Semi-Decidability of First Order Logic

Marco Piastra

Decidability and automation of L_{FO}

- L_{FO} is <u>not</u> decidable

  No Turing machine can tell whether $\Gamma \models \varphi$

  Are there any hopes for automating the calculus?

- L_{FO} is ***semi-decidable*** **(Herbrand, 1930)**

  In general, Turing machine can tell (in *finite* time) that

  $$\Gamma \models \varphi$$

  ... but <u>not</u> that

  $$\Gamma \not\models \varphi$$

  In other words, the above Turing machine, when facing the problem "$\Gamma \models \varphi$?" :

  1) it will terminate with success if $\Gamma \models \varphi$
  2) it *<u>might</u>* diverge if $\Gamma \not\models \varphi$

Herbrand's System

Given a universal sentence of the form:

$\forall x_1 \forall x_2 \ldots \forall x_n \varphi$ (where φ does not contain quantifiers)

the ***Herbrand's System*** is the set (possibly *infinite*) of *ground* wffs generated by replacing the variables

$\varphi[x_1/t_1, x_2/t_2 \ldots x_n/t_n]$

A term (or a wff) is *ground* does not contain variables

with all possible combinations of *ground* terms $<t_1, t_2 \ldots t_n>$ of the *signature* Σ

Examples:

$H(\forall x\, P(x) \to Q(x))) = \{P(f(a)) \to Q(f(a)), P(g(a, b)) \to Q(g(a, b)), \ldots \}$

$H(\forall x\, \forall y\, R(x, y)) = \{R(f(a), f(a)), R(g(a, b), f(a)), R(f(a), g(a, b)), \ldots \}$

- ***Herbrand's System*** of a theory

Given a theory Φ of universal sentences, the Herbrand's system $H(\Phi)$ is the union of all Herbrand's systems of the sentences in Φ

Example:

$\Phi = \{\varphi, \psi, \chi\}$

$H(\Phi) = H(\psi) \cup H(\varphi) \cup H(\chi)$

Herbrand's Theorem

- **Herbrand's Theorem**

  Given a theory of universal sentences Φ,
  $H(\Phi)$ has a model iff Φ has a model

  ... but what is the utility of that?
  $H(\Phi)$ may well be infinite even when Φ is finite,
  Furthermore, the theorem applies only to sets of <u>universal</u> sentences…

Prenex normal form (PNF)

Any wff φ can be transformed into an equivalent formula of the form

$Q_1 x_1 Q_2 x_2 \ldots Q_n x_n \ \psi$ (ψ is called the ***matrix***)

where Q_i is either $\forall$ or $\exists$ and ψ does not contain quantifiers

1) Replace $\rightarrow$ and $\leftrightarrow$:

$$\varphi \leftrightarrow \psi \quad \Leftrightarrow \quad (\varphi \rightarrow \psi) \land (\psi \rightarrow \varphi)$$

$$\varphi \rightarrow \psi \quad \Leftrightarrow \quad (\neg \varphi \lor \psi)$$

2) Push negation $\neg$ inwards, as much as possible:

$$\neg(\varphi \land \psi) \Leftrightarrow (\neg \varphi \lor \neg \psi)$$

$$\neg(\varphi \lor \psi) \Leftrightarrow (\neg \varphi \land \neg \psi)$$

$$\neg \neg \varphi \Leftrightarrow \varphi$$

$$\neg \forall x \, \varphi \Leftrightarrow \exists x \, \neg \varphi$$

$$\neg \exists x \, \varphi \Leftrightarrow \forall x \, \neg \varphi$$

3) Move all quantifiers outwards, respecting order

CAUTION: *variables MUST be renamed - when required - to avoid name clashes*

Prenex normal form (PNF)

Any wff φ can be transformed into an equivalent formula of the form

$$Q_1 x_1 Q_2 x_2 \dots Q_n x_n \, \psi \qquad (\psi \text{ is called the } \textbf{\textit{matrix}})$$

where Q_i is either $\forall$ or $\exists$ and ψ does not contain quantifiers

Examples:

$\exists y \, (P(y) \rightarrow \forall x \, P(x))$	
$\exists y \forall x \, (\neg P(y) \lor P(x))$	PNF
$\exists y \, (\forall x \, P(x) \rightarrow P(y))$	
$\exists y \, (\neg \forall x \, P(x) \lor P(y))$	
$\exists y \exists x \, (\neg P(x) \lor P(y))$	PNF
$\forall x \exists y \, (Q(x,y) \rightarrow P(y)) \land \neg \forall x \, P(x)$	
$\forall x \exists y \, (\neg Q(x,y) \lor P(y)) \land \exists x \, \neg P(x)$	
$\forall x \exists y \, (\neg Q(x,y) \lor P(y)) \land \exists z \, \neg P(z)$	(renaming variable)
$\forall x \exists y \exists z \, ((\neg Q(x,y) \lor P(y)) \land \neg P(z))$	PNF

Skolemization

In a sentence in PNF, existential quantifiers can be eliminated by extending the *signature* Σ of the *language*

Consider a sentence in PNF $Q_1 x_1 Q_2 x_2 \dots Q_n x_n \psi$

From left to right, for each $Q_i x_i$ of type $\exists x_i$:

- Apply to ψ the *substitution* $[x_i/k(x_1, \dots, x_j)]$
  where k is a <u>new</u> function and $x_1, \dots, x_j$ are the variables of j the universal quantifiers that come *before* $\exists x_i$ (k is an individual constant if $j = 0$)
- $\exists x_i$ is simply removed

Examples:

$\exists y \, \forall x \, (\neg P(y) \lor P(x))$

$\quad \forall x \, (\neg P(k) \lor P(x))$ (k Skolem's constant)

$\forall x \, \exists y \, \exists z \, ((\neg Q(x,y) \lor P(y)) \land \neg P(z))$

$\quad \forall x \, ((\neg Q(x, k(x)) \lor P(k(x))) \land \neg P(m(x)))$ ($k/1$ and $m/1$ Skolem's functions)

- Theorem

For any sentence φ

φ has a model iff $sko(\varphi)$ (i.e. Skolemization of φ) has a model

Semi-decidability of L_{PO}

- Corollary of Herbrand's theorem

  For any set of sentences Γ and sentence φ
  these three statements are equivalent:

  - $\Gamma \models \varphi$
  - $\Gamma \cup \{\neg\varphi\}$ is not satisfiable (= it has no model)
  - There exists a ***finite*** subset
    of $\mathrm{H}(sko(\Gamma \cup \{\neg\varphi\}))$ (= Herbrand's system of the Skolemitazion of $\Gamma \cup \{\neg\varphi\}$)
    that is ***inconsistent***

  Therefore:

  When $\Gamma \models \varphi$, a procedure that generates the finite *subsets* of $\mathrm{H}(sko(\Gamma \cup \{\neg\varphi\}))$
  will certainly discover a contradiction (*in finite time*)