Lecture 11

Applications of Herbrand's theorem

Ground resolution proofs, semi-decidability of validity, undecidability of validity

Dr Christoph Haase
University of Oxford
(with small changes by Javier Esparza)

Recap

Theorem (Herbrand's theorem)

A closed formula in Skolem form is satisfiable if and only if it has a Herbrand model.

Recap

Theorem (Herbrand's theorem)

A closed formula in Skolem form is satisfiable if and only if it has a Herbrand model.

Theorem

A closed formula in Skolem form is satisfiable iff its Herbrand expansion is satisfiable when considered as a set of propositional formulas.

Recap

Theorem (Herbrand's theorem)

A closed formula in Skolem form is satisfiable if and only if it has a Herbrand model.

Theorem

A closed formula in Skolem form is satisfiable iff its Herbrand expansion is satisfiable when considered as a set of propositional formulas.

Theorem (Ground resolution theorem)

A closed formula in Skolem form is unsatisfiable if and only if there is a propositional resolution proof of $\square$ *from its Herbrand expansion (in clause form).*

Generalisation of the ground resolution theorem

Theorem

Let $F_1, \ldots, F_n$ be closed rectified formulas in prenex form with Skolem forms $G_1, \ldots, G_n$. Assume each G_i is obtained using different Skolem functions and constants. Then

$$F_1 \wedge F_2 \wedge \cdots \wedge F_n \textit{ is satisfiable}$$
$$\textit{iff } G_1 \wedge G_2 \wedge \cdots \wedge G_n \textit{ is satisfiable}$$

Generalisation of the ground resolution theorem

Theorem

Let $F_1, \ldots, F_n$ be closed rectified formulas in prenex form with Skolem forms $G_1, \ldots, G_n$. Assume each G_i is obtained using different Skolem functions and constants. Then

$$F_1 \wedge F_2 \wedge \cdots \wedge F_n \text{ is satisfiable}$$
$$\text{iff } G_1 \wedge G_2 \wedge \cdots \wedge G_n \text{ is satisfiable}$$

Theorem (Ground resolution theorem)

Let $F_1, \ldots, F_n$ be closed formulas in Skolem form whose respective matrices $F_1^*, F_2^*, \ldots, F_n^*$ are in CNF. Then $F_1 \wedge F_2 \wedge \cdots \wedge F_n$ is unsatisfiable if and only if there is a propositional resolution proof of $\Box$ starting from the set of ground instances of clauses from $F_1^*, \ldots, F_n^*$.

An example

An example

Example

Consider the following hypothetical scenario:

(a) Everyone at Oriel is lazy, a rower or drunk.

(b) All rowers are lazy.

(c) Someone at Oriel is not drunk.

(d) Someone at Oriel is lazy.

Show that (a), (b) and (c) together entail (d).

An example

Translation into first-order logic:

An example

Translation into first-order logic:

$$F_1 := \forall x(O(x) \rightarrow (L(x) \vee R(x) \vee D(x)))$$
$$F_2 := \forall x(R(x) \rightarrow L(x))$$
$$F_3 := \exists x(O(x) \wedge \neg D(x))$$
$$F_4 := \neg\exists x(O(x) \wedge L(x))$$

Transformation into CNF Skolem form:

An example

Translation into first-order logic:

$$F_1 := \forall x(O(x) \to (L(x) \vee R(x) \vee D(x)))$$
$$F_2 := \forall x(R(x) \to L(x))$$
$$F_3 := \exists x(O(x) \wedge \neg D(x))$$
$$F_4 := \neg\exists x(O(x) \wedge L(x))$$

Transformation into CNF Skolem form:

$$G_1 := \forall x(\neg O(x) \vee L(x) \vee R(x) \vee D(x))$$
$$G_2 := \forall x(\neg R(x) \vee L(x))$$
$$G_3 := O(a) \wedge \neg D(a)$$
$$G_4 := \forall x(\neg O(x) \vee \neg L(x))$$

Resolution proof for the example

$$
\begin{aligned}
G_1 &:= \forall x(\neg O(x) \vee L(x) \vee R(x) \vee D(x)) \\
G_2 &:= \forall x(\neg R(x) \vee L(x)) \\
G_3 &:= O(a) \wedge \neg D(a) \\
G_4 &:= \forall x(\neg O(x) \vee \neg L(x))
\end{aligned}
$$

$$
\dfrac{\dfrac{\dfrac{\dfrac{\{\neg R(a), L(a)\} \qquad \{\neg O(a), L(a), R(a), D(a)\}}{\{L(a), \neg O(a), D(a)\}} \qquad \{\neg O(a), \neg L(a)\}}{\{\neg O(a), D(a)\}} \qquad \{\neg D(a)\}}{\{\neg O(a)\}} \qquad \{O(a)\}}{\square}
$$

