TECHNOSUP

Les filières technologiques des enseignements supérieurs

Informatique théorique

Logique et démonstration automatique

Une introduction à la logique propositionnelle et à la logique du premier ordre

Stéphane Devismes, Pascal Lafourcade et Michel Lévy

Maîtres de conférences, Université Joseph Fourier, Grenoble.

Avant propos

DURANT des siècles les mathématiciens ont raisonné *correctement sans connaître* la logique formelle. Quand ils ont réalisé que l'absence de formalisation pouvait conduire à des contradictions, ils ont commencé à formaliser le raisonnement mathématique, donnant ainsi naissance à la logique moderne. Dans ce livre, nous nous consacrons à l'étude de la logique *classique* (par opposition à la logique intuitionniste). La logique classique est la logique à deux valeurs de vérité. Elle permet de modéliser les circuits combinatoires, ce qui explique sa grande importance pratique. Il existe de nombreux autres domaines où la logique formelle intervient, ce qui explique pourquoi la recherche dans ce domaine est plus que jamais d'actualité. Les avancées actuelles de la logique sont principalement guidées par les deux objectifs suivants :

– Le besoin de prouver la correction automatique des programmes, particulièrement quand ces programmes sont utilisés dans des domaines où la sécurité est en jeu. Un exemple récent est l'utilisation d'outils de vérification formels dans la réalisation de la ligne automatique 14 du métro parisien lors du projet *Meteor*.
– L'ambition de représenter en machine des connaissances mathématiques. Nous noterons par exemple, le codage de problème de décision sous la forme d'une expression booléenne. Ce codage est ensuite traité par des programmes dédiés appelés *solveurs SAT*. Un autre exemple est la représentation de connaissance en intelligence artificielle pour la réalisation d'outils d'aide à la décision, par exemple les systèmes experts tels que *Mycin* dans les années 70.

Notre objectif est non seulement de faire découvrir la logique au lecteur, mais aussi de l'aider à développer sa rigueur mathématique et conforter son aptitude à raisonner. Ces compétences constituent pour nous la base d'un raisonnement scientifique de qualité. Nous avons choisi de présenter uniquement des résultats et algorithmes de la logique qui sont à la portée d'un novice et dont il existe une réalisation logicielle permettant d'appliquer automatiquement ces résultats sur des exemples. Plus précisément les compétences et connaissances que nous souhaitons transmettre sont les suivantes :

– *Comprendre un raisonnement* : être capable de déterminer si un raisonnement logique est correct ou non.
– *Raisonner* : savoir construire un raisonnement correct utilisant les outils de la logique propositionnelle et du premier ordre.
– *Modéliser et formaliser un problème* : à partir d'une description en français d'un problème proposer un ensemble de formules le caractérisant afin de pouvoir le résoudre.
– *Écrire une preuve rigoureuse* : être capable de rédiger une preuve formelle d'un résultat à l'aide de la logique.

La première partie de ce livre est consacrée à la *logique propositionnelle*. Plus précisément, dans le premier chapitre, nous présentons les définitions et résultats de base de la logique des prédicats (syntaxe, sémantique, substitution, algèbre de Boole ...). Nous montrons également dans ce chapitre comment utiliser la logique propositionnelle pour formaliser les circuits combinatoires. Dans le second chapitre, nous étudions un système de déduction à une seule règle : *la résolution*. Nous proposons un algorithme original basé sur la résolution : la *stratégie complète*. Enfin, nous présentons un autre algorithme basé sur la résolution : l'algorithme `DPLL`, proposé dans les années soixante par Martin Davis, Hilary Putnam, George Logemann et Donald Loveland [**?**, **?**]. Ce dernier est encore actuellement à la base des solveurs SAT complets. Enfin, dans le dernier chapitre de cette partie, nous étudions un système formel à plusieurs règles : la *déduction naturelle* introduite dans les années trente par Gerhard Gentzen [**?**, **?**]. Pour cette dernière, nous avons préféré suivre la manière dont les mathématiciens font des déductions plutôt que d'être fidèles à l'appellation déposée par le logicien Gerhard Gentzen. Cela nous a conduit à introduire un système formel avec une seule règle d'élimination des hypothèses. Pour cette raison, notre système formel est plus simple que la majorité des présentations de la déduction naturelle faite dans la littérature. À partir de notre système formel, nous avons réalisé un logiciel permettant de construire de manière automatique des preuves de formules propositionnelles en déduction naturelle [1].

