Licence informatique & vidéoludisme
Semestre 4

Histoire de l'informatique

Chapitre 5

Histoire de la logique

Pablo Rauzy <pr@up8.edu>
pablo.rauzy.name/teaching/hi

Histoire de la logique

La logique

- La *logique* est l'étude du raisonnement.
- Elle est à l'origine une branche de la philosophie (au sens grec).
- À partir du 19ème siècle se développe une approche mathématique de la logique.
- Depuis, son rapprochement avec l'informatique a fortement relancé son développement (ou c'est l'inverse ?).
- Les branches actuelles de la logiques sont
  - la théorie des ensembles,
  - la théorie de la démonstration,
  - la théorie des modèles,
  - la théorie de la calculabilité, et
  - la théorie des types.

La logique

- La *logique* est l'étude du raisonnement.
- Elle est à l'origine une branche de la philosophie (au sens grec).
- À partir du 19ème siècle se développe une approche mathématique de la logique.
- Depuis, son rapprochement avec l'informatique a fortement relancé son développement (ou c'est l'inverse ?).
- Les branches actuelles de la logiques sont
  - la théorie des ensembles,
  - la théorie de la démonstration,
  - la théorie des modèles,
  - la théorie de la calculabilité, et
  - la théorie des types.
- Attention : l'histoire de la logique est longue et compliquée.
- Aujourd'hui, on ne s'intéresse qu'à une vision partielle et subjective d'informaticien, parfois réductrice et certainement simpliste.
- Cette séance devrait peut-être plutôt s'appeler "un peu d'histoire *via* la logique".

Antiquité

- Les bases de la logique qui ont été formalisées par Aristote et Euclide perdurent jusqu'à aujourd'hui.
- Aristote cherche à analyser les formes de pensée permettant de construire un discours (*logos*) philosophique cohérent.
- Euclide, dans ses *Éléments*, cherche "juste" à fonder un corpus logique suffisant pour les mathématiques.
  - Les éléments fondamentaux qu'il appelle « notion ordinaire » ou postulat sont spécifiques aux mathématiques, par exemple « Un segment de droite peut être tracé en joignant deux points quelconques. ».
  - Ils ne peuvent prétendre à la généralité que couvre la logique d'Aristote, qui est essentiellement à finalité philosophique.

Période Classique

- À partir de la Renaissance, la science de manière générale subit un profond bouleversement : la révolution copernicienne.
- Les systèmes formels ne résistent pas au besoin du développement de la science.
- On fait alors une plus grande place à l'intuition et surtout à l'empirisme.
- Au siècle des Lumières, on appelle « raison éclairée de l'homme » l'approche qui construit sur l'expérience et l'induction.

Ère contemporaine

- Pendant la première moitié du 19ème siècle, la logique, héritée de la Grèce antique, est vue comme un outil philosophique.
- C'est durant le 19ème siècle que naît vraiment la logique mathématique.
- D'un côté il y a une volonté (Frege, Russel, Peano, Zermalo, Hilbert, …) de donner une fondation axiomatique aux mathématiques, de l'autre Boole introduit des structures algébriques permettant de définir un *calcul de vérité*.
- L'idée commence alors à s'installer que *le langage mathématique peut se définir mathématiquement, et donc être un objet d'étude des mathématiques*.
- Apparition du dualisme *syntaxe* / *sémantique*.

Algèbre de Boole

- En 1847, Boole publie son algèbre.
- La même année, De Morgan publie ses lois.
- La logique devient une branche à part entière des mathématiques.

Idéographie

- En 1879, Frege publie *Begriffsschrift* (*Idéographie*), un langage entièrement formalisé qui a pour but de représenter de manière parfaite la logique mathématique.
- Cette clarification permet de mettre en avant les trois caractéristiques qu'une théorie mathématique devrait avoir :
  - la cohérence,
  - la complétude, et
  - la décidabilité.