Another example

Show that the following formula is valid:

$$F = \forall x \exists y (P(x) \to Q(y)) \to \exists y \forall x (P(x) \to Q(y))$$

Another example

Show that the following formula is valid:

$$F = \forall x \exists y (P(x) \to Q(y)) \to \exists y \forall x (P(x) \to Q(y))$$

F is valid if and only if F is unsatisfiable:

$$\neg F \equiv \forall x \exists y (P(x) \to Q(y)) \wedge \neg \exists y \forall x (P(x) \to Q(y))$$

Another example

Show that the following formula is valid:

$$F = \forall x \exists y (P(x) \rightarrow Q(y)) \rightarrow \exists y \forall x (P(x) \rightarrow Q(y))$$

F is valid if and only if F is unsatisfiable:

$$\neg F \equiv \forall x \exists y (P(x) \rightarrow Q(y)) \wedge \neg \exists y \forall x (P(x) \rightarrow Q(y))$$

Skolemise:

$$F_1 = \forall x (\neg P(x) \vee Q(f(x)) \qquad F_2 = \forall y (P(g(y)) \wedge \neg Q(y))$$

Another example

Show that the following formula is valid:

$$F = \forall x \exists y (P(x) \to Q(y)) \to \exists y \forall x (P(x) \to Q(y))$$

F is valid if and only if F is unsatisfiable:

$$\neg F \equiv \forall x \exists y (P(x) \to Q(y)) \land \neg \exists y \forall x (P(x) \to Q(y))$$

Skolemise:

$$F_1 = \forall x(\neg P(x) \lor Q(f(x)) \qquad F_2 = \forall y(P(g(y)) \land \neg Q(y))$$

No ground terms due to lack of constant symbols, introduce some constant symbol a

$$\dfrac{\dfrac{\{P(g(a))\} \qquad \{\neg P(g(a)), Q(f(g(a)))\}}{\{Q(f(g(a)))\}} \qquad \{\neg Q(f(g(a)))\}}{\square}$$

Semi-decidability of validity

Theorem

Validity of first-order logic is semi-decidable.

Semi-decidability of validity

Theorem

Validity of first-order logic is semi-decidable.

Semi-Decision Procedure for Validity
Input: Closed formula F
Output: Either that F is valid or compute forever

Compute a Skolem-form formula G equisatisfiable with $\neg F$
Let $G_1, G_2, \ldots$ be an enumeration of the Herbrand expansion $E(G)$
for $n = 1$ to ∞ **do**
begin
  if $\square \in \mathrm{Res}^*(G_1 \cup \ldots \cup G_n)$ **then** stop and output "F is valid"
end

How to show undecidability

Principle:

- Take an undecidable problem P
- Provide a computable function f that translates an instance I of P into a satisfiability problem for first order logic $f(I)$
- "Satisfiability for first-order logic is at least as difficult as P and hence undecidable"

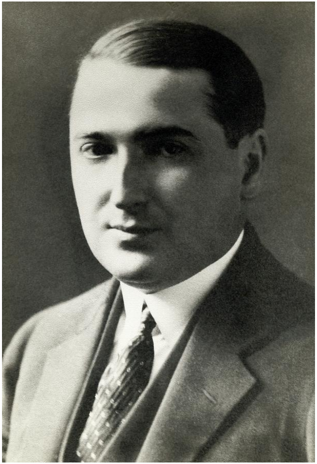
How to show undecidability

Principle:

- Take an undecidable problem P
- Provide a computable function f that translates an instance I of P into a satisfiability problem for first order logic $f(I)$
- "Satisfiability for first-order logic is at least as difficult as P and hence undecidable"

We choose P to be the **Post Correspondence Problem (PCP)**

Emil Post (1897 – 1954)



The post correspondence problem

In PCP, given set of **tiles** $(x_i, y_i) \in \{0,1\}^* \times \{0,1\}^*$, e.g.:

$$\left\{ \begin{bmatrix} 1 \\ 101 \end{bmatrix}, \begin{bmatrix} 10 \\ 00 \end{bmatrix}, \begin{bmatrix} 011 \\ 11 \end{bmatrix} \right\}$$