1. Ce logiciel est accessible à l'adresse suivante : `http://teachinglogic.liglab.fr/DN/`

La seconde partie est consacrée à la logique du premier ordre. Le premier chapitre introduit la syntaxe, la sémantique, les définitions et résultats de base au premier ordre. Pour expliquer la notion de modèle en logique du premier ordre, nous montrons comment construire automatiquement des modèles finis de formules. Ceci permet aux personnes découvrant la logique d'appréhender le sens des quantificateurs et les relations entre la logique propositionnelle et la logique du premier ordre. Dans le second chapitre, nous présentons les bases de la démonstration automatique via la résolution du premier ordre. En particulier nous aborderons les notions d'*unification* et *skolémisation*. Nous présentons la skolémisation à l'aide d'étapes élémentaires sur les formules ce qui ne nous oblige pas à passer par des formes *prénexes*. Enfin, dans le dernier chapitre, nous généralisons au premier ordre la déduction naturelle étudié dans la première partie.

Nous tenons à souligner que tous les résultats de cet ouvrage sont prouvés, sauf un : la complétude de déduction naturelle au premier ordre. Cette preuve nous paraît hors de portée du public visé. Une preuve de complétude pour un système analogue est disponible dans le livre de Herbert B. Enderton [**?**]. Nous sommes convaincus que le fait de proposer des démonstrations utilisant les concepts et notions introduites uniquement dans ce livre renforce la compréhension du lecteur mais aussi accroît sa rigueur et ses facultés de raisonnement. De plus l'ensemble des notions introduites dans ce livre (déduction naturelle, résolution, algorithme DPLL, construction de modèles etc ...) sont confortées par l'utilisation de logiciels.

À la fin de chaque chapitre, nous proposons une série d'exercices portant sur le contenu du chapitre. Nous indiquons par des étoiles la difficulté des exercices proposés, plus il y a d'étoiles plus l'exercice est jugé difficile. Nous indiquons aussi par le symbole ⊶ les exercices qui complètent les preuves énoncées dans le même chapitre. Les exercices proposés sont de trois catégories :
- Des exercices de base : aidant le lecteur à se familiariser avec le vocabulaire et à manipuler les notions introduites.
- Des exercices techniques : applications directes ou immédiates des résultats énoncés.
- Des exercices de réflexion : permettant à l'apprenant à raisonner, démontrer, prouver des résultats à partir des techniques et notions apprises grâce aux deux autres types d'exercices.

Ceci offre aux novices comme aux lecteurs plus expérimentés la possibilité de tester leurs connaissances et leur compréhension des différentes notions présentées.

Ce livre s'adresse principalement aux étudiants de premier cycle préparant une licence scientifique, c'est-à-dire, des étudiants de niveaux L1, L2 ou L3. Il ne suppose pas de connaissances *a priori* ni en logique, ni en mathématiques, mais une pratique des mathématiques correspondant au baccalauréat série S. Il peut servir également aux étudiants de second cycle, préparant une maîtrise scientifique ou l'agrégation de Mathématiques, souhaitant actualiser leurs connaissances en logique.

Remerciements : Nous sommes très reconnaissant envers Nicolas Peltier, chargé de recherche au CNRS. Plusieurs preuves et exercices sont inspirés de ses travaux. Nous sommes aussi redevables à Sophie Quinton et Mathilde Duclos pour leurs contributions. Enfin nous remercions Fabien Eloy pour le temps passé à la relecture de ces quelques pages.

Table des matières

Chapitre IV

Logique du premier ordre

Sommaire

CE célèbre syllogisme [1] ne peut pas se formaliser en logique propositionnelle :

Tous les hommes sont mortels ;
or Socrate est un homme ;
donc Socrate est mortel.

Pour formaliser un tel raisonnement nous avons besoin d'enrichir la logique propositionnelle avec de nouveaux connecteurs appelés *quantificateurs*, cette logique étendue s'appelle logique du premier ordre. Elle permet de parler de structures comportant un seul domaine non vide (contrairement à la logique propositionelle, ce domaine peut avoir plus de deux valeurs), des fonctions et relations sur ce domaine. Le langage de cette logique comporte plusieurs catégories : les termes qui représentent les éléments du domaine ou des fonctions sur ces éléments, des relations qui relient des termes entre eux et les formules qui décrivent les interactions entre les relations grâce aux connecteurs et aux quantificateurs. Par exemple, la relation $mortel(x)$ désigne x est mortel, la relation $homme(x)$ signifie que x est un homme. De plus, $Socrate$ est une constante du domaine (c'est-à-dire, elle a pour valeur un élément du domaine) et $homme(Socrate)$ signifie que Socrate est un homme. Enfin, la formule $\forall x\ (homme(x) \Rightarrow mortel(x))$ indique que tous les hommes sont mortels. À partir de ces deux hypothèses, il est possible de conclure que Socrate est mortel ($mortel(Socrate)$) par différentes méthodes

