CS 228 : Logic in Computer Science

Krishna. S

Satisfiability of FOL

Given a formula in FOL over some signature τ, is it satisfiable?

Herbrand Theory

- Named after the French mathematician Jacques Herbrand
- Famous for Herbrand's Theorem, which allows a certain reduction from FOL to propositional logic
- Herbrand's theorem allows reducing a FOL formula φ in Skolem Normal Form to an infinite set $E(\varphi)$ of propositional formulae s.t. φ is satisfiable iff $E(\varphi)$ is satisfiable
- If $E(\varphi)$ is not satisfiable, then $\emptyset \in Res^*(E(\varphi))$, and we can derive this in finite number of steps
- As $E(\varphi)$ may be infinite, there is no way to say $\emptyset \notin Res^*(E(\varphi))$.

Herbrand Universe

Herbrand Universe

Let τ be a signature. The Herbrand universe for τ is the set of all ground terms.

Herbrand Universe

Herbrand Universe

Let τ be a signature. The Herbrand universe for τ is the set of all ground terms.

- If τ contains a constant c and unary function f, then the Herbrand universe contains $c, f(c), f(f(c)), f(f(f(c))), \ldots$

Herbrand Universe

Herbrand Universe

Let τ be a signature. The Herbrand universe for τ is the set of all ground terms.

- If τ contains a constant c and unary function f, then the Herbrand universe contains $c, f(c), f(f(c)), f(f(f(c))), \ldots$
- If τ contains a constant c and unary function f and binary function g, then the Herbrand universe contains distinct ground terms $c, g(c, c), f(c), g(c, f(c)), g(g(f(f(c)), c), f(c)), \ldots$

Herbrand Universe

- If τ has no constants, then the Herbrand universe is empty
- If τ has no functions, then the Herbrand universe consists of the constants of τ and is finite
- If τ has constants and functions, then the Herbrand universe is infinite

Herbrand Structures

Herbrand Structures

A structure $\mathcal{A}$ over a signature τ is a Herbrand structure if its underlying universe is a Herbrand universe.

- A Herbrand structure gives the natural interpretation to the constants and functions in τ : a constant c is interpreted as the element c in the universe,

Herbrand Structures

Herbrand Structures

A structure $\mathcal{A}$ over a signature τ is a Herbrand structure if its underlying universe is a Herbrand universe.

- A Herbrand structure gives the natural interpretation to the constants and functions in τ : a constant *c* is interpreted as the element *c* in the universe,
- If the signature τ has no relations or no constants, there is a unique Herbrand structure for τ

Herbrand Structures

- If the signature τ has relations and constants, then there are many Herbrand structures over τ depending on how you interpret them.

Herbrand Structures

- If the signature τ has relations and constants, then there are many Herbrand structures over τ depending on how you interpret them.
  - If τ contains a constant c and a binary relation R, then
  - $\mathcal{A} = (U^{\mathcal{A}} = \{c\}, R^{\mathcal{A}} = \{(c, c)\})$ is a Herbrand structure for τ.
  - $\mathcal{A} = (U^{\mathcal{A}} = \{c\}, R^{\mathcal{A}} = \{\})$ is a Herbrand structure for τ.
- If τ contains a constant c, function f and a unary relation R, then
  - $\mathcal{A} = (U^{\mathcal{A}} = \{c, f(c), f(f(c)), \ldots,\}, R^{\mathcal{A}} = \{c, f(c)\})$ is a Herbrand structure for τ.
  - $\mathcal{A} = (U^{\mathcal{A}} = \{c, f(c), f(f(c)), \ldots,\}, R^{\mathcal{A}} = \{c, f(c), f(f(f(f(c))))\})$ is a Herbrand structure, and so on.

Herbrand Signature

Let Γ be a set of sentences over a signature τ.

- The Herbrand signature for Γ is denoted τ_H.
- $\tau_H = \tau \cup \{c\}$ if τ contains no constants, else it is τ.
- The Herbrand universe for Γ denoted $H(\Gamma)$ is the Herbrand universe for τ_H.

Herbrand Model

A Herbrand model for Γ is a Herbrand structure M over τ_H such that $M \models \varphi$ for all $\varphi \in \Gamma$.

FO without equality

Let us focus on FO without "=". Recall that "=" is always interpreted as equality.

Herbrand Theorem

Let $\Gamma = \{\varphi_1, \varphi_2, \dots\}$ be a set of equality-free sentences in Skolem Normal Form. Then Γ is satisfiable iff Γ has a Herbrand model.

If Γ has a Herbrand model, clearly Γ is satisfiable. The converse needs a proof.

The converse

Assume Γ is satisfiable. Let τ_H be the Herbrand signature for Γ.