The post correspondence problem

In PCP, given set of **tiles** $(x_i, y_i) \in \{0,1\}^* \times \{0,1\}^*$, e.g.:

$$\left\{ \begin{bmatrix} 1 \\ 101 \end{bmatrix}, \begin{bmatrix} 10 \\ 00 \end{bmatrix}, \begin{bmatrix} 011 \\ 11 \end{bmatrix} \right\}$$

Solution is sequence of tiles such that top string equals bottom string:

$$\begin{bmatrix} 1 \\ 101 \end{bmatrix} \begin{bmatrix} 011 \\ 11 \end{bmatrix} \begin{bmatrix} 10 \\ 00 \end{bmatrix} \begin{bmatrix} 011 \\ 11 \end{bmatrix}$$

The post correspondence problem

Definition (Post Correspondence Problem (PCP))

An **instance of PCP** is a finite set

$$P = \{(x_1, y_1), \ldots, (x_k, y_k)\} \subseteq \{0,1\}^* \times \{0,1\}^*.$$

A **solution of** P is a sequence of indices $i_1, i_2, \ldots, i_n$ such that $i_j \in \{1, \ldots, k\}$, $1 \leq j \leq n$ and

$$x_{i_1} x_{i_2} \cdots x_{i_n} = y_{i_1} y_{i_2} \cdots y_{i_n}.$$

Reduction to first-order logic

$$\left\{ \begin{bmatrix} 1 \\ 101 \end{bmatrix}, \begin{bmatrix} 10 \\ 00 \end{bmatrix}, \begin{bmatrix} 011 \\ 11 \end{bmatrix} \right\}$$

Reduction to first-order logic

$$\left\{ \begin{bmatrix} 1 \\ 101 \end{bmatrix}, \begin{bmatrix} 10 \\ 00 \end{bmatrix}, \begin{bmatrix} 011 \\ 11 \end{bmatrix} \right\}$$

- introduce constant symbol e
- introduce unary function symbols f_0 and f_1
- introduce binary predicate symbol P
- write e.g. $f_{10110}(e)$ instead of $f_1(f_0(f_1(f_1(f_0(e)))))$

Reduction to first-order logic

$$\left\{ \begin{bmatrix} 1 \\ 101 \end{bmatrix}, \begin{bmatrix} 10 \\ 00 \end{bmatrix}, \begin{bmatrix} 011 \\ 11 \end{bmatrix} \right\}$$

- introduce constant symbol e
- introduce unary function symbols f_0 and f_1
- introduce binary predicate symbol P
- write e.g. $f_{10110}(e)$ instead of $f_1(f_0(f_1(f_1(f_0(e)))))$

$$\begin{aligned}
F_1 &= P(f_1(e), f_{101}(e)) \wedge P(f_{10}(e), f_{00}(e)) \wedge P(f_{011}(e), f_{11}(e)) \\
F_2 &= \forall u \forall v\, (P(u, v) \rightarrow P(f_1(u), f_{101}(v)) \wedge P(f_{10}(u), f_{00}(v)) \wedge \\
&\quad \wedge P(f_{011}(u), f_{11}(v))) \\
F_3 &= \exists u\, P(u, u)\,.
\end{aligned}$$

Reduction to first-order logic

Given instance P of PCP

$$P = \{(x_1, y_1), \ldots, (x_k, y_k)\} \subseteq \{0,1\}^* \times \{0,1\}^*.$$

Reduction to first-order logic

Given instance P of PCP

$$P = \{(x_1, y_1), \ldots, (x_k, y_k)\} \subseteq \{0,1\}^* \times \{0,1\}^*.$$

Define

$$\begin{aligned}
F_1 &= \bigwedge_{i=1}^{k} P(f_{x_i}(e), f_{y_i}(e)) \\
F_2 &= \forall u \, \forall v \, \bigwedge_{i=1}^{k} (P(u, v) \rightarrow P(f_{x_i}(u), f_{y_i}(v))) \\
F_3 &= \exists u \, P(u, u) \, .
\end{aligned}$$

Reduction to first-order logic

Given instance P of PCP

$$P = \{(x_1, y_1), \ldots, (x_k, y_k)\} \subseteq \{0,1\}^* \times \{0,1\}^*.$$

Define

$$\begin{aligned}
F_1 &= \bigwedge_{i=1}^{k} P(f_{x_i}(e), f_{y_i}(e)) \\
F_2 &= \forall u \forall v \bigwedge_{i=1}^{k} (P(u,v) \rightarrow P(f_{x_i}(u), f_{y_i}(v))) \\
F_3 &= \exists u \, P(u,u) \,.
\end{aligned}$$

Proposition

P has a solution if and only if $F_1 \land F_2 \rightarrow F_3$ is valid.