Théorie naïve des ensembles

- Au début des années 1880, Cantor introduit la *théorie des ensembles*.
- L'idée fondamentale est de définir l'*équipotence*.
  - Deux ensembles sont équipotents lorsqu'il existe une bijection entre eux.
  - Cette notion permet de définir la cardinalité, c'est-à-dire le nombre d'éléments d'un ensemble, qu'il soit fini ou infini.
- La théorie de Cantor met en avant le cas des ensembles infinis, objets aux propriétés particulières qui demandent une nouvelle approche.
- Cantor a approfondi la théorie et a construit des hiérarchies infinies d'ensembles infinis : les nombres ordinaux et les nombres cardinaux.
- La théorie de Cantor est considérée comme « naïve » parce qu'elle n'emploie pas encore une axiomatique précise, et parce que pour lui il n'y avait qu'une seule théorie des ensembles, un seul univers ensembliste attendu.

Le programme de Hilbert

- En 1898, Hilbert propose de réduire l'arithmétique à la logique. Son but est de montrer que les nombres sont des objets qui se déduisent d'un système d'axiomes non-contradictoires.
- En 1900, Hilbert présente son fameux programme en 23 questions pour le siècle à venir, dont la deuxième est :

  « *Peut-on prouver la cohérence de l'arithmétique ? En d'autres termes, peut-on démontrer que les axiomes de l'arithmétique ne sont pas contradictoires et, subséquemment, sont-ils indépendants ?* »

Paradoxe de Russel

- Le paradoxe de Russel (1902) est l'instance la plus facilement compréhensible de la série de problèmes et paradoxes que pose la théorie naïve des ensembles (et similairement, l'idéographie).
- Le paradoxe de Russel consiste en la construction de l'ensemble des ensembles qui n'appartiennent pas à eux-mêmes.
- Mathématiquement, si on pose $P = \{E | E \notin E\}$, on a immédiatement que $P \in P \iff P \notin P$.
- De manière plus imagée, il s'agit du paradoxe du barbier qui rase tous les hommes qui ne se rasent pas eux-mêmes, qu'on peut résoudre en affirmant qu'un tel barbier ne peut exister.

Paradoxe de Russel

- Le paradoxe de Russel (1902) est l'instance la plus facilement compréhensible de la série de problèmes et paradoxes que pose la théorie naïve des ensembles (et similairement, l'idéographie).
- Le paradoxe de Russel consiste en la construction de l'ensemble des ensembles qui n'appartiennent pas à eux-mêmes.
- Mathématiquement, si on pose $P = \{E | E \notin E\}$, on a immédiatement que $P \in P \iff P \notin P$.
- De manière plus imagée, il s'agit du paradoxe du barbier qui rase tous les hommes qui ne se rasent pas eux-mêmes, qu'on peut résoudre en affirmant qu'un tel barbier ne peut exister.
- En théorie des ensembles c'est plus compliqué.
  - Il est naturel de considérer que toute propriété exprimable définit un ensemble : celui des objets qui satisfont la propriété (c'est le *principe de compréhension*).
  - Mais si on admet ce principe, alors on doit admettre l'existence de l'ensemble paradoxal.

Diagonalisation

- La démonstration du paradoxe de Russel repose sur une *diagonalisation*.
  - Il s'agit d'un genre de démonstration par l'absurde.
  - Le principe est d'avoir une autoréférence et une négation (« Je mens. »).
- Elle est de fait très semblable à la démonstration du *théorème de Cantor* (1891).
  - Le théorème de Cantor dit que le cardinal d'un ensemble E est toujours strictement inférieur au cardinal de l'ensemble de ses parties $P(E)$.
  - L'idée est de montrer qu'il n'existe pas de bijection $f : E \to P(E)$.
  - Pour cela, on construit un ensemble $D = \{x \in E | x \notin f(x)\}$ qui n'a pas d'antécédent (c'est à dire qu'il y a pas d'élément $e \in E$ tel que $D = f(e)$.
  - En effet, si on suppose le contraire, soit e est dans D et donc n'appartient pas à D ; soit e n'est pas dans D et donc appartient à D.
  - Il n'existe donc aucune fonction surjective de E dans $P(E)$.
- La conséquence de ce théorème est qu'il n'existe pas de plus grand cardinal (paradoxe de Cantor).

La crise des fondements

- Ces histoires de paradoxes dans ce qui se veut être les fondations des mathématiques provoquent le début de ce qu'on appelle la *crise des fondements*.
- En effet, la cohérence des mathématiques est sous la menace du *principe d'explosion* (« *ex falso quodlibet* »).
- Du coup, plus aucune prédiction ne peut être réalisée par le calcul de l'évolution d'un système physique puisque n'importe quel calcul contradictoire et incompatible devenait possible.