- Let $\mathcal{A}$ be a τ_H structure such that $\mathcal{A} \models \Gamma$. ($U^{\mathcal{A}}$ need not be the Herbrand universe)
- Let $\mathcal{B}$ be a Herbrand structure over τ_H. ($U^{\mathcal{B}}$ is the Herbrand universe)

The converse

Assume Γ is satisfiable. Let τ_H be the Herbrand signature for Γ.

- Let $\mathcal{A}$ be a τ_H structure such that $\mathcal{A} \models \Gamma$. ($U^{\mathcal{A}}$ need not be the Herbrand universe)
- Let $\mathcal{B}$ be a Herbrand structure over τ_H. ($U^{\mathcal{B}}$ is the Herbrand universe)
- Try "merging" $\mathcal{A}$ and $\mathcal{B}$ to obtain a Herbrand model M for Γ.

The converse

Assume Γ is satisfiable. Let τ_H be the Herbrand signature for Γ.

- Let $\mathcal{A}$ be a τ_H structure such that $\mathcal{A} \models \Gamma$. ($U^{\mathcal{A}}$ need not be the Herbrand universe)
- Let $\mathcal{B}$ be a Herbrand structure over τ_H. ($U^{\mathcal{B}}$ is the Herbrand universe)
- Try "merging" $\mathcal{A}$ and $\mathcal{B}$ to obtain a Herbrand model M for Γ.
  - Define M so that its universe is the Herbrand universe over τ_H.
  - Let M interpret functions and constants like $\mathcal{B}$ (both have the same Herbrand universe)
  - Let M interpret relations like $\mathcal{A}$ (not obvious, their universes are not the same.)

Building the Herbrand Model M

- Let R be ant n-ary relation in τ_H (hence in τ).
- For each n-tuple $(t_1, \ldots, t_n)$ with t_i coming from the Herbrand universe $H(\Gamma)$, we must say whether $(t_1, \ldots, t_n) \in R^M$ or not
- Each $t_i \in H(\Gamma)$ is a ground term in τ_H (so variable free).
- Since $\mathcal{A}$ is a structure over τ_H, if $t \in H(\Gamma)$ is a ground term from τ_H, $\mathcal{A}$ interprets t as an element of $U^{\mathcal{A}}$.
- For each n-tuple $(t_1, \ldots, t_n)$, we know whether $(t_1, \ldots, t_n) \in R^{\mathcal{A}}$ or not
- Define $R^M = R^{\mathcal{A}}$.
- Now we prove that if $\mathcal{A} \models \varphi$ for any $\varphi \in \Gamma$, then $M \models \varphi$.
- The proof is by induction on the number of quantifiers in φ. Recall that each φ is in Skolem Normal Form.

Base case : φ has 0 quantifiers

$\mathcal{A} \models \varphi$ iff $M \models \varphi$. Do structural induction on φ.

- Assume φ is an atomic formula. Then φ is $R(t_1, \ldots, t_n)$ where R is an n-ary relation from τ_H, and $t_1, \ldots, t_n$ are all terms from $H(\Gamma)$.
- By the construction of M, $R^M = R^{\mathcal{A}}$.
- Hence $M \models \varphi$ iff $\mathcal{A} \models \varphi$.
- Same reasoning holds for $\varphi_1 \wedge \varphi_2$, $\varphi_1 \wedge \varphi_2$ and $\neg\varphi$.
- Hence, $\mathcal{A} \models \varphi$ iff $M \models \varphi$.

Post Inductive Hypothesis

Assume that for any $\psi \in \Gamma$ with $\leqslant k-1$ quantifiers, if $\mathcal{A} \models \psi$, then $M \models \psi$. Let $\varphi \in \Gamma$ have k quantifiers, $\varphi = \forall x_1 \forall x_2 \ldots \forall x_k\ \zeta(x_1, \ldots, x_k)$ where ζ is quantifier free.

- Let $\kappa(x_1) = \forall x_2 \ldots \forall x_k\ \zeta(x_1, \ldots, x_k)$, and $\varphi = \forall x_1 \kappa(x_1)$.
- $\mathcal{A} \models \varphi$ implies $\mathcal{A} \models \forall x_1 \kappa(x_1)$. That is, $\mathcal{A} \models \kappa(a)$ for any $a \in U^{\mathcal{A}}$.
- Since $\mathcal{A}$ is a structure over τ_H, if $t \in H(\Gamma)$ is a ground term from τ_H, $\mathcal{A}$ interprets t as an element of $U^{\mathcal{A}}$.
- Thus, $\mathcal{A} \models \kappa(t)$ for any $t \in H(\Gamma)$.
- By induction hypothesis, $M \models \kappa(t)$ for any $t \in H(\Gamma)$.
- Since $H(\Gamma)$ is the universe of M, $M \models \forall x_1 \kappa(x_1)$. That is, $M \models \varphi$.

