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 8 février 2024 à 8 h 39 min 38 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

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 :

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

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*

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

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.

$$\begin{tikzcd}
X^{\alpha(c)} \arrow[r, "f^{\alpha(c)}"] \arrow[d, "c"] & Y^{\alpha(c)} \arrow[d, "f_c"] \\
X \arrow[r, "f"] & Y
\end{tikzcd}$$

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

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

- 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

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

  On définit alors

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

  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 $x_1, \ldots, x_p \in X_n$ tels que $x = (c, x_1, \ldots, x_p)$, on a alors

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

  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 $(c, x_1, \ldots, x_p)$ et $c(x_1, \ldots, x_p)$ 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 :

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

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

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

Exercice 1.1.2.3. Soit A un ensemble. Donner une signature définissant l'ensemble $\mathrm{BinTree}(A)$ 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 $h : \mathrm{BinTree}(A) \to \mathbb{N}$ 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 $h(\mathrm{nil}) = 0$.

Définir une fonction $|-| : \mathrm{BinTree}(A) \to \mathbb{N}$ 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 $x_1, \dots, x_p \in X^{\alpha(c)}$, la propriété*

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

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 $x_1, \dots, x_p \in X_n$ tels que $x = c(x_1, \dots, x_p)$. Par hypothèse de récurrence, on en déduit que pour tout $i \in \{1, \dots, p\}$, $x_i \in P$. Il vient donc, avec l'hypothèse du théorème sur P, que $c(x_1, \dots, x_p) \in P$.

  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 $\mathrm{pair}(n) \overset{\text{déf}}{=} \exists m \in \mathbb{N}, n = 2 \times m$.

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

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

Montrer que

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

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

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

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

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

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

Pour cela, il suffit de montrer que $f(\alpha) \subseteq \alpha$, mais comme α est une intersection, il suffit de montrer que pour tout $A \in \mathrm{prefix}(f), f(\alpha) \subseteq A$. Pour montrer cela, on remarque par transitivité de $\subseteq$ qu'il suffit de montrer que $f(\alpha) \subseteq f(A)$ pour tout $A \in \mathrm{prefix}(f)$, mais cela est direct en utilisant le fait que f est croissante et que pour tout $A \in \mathrm{prefix}(f)$, $\alpha \subseteq A$. 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 $\mathrm{prefix}(f)$: 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 $\mathrm{r}_1, \dots, \mathrm{r}_n$ 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 $\mathrm{r}_1, \dots, \mathrm{r}_n$, une règle r est dérivable s'il existe une dérivation de la conclusion de r utilisant les règles $\mathrm{r}_1, \dots, \mathrm{r}_n$ 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 $\mathrm{r}_1, \dots, \mathrm{r}_n$, 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 $\mathrm{r}_1, \dots, \mathrm{r}_n$, un ensemble E et R la relation définie sur R par ces règles. Si $R(x_1, \dots, x_p)$ pour un tuple $(x_1, \dots, x_p) \in E$, alors il existe $i \in \{1, \dots, n\}$ et des instances $P_1, \dots, P_k$ des prémisses de r_i tels que*

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

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.

Deuxième partie

Théorie des modèles

Chapitre 2

Langage, structures et théories