Théorie des types

- En 1908, Russel présente la *théorie des types* comme une solution à son paradoxe (une autre solution basée sur la restriction du principe de compréhension est due à Zermalo, la même année).
- Dans cette théorie, les ensembles sont hiérarchisés par leur type : les éléments d'un ensemble ne peuvent être que des objets (possiblement des ensembles) de type strictement inférieur à celui de l'ensemble initial.
- Cela abouti aux *Principia Mathematica*, de Russell et Whitehead.
- Cela résout effectivement le paradoxe en limitant les propriétés exprimables, et en fait un candidat à l'axiomatisation des mathématiques.
- Il reste à savoir ce qui en est de la cohérence, de la complétude, et de la décidabilité.

Théorie des ensembles

- La même année, en 1908, Zermalo construit un système d'axiomes pour la théorie des ensembles.
- En dehors de l'*axiome d'extensionnalité*, on peut voir ces axiomes comme une restriction de la version contradictoire du *schéma d'axiomes de compréhension* aux cas particuliers utiles, qui ne permettent pas de dériver les paradoxes.
- Cela abouti plus tard au système ZF(C) (Zermalo-Fraenkel et axiome du choix), encore le plus prisé de nos jours.
- Cela résout effectivement le paradoxe en limitant les propriétés exprimables, et en fait un candidat à l'axiomatisation des mathématiques.
- À nouveaux, il reste à savoir ce qui en est de la cohérence, de la complétude, et de la décidabilité.

Entscheidungsproblem

- En 1922, Hilbert précise son projet de formalisation de l'arithmétique en posant le Entscheidungsproblem (le *problème de la décision*).
- Il demande si il existe une procédure (un algorithme) permettant de vérifier si une expression formelle peut se déduire d'un système d'axiomes donnés.
- Cela aboutit en 1928 à un ambitieux programme de recherche qui s'articule autour de trois questions :
  - Les mathématiques sont-elles complètes ?
  - Les mathématiques sont-elles cohérentes ?
  - Les mathématiques sont-elles décidables ?

Théorèmes de Gödel

- En 1931, Gödel publie un article dans lequel il énonce ses deux théorèmes d'incomplétude, qui viennent répondre par la négative aux questions de Hilbert.
- Le premier théorème dit :

  « *Dans n'importe quelle théorie récursivement axiomatisable, cohérente et capable de "formaliser l'arithmétique", on peut construire un énoncé arithmétique qui ne peut être ni prouvé ni réfuté dans cette théorie.* »
- Le second théorème dit :

  « *Si T est une théorie cohérente qui satisfait des hypothèses analogues, la cohérence de T, qui peut s'exprimer dans la théorie T, n'est pas démontrable dans T.* »
- Ces deux théorèmes ont été démontrés pour l'arithmétique de Peano et donc pour les théories plus fortes que celle-ci, en particulier les théories destinées à fonder les mathématiques, telles que la théorie des ensembles ou les *Principia Mathematica*.

Liens avec l'informatique

- Les liens entre la logique et l'informatique sont très forts.
- Ils existent notamment au travers du λ-calcul.

λ-calcul

- Le λ-calcul est un système formel inventé par Church dans les années 1930.
- Il s'agit du premier formalisme permettant de caractériser les *fonctions récursives*, il a donc une grande importance en calculabilité, à l'égal des machines de Turing.

- L'idée de base du λ-calcul est que *tout est fonction*.
- Ces fonctions peuvent contenir des fonctions qui ne sont pas prédéfinie et qui sont alors des variables.
- La seule chose qu'on peut faire avec ces fonctions est de les *appliquer* à des valeurs (qui sont elles-mêmes des fonctions).

- La syntaxe du λ-calcul est minimaliste.
- Si E est une expression du λ-calcul (un λ-terme) alors E est soit :
  - une *variable* : x, y, ... sont des λ-termes ;
  - une *application* : uv est un λ-terme si u et v sont des λ-termes ;
  - une *abstraction* : $\lambda x.v$ est un λ-terme si x est une variable et v est un λ-terme.
- L'application uv peut être vue comme l'application d'une fonction u à la valeur v.
- L'abstraction $\lambda x.v$ peut être vue comme la fonction qui à x associe v (où v contient généralement des occurrences de x).

