')$, d'où $n \leq p$.

Ainsi, $\leq$ est une relation d'ordre. □

Troisième partie

Théorie des modèles

Chapitre 6

Langage, structures et théories