Reduction to first-order logic

$$F_1 = \bigwedge_{i=1}^{k} P(f_{x_i}(e), f_{y_i}(e)) \qquad\qquad F_3 = \exists u\, P(u,u)$$

$$F_2 = \forall u \forall v \bigwedge_{i=1}^{k} (P(u,v) \to P(f_{x_i}(u), f_{y_i}(v)))$$

- If $F_1 \wedge F_2 \to F_3$ is valid, consider structure $\mathcal{A}$ with universe $\{0,1\}^*$, $e_{\mathcal{A}} = \varepsilon$, $(f_0)_{\mathcal{A}}(\sigma) = \sigma 0$, $(f_1)_{\mathcal{A}}(\sigma) = \sigma 1$, and

$$P_{\mathcal{A}} = \{(\sigma,\tau) : \exists i_1 \ldots \exists i_t \,.\, \sigma = x_{i_1} \ldots x_{i_t} \text{ and } \tau = y_{i_1} \ldots y_{i_t}\}\,.$$

Reduction to first-order logic

$$F_1 = \bigwedge_{i=1}^{k} P(f_{x_i}(e), f_{y_i}(e)) \qquad\qquad F_3 = \exists u\, P(u,u)$$

$$F_2 = \forall u \forall v \bigwedge_{i=1}^{k} (P(u,v) \rightarrow P(f_{x_i}(u), f_{y_i}(v)))$$

- If $F_1 \wedge F_2 \rightarrow F_3$ is valid, consider structure $\mathcal{A}$ with universe $\{0,1\}^*$, $e_{\mathcal{A}} = \varepsilon$, $(f_0)_{\mathcal{A}}(\sigma) = \sigma 0$, $(f_1)_{\mathcal{A}}(\sigma) = \sigma 1$, and

  $$P_{\mathcal{A}} = \{(\sigma, \tau) : \exists i_1 \ldots \exists i_t \,.\, \sigma = x_{i_1} \ldots x_{i_t} \text{ and } \tau = y_{i_1} \ldots y_{i_t}\}\,.$$

  Now $\mathcal{A}$ satisfies $F_1 \wedge F_2$ and $F_1 \wedge F_2 \rightarrow F_3$, and so $\mathcal{A}$ satisfies F_3. But then P has solution.

Reduction to first-order logic

$$F_1 = \bigwedge_{i=1}^{k} P(f_{x_i}(e), f_{y_i}(e)) \qquad F_3 = \exists u\, P(u,u)$$

$$F_2 = \forall u \forall v \bigwedge_{i=1}^{k} (P(u,v) \rightarrow P(f_{x_i}(u), f_{y_i}(v)))$$

- If $F_1 \wedge F_2 \rightarrow F_3$ is valid, consider structure $\mathcal{A}$ with universe $\{0,1\}^*$, $e_\mathcal{A} = \varepsilon$, $(f_0)_\mathcal{A}(\sigma) = \sigma 0$, $(f_1)_\mathcal{A}(\sigma) = \sigma 1$, and

  $$P_\mathcal{A} = \{(\sigma, \tau) : \exists i_1 \ldots \exists i_t \,.\, \sigma = x_{i_1} \ldots x_{i_t} \text{ and } \tau = y_{i_1} \ldots y_{i_t}\}\,.$$

  Now $\mathcal{A}$ satisfies $F_1 \wedge F_2$ and $F_1 \wedge F_2 \rightarrow F_3$, and so $\mathcal{A}$ satisfies F_3. But then P has solution.
- If P has solution, consider any $\mathcal{A}$ that satisfies $F_1 \wedge F_2$. Show by induction on t that for any sequence of tiles $i_1 \ldots i_t$,

  $$\mathcal{A} \models P(f_u(e), f_v(e)), \text{ where } u = x_{i_1} \ldots x_{i_t} \text{ and } v = y_{i_1} \ldots y_{i_t}.$$

Reduction to first-order logic

$$F_1 = \bigwedge_{i=1}^{k} P(f_{x_i}(e), f_{y_i}(e)) \qquad\qquad F_3 = \exists u\, P(u,u)$$

$$F_2 = \forall u \forall v \bigwedge_{i=1}^{k} (P(u,v) \to P(f_{x_i}(u), f_{y_i}(v)))$$