Variables libres et variables liées

- Comme avec les quantificateurs classiques ∃ et ∀, les variables sont *liées* par le λ.
- Les variables qui ne sont pas liées sont dites *libres*.
- Tout comme $\forall x, P(x)$ est équivalent à $\forall y, P(y)$, les deux termes $\lambda x.P$ et $\lambda y.P[x/y]$ sont équivalent.
- Exemples :
  - Dans $\lambda x.xy$, x est liée et y est libre, du coup il est équivalent à $\lambda z.zy$ mais pas à $\lambda x.xz$.
  - Le terme $\lambda b.\lambda n.banane$ est équivalent à $\lambda p.\lambda t.patate$.
- L'opération de renommage des variables liées, qui est parfois utiles pour clarifier un terme, s'appelle la *α-conversion*.

- Le calcul se fait par *réduction*, défini comme une réécriture :
  - $(\lambda x.v)y \mapsto v[x/y]$.
- Exemples :
  - $(\lambda x.xy)a \mapsto ay$.
  - $(\lambda b.\lambda n.banane)p \mapsto \lambda n.panane$.
  - $(\lambda n.panane)t \mapsto patate$.
- On appelle cette opération la *β-réduction*.

λ-calcul simplement typé

- Sans typage, le λ-calcul peut exprimer des paradoxes, ou du moins des calculs qui ne terminent pas, par exemple $\lambda x.xx$ appliqué à lui même.
- En donnant au λ-calcul un typage simple, on peut réduire son expressivité.
- Les types peuvent être :
  - un type de base : ι,
  - si τ_1 et τ_2 sont des types, alors $\tau_1 \rightarrow \tau_2$ est un type.

Jugement de typage

- On appelle un *jugement de typage* les règles qui définissent le type d'un λ-terme en fonction de sa structure et des types des termes qui le compose.
- Dans le cas du λ-calcul simplement typé, ces règles sont :

$$\frac{}{\Gamma \vdash x : \tau \quad (\text{si } (x, \tau) \in \Gamma)} \;(\mathrm{ax})$$

$$\frac{\Gamma \vdash u : \tau_1 \to \tau_2 \quad \Gamma \vdash v : \tau_1}{\Gamma \vdash uv : \tau_2} \;(\mathrm{app})$$

$$\frac{\Gamma \cup \{(x, \tau_1)\} \vdash v : \tau_2}{\Gamma \vdash \lambda x.v : \tau_1 \to \tau_2} \;(\mathrm{abs})$$

- La règle ax (axiome) se lit comme "dans l'environnement Γ, le terme x a le type τ".
- La règle app (application) si lit comme "si dans l'environnement Γ le terme u est de type $\tau_1 \to \tau_2$ et que dans ce même environnement Γ le terme v est de type τ_1, alors le type du terme uv est τ_2".
- La règle abs (abstraction) si lit comme "si dans l'environnement Γ, auquel on ajoute la règle que le terme x est de type τ_1 le terme v est de type τ_2 alors le type du terme $\lambda x.v$ est $\tau_1 \to \tau_2$".

Termes correctement typés

- Si on donne (comme axiome) des types de base à nos variables on peut donc vérifier qu'un terme est correctement typé.
- Pour cela il faut qu'en partant du terme, on puisse remonter aux axiomes en utilisant seulement les règles de jugement de typage.
- On peut aisément comprendre qu'un terme tel que $\lambda x.xx$ n'est pas correctement typable.

Isomorphisme de Curry-Howard

- L'isomorphisme de Curry-Howard définit une *correspondance entre preuves et programmes*.
- Une façon simple d'en avoir un aperçu est de regarder la règle app du λ-calcul simplement typé et de constater qu'il correspond tout à fait à la règle de déduction du *modus ponens*.
- Le *modus ponens* est la règle de déduction qui dit que si dans une théorie on a comme axiomes ou théorèmes
  - A, et
  - $A \implies B$,

  alors on peut en déduire que B est aussi un théorème de cette théorie.

Logicomix

- Je vous recommande fortement la lecture de *Logicomix*.