1. La notion de syllogisme fut introduite par Aristote et signifie littéralement « parole (qui va) avec (une autre) », aussi connu en latin par *modus ponendo ponens* = « manière d'affirmer, d'établir en affirmant » et plus brièvement *modus ponens*.

que nous détaillerons dans cette seconde partie. Ce même raisonnement s'applique également pour prouver le syllogisme suivant :

Un cheval bon marché est rare.
Tout ce qui est rare est cher.
Donc un cheval bon marché est cher.

Malgré la conclusion qui semble contradictoire, le raisonnement est correct. Afin d'obtenir une contradiction il faut ajouter l'hypothèse suivante : $\forall x($ bonmarché $(x) \Leftrightarrow \neg$ cher $(x))$. Ceci traduit la relation communément admise entre la notion de bon marché et de cher, *i.e.*, tout ce qui est cher n'est pas bon marché et réciproquement. Avec cette hypothèse supplémentaire nous pouvons montrer que ce raisonnement est contradictoire, sans cette hypohèse le raisonnement est correct. Ceci élucide le paradoxe de ce syllogisme.

Notre objectif dans ce chapitre est d'introduire les concepts et notions élémentaires de la logique du premier ordre afin de pouvoir proposer des méthodes de raisonnement dans les chapitres suivants.

Plan : Nous commençons par décrire la syntaxe de la logique du premier ordre. Puis, nous définissons les notions de *libre* et *liée* qui sont essentielles pour pouvoir interpréter le sens des formules. Ensuite nous définissons le sens des formules que nous pouvons construire à partir de la syntaxe introduite précédemment. Enfin, nous énonçons des propriétés remarquables de la logique du premier ordre, qui pourront être utilisées dans les raisonnements.

1 Syntaxe

Nous ajoutons à la logique propositionnelle deux symboles : le symbole *existentiel* ($\exists$) et le symbole *universel* ($\forall$). Le symbole existentiel signifie qu'il existe un élément ayant une certaine propriété alors que le symbole universel permet de parler de tous les éléments ayant une propriété. Ces changements impliquent qu'un élément de ce langage peut désormais dépendre de plusieurs variables. Il faut donc rajouter dans la syntaxe un délimiteur de variables. Nous avons choisi classiquement la virgule. Ces changements introduisent de nouveaux symboles en plus des variables, ces symboles peuvent par exemple être des nombres sur lesquels nous pouvons créer des fonctions (comme l'addition) ou encore définir des relations (comme l'égalité). Toutes ces modifications rendent la présentation de la syntaxe légèrement plus subtile que celle de la logique propositionnelle.

1.1 Formules strictes

Pour écrire les formules de la logique du premier ordre nous étendons la syntaxe de la logique propositionnelle, ainsi nous disposons du vocabulaire suivant :
- Deux constantes propositionnelles : $\perp$ et $\top$ représentant respectivement le faux et le vrai.
- Variables : suite de lettres et de chiffres commençant par une des minuscules u, v, w, x, y, z.
- Connecteurs : $\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow$.
- Quantificateurs : $\forall$ le quantificateur universel et $\exists$ le quantificateur existentiel.
- Ponctuations : la virgule « , » et les parenthèses ouvrantes « (» et fermantes «) ».
- Symbole :
  - ordinaire : suite de lettres et de chiffres ne commençant pas par une des minuscules u, v, w, x, y, z.
  - spécial : $+, -, *, /, =, \neq, <, \leq, >, \geq$

L'ensemble des symboles spéciaux peut être augmenté de nouveaux éléments, suivant les domaines mathématiques étudiés, à condition de respecter la contrainte que ces nouveaux éléments ne soient pas dans les ensembles des constantes propositionnelles, des variables, des connecteurs, des quantificateurs, des ponctuations ou des symboles ordinaires.

Exemple 1.1 *Ci-dessous, nous illustrons ces notions :*
- $x, x1, x2, y$ *sont des variables,*
- *homme, parent, succ, 12, 24,* $f1$ *sont des symboles ordinaires, les symboles ordinaires représenteront des fonctions (constantes numériques ou fonctions à plusieurs arguments) ou des relations (variables propositionelles ou relations à plusieurs arguments).*
- $x = y$*,* $z > 3$ *sont des exemples d'application de symboles spéciaux.*

Nous définissons en toute généralité la notion de terme qui est essentielle en logique du premier ordre. Nous étendrons dans la suite cette notion en définissant ce qu'est un terme associé à un ensemble de symboles.