- If $F_1 \wedge F_2 \to F_3$ is valid, consider structure $\mathcal{A}$ with universe $\{0,1\}^*$, $e_{\mathcal{A}} = \varepsilon$, $(f_0)_{\mathcal{A}}(\sigma) = \sigma 0$, $(f_1)_{\mathcal{A}}(\sigma) = \sigma 1$, and

  $$P_{\mathcal{A}} = \{(\sigma, \tau) : \exists i_1 \ldots \exists i_t \,.\, \sigma = x_{i_1} \ldots x_{i_t} \text{ and } \tau = y_{i_1} \ldots y_{i_t}\}\,.$$

  Now $\mathcal{A}$ satisfies $F_1 \wedge F_2$ and $F_1 \wedge F_2 \to F_3$, and so $\mathcal{A}$ satisfies F_3. But then P has solution.
- If P has solution, consider any $\mathcal{A}$ that satisfies $F_1 \wedge F_2$. Show by induction on t that for any sequence of tiles $i_1 \ldots i_t$,

  $$\mathcal{A} \models P(f_u(e), f_v(e)), \text{ where } u = x_{i_1} \ldots x_{i_t} \text{ and } v = y_{i_1} \ldots y_{i_t}\,.$$

Reduction to first-order logic

$$F_1 = \bigwedge_{i=1}^{k} P(f_{x_i}(e), f_{y_i}(e)) \qquad\qquad F_3 = \exists u\, P(u,u)$$

$$F_2 = \forall u \forall v \bigwedge_{i=1}^{k} (P(u,v) \rightarrow P(f_{x_i}(u), f_{y_i}(v)))$$

- If $F_1 \wedge F_2 \rightarrow F_3$ is valid, consider structure $\mathcal{A}$ with universe $\{0,1\}^*$, $e_{\mathcal{A}} = \varepsilon$, $(f_0)_{\mathcal{A}}(\sigma) = \sigma 0$, $(f_1)_{\mathcal{A}}(\sigma) = \sigma 1$, and

  $$P_{\mathcal{A}} = \{(\sigma,\tau) : \exists i_1 \ldots \exists i_t\,.\,\sigma = x_{i_1} \ldots x_{i_t} \text{ and } \tau = y_{i_1} \ldots y_{i_t}\}\,.$$

  Now $\mathcal{A}$ satisfies $F_1 \wedge F_2$ and $F_1 \wedge F_2 \rightarrow F_3$, and so $\mathcal{A}$ satisfies F_3. But then P has solution.
- If P has solution, consider any $\mathcal{A}$ that satisfies $F_1 \wedge F_2$. Show by induction on t that for any sequence of tiles $i_1 \ldots i_t$,

  $$\mathcal{A} \models P(f_u(e), f_v(e)), \text{ where } u = x_{i_1} \ldots x_{i_t} \text{ and } v = y_{i_1} \ldots y_{i_t}\,.$$

  But since P has solution, $\mathcal{A} \models P(f_u(e), f_u(e))$ for some string u. Thus $\mathcal{A} \models F_3$.

Reduction to first-order logic

$$F_1 = \bigwedge_{i=1}^{k} P(f_{x_i}(e), f_{y_i}(e)) \qquad\qquad F_3 = \exists u\, P(u,u)$$

$$F_2 = \forall u \forall v \bigwedge_{i=1}^{k} (P(u,v) \to P(f_{x_i}(u), f_{y_i}(v)))$$

- If $F_1 \wedge F_2 \to F_3$ is valid, consider structure $\mathcal{A}$ with universe $\{0,1\}^*$, $e_{\mathcal{A}} = \varepsilon$, $(f_0)_{\mathcal{A}}(\sigma) = \sigma 0$, $(f_1)_{\mathcal{A}}(\sigma) = \sigma 1$, and

  $$P_{\mathcal{A}} = \{(\sigma, \tau) : \exists i_1 \ldots \exists i_t \,.\, \sigma = x_{i_1} \ldots x_{i_t} \text{ and } \tau = y_{i_1} \ldots y_{i_t}\}\,.$$

  Now $\mathcal{A}$ satisfies $F_1 \wedge F_2$ and $F_1 \wedge F_2 \to F_3$, and so $\mathcal{A}$ satisfies F_3. But then P has solution.
- If P has solution, consider any $\mathcal{A}$ that satisfies $F_1 \wedge F_2$. Show by induction on t that for any sequence of tiles $i_1 \ldots i_t$,

  $$\mathcal{A} \models P(f_u(e), f_v(e)), \text{ where } u = x_{i_1} \ldots x_{i_t} \text{ and } v = y_{i_1} \ldots y_{i_t}\,.$$

  But since P has solution, $\mathcal{A} \models P(f_u(e), f_u(e))$ for some string u. Thus $\mathcal{A} \models F_3$.

Theorem

Validity and satisfiability in first-order logic are undecidable.