5 Exercices

Exercice 68 (Structure et variables libres)
Pour chaque formule ci-dessous, indiquer sa structure et ses variables libres.

1. $\forall x(P(x) \Rightarrow \exists y Q(x,y))$.
2. $\forall a \forall b(b \neq 0 \Rightarrow \exists q \exists r(a = b * q + r \wedge r < b))$ [4].
3. $Pair(x) \Leftrightarrow \exists y(x = 2 * y)$.
4. x *Divise* $y \Leftrightarrow \exists z(y = z * x)$.
5. $Premier(x) \Leftrightarrow \forall y(y \text{ Divise } x \Rightarrow y = 1 \vee y = x)$.

□

Exercice 69 (Formalisation, symbole de fonction et de relation) *Nous considérons* $\Sigma = \{f^{r2}, o^{r2}, c^{r2}, j^{r2}, r^{f0}, p^{f1}\}$ *la signature ayant la sémantique suivante :*

- $f(x,y) :=$ *x est frère de y.*
- $o(x,y) :=$ *x est l'oncle de y.*
- $c(x,y) :=$ *x est le cousin de y.*
- $j(x,y) :=$ *x est plus jeune que y.*
- *r est le diminutif de Robert.*
- $p(x)$ *est le père de x.*

Exprimer en logique du premier ordre et en utilisant la signature Σ *les phrases suivantes :*

1. *Tout frère du père de Robert est un oncle de Robert.*
2. *Si les pères de deux enfants sont des frères alors ces deux enfants sont des cousins.*
3. *Robert a un cousin plus jeune qu'un des frères de Robert.*

□

Exercice 70 (Formalisation) *Considérons la signature* $\Sigma = \{a^{f0}, f^{f0}, J^{r2}, G^{r2}\}$*, où les symboles ont le sens suivant :*

- $a :=$ *l'équipe d'Allemagne.*
- $f :=$ *l'équipe de France.*
- $J(x,y) :=$ *x a joué un match contre y.*
- $G(x,y) :=$ *x a gagné contre y.*

Exprimer en logique du premier ordre en utilisant la signature Σ *les assertions suivantes :*

1. *L'équipe de France a gagné un match et en a perdu un.*
2. *L'équipe de France et l'équipe d'Allemagne ont fait match nul.*
3. *Une équipe a gagné tous ses matchs.*
4. *Aucune équipe n'a perdu tous ses matchs.*
5. *Considérons l'assertion suivante : « Tous ceux qui ont joué contre une équipe qui a gagné tous ses matchs, ont gagné au moins un match ». Parmi les formules suivantes, lesquelles expriment la phrase ci-dessus, et lesquelles sont équivalentes ?*
   (a) $\forall x \exists y (J(x,y) \wedge \forall z(J(y,z) \Rightarrow G(y,z)) \Rightarrow \exists v G(x,v))$.
   (b) $\forall x(\exists y (J(x,y) \wedge \forall z(J(y,z) \Rightarrow G(y,z))) \Rightarrow \exists v G(x,v))$.
   (c) $\exists x(\forall y (J(x,y) \Rightarrow G(x,y)) \Rightarrow \forall z(J(x,z) \Rightarrow \exists v G(x,v)))$.
   (d) $\forall x \forall y (J(x,y) \wedge \forall z(J(y,z) \Rightarrow G(y,z)) \Rightarrow \exists v G(x,v))$.
   (e) $\forall x(\forall y (J(x,y) \wedge \forall z(J(y,z) \Rightarrow G(y,z))) \Rightarrow \exists v G(x,v))$.

□

Exercice 71 (Formaliser,)** *Soient les constantes s pour Serge, t pour Tobby et les symboles de relation* $A(x,y)$*, x aime y,* $C(x)$*, x est un chien,* $D(x)$*, x est un animal domestique,* $E(x)$*, x est un enfant,* $O(x)$*, x est un oiseau et* $P(x,y)$*, x a peur de y, donner la signature* Σ *associée à ces symboles. Formaliser en logique du premier ordre en utilisant* Σ *les énoncés suivants :*

1. *Les chiens et les oiseaux sont des animaux domestiques.*

4. Afin de respecter la notation usuelle pour la division euclidienne nous prenons exceptionnellement a, b, q et r comme nom de variable.