Equality

$$\varphi = \forall x[(f(x) \neq x) \wedge (f(f(x)) = x)].$$

Equality

$\varphi = \forall x[(f(x) \neq x) \wedge (f(f(x)) = x)]$.

- Herbrand universe over $\tau_H = \{c, f\}$ is $\{c, f(c), f(f(c)), \dots\}$ all distinct
- Then $f(f(c)) \neq c$, and a Herbrand structure cannot satisfy φ

Equality

$\varphi = \forall x[(f(x) \neq x) \wedge (f(f(x)) = x)]$.

- Herbrand universe over $\tau_H = \{c, f\}$ is $\{c, f(c), f(f(c)), \dots\}$ all distinct
- Then $f(f(c)) \neq c$, and a Herbrand structure cannot satisfy φ
- However, φ is satisfiable. Define a structure $\mathcal{A} = (\{0, 1\}, f^{\mathcal{A}}(0) = 1, f^{\mathcal{A}}(1) = 0)$, $\mathcal{A} \models \varphi$
- For formulae which have equality, Herbrand's Theorem does not apply directly
- If φ has equality, convert it to an equisatisfiable sentence without equality and apply Herbrand

Equality

$\varphi = \forall x[(f(x) \neq x) \wedge (f(f(x)) = x)]$.

- Herbrand universe over $\tau_H = \{c, f\}$ is $\{c, f(c), f(f(c)), \dots\}$ all distinct
- Then $f(f(c)) \neq c$, and a Herbrand structure cannot satisfy φ
- However, φ is satisfiable. Define a structure $\mathcal{A} = (\{0, 1\}, f^{\mathcal{A}}(0) = 1, f^{\mathcal{A}}(1) = 0)$, $\mathcal{A} \models \varphi$
- For formulae which have equality, Herbrand's Theorem does not apply directly
- If φ has equality, convert it to an equisatisfiable sentence without equality and apply Herbrand

Let φ be in Skolem normal form with equality. Then φ is satisfiable iff there is an equisatisfiable formula ψ in Skolem normal form without equality which has a Herbrand model.

Dealing with Equality

Assume φ is in Skolem Normal Form and uses "=". We define a equisatisfiable formula φ_E which does not use "=".

- Let τ be the signature of φ. Let E be a binary relation not in τ.
- Let $\varphi_{\neq}$ be the sentence obtained by replacing all occurrences of $t_1 = t_2$ in φ with $E(t_1, t_2)$.
- Define φ_{ER} to be the sentence

$$\forall x \forall y \forall z (E(x,x) \wedge ((E(x,y) \leftrightarrow E(y,x)) \wedge (E(x,y) \wedge E(y,z) \rightarrow E(x,z)))$$

- For each relation R in τ, define φ_R as

$$\forall x_1 \ldots \forall x_n \forall y_1 \ldots \forall y_n ((\bigwedge_{i=1}^{n} E(x_i, y_i) \wedge R(x_1, \ldots, x_n)) \rightarrow R(y_1, \ldots, y_n))$$

- Let $\varphi_1 = \bigwedge_{R \in \tau} \varphi_R$

Dealing with Equality

- For each function f in τ, define φ_f as

$$\forall x_1 \ldots \forall x_n \forall y_1 \ldots \forall y_n ((\bigwedge_{i=1}^{n} E(x_i, y_i) \rightarrow E(f(x_1, \ldots, x_n), f(y_1, \ldots, y_n)))$$

- Let $\varphi_2 = \bigwedge_{f \in \tau} \varphi_f$
- Let $\psi_E = \varphi_{\neq} \wedge \varphi_{ER} \wedge \varphi_1 \wedge \varphi_2$
- Convert ψ_E to Prenex normal form to obtain φ_E in Skolem normal form

For any formula φ in Skolem normal form, φ is satisfiable iff φ_E is satisfiable

An Example

$\varphi = \forall x[(f(x) \neq x) \wedge (f(f(x)) = x)]$.

- φ is satisfiable : $\mathcal{A} = (\{c, d\}, f^{\mathcal{A}}(c) = d, f^{\mathcal{A}}(d) = c)$ and $\mathcal{A} \models \varphi$.
- $\varphi_{\neq} = \forall x[\neg E(f(x), x) \wedge E(f(f(x)), x)]$
- $\varphi_2 = \forall x \forall y[E(x, y) \rightarrow E(f(x), f(y))]$
- Conjunct $\varphi_{\neq}, \varphi_2$ and φ_{ER} and convert to Prenex normal form
- $$\begin{aligned} \varphi_E = \forall x \forall y \forall z[&(\neg E(f(x), x) \wedge E(f(f(x)), x)) \wedge \\ &(E(x, y) \rightarrow E(f(x), f(y))) \wedge E(x, x) \wedge \\ &(E(x, y) \wedge E(y, z) \rightarrow E(x, z))] \end{aligned}$$
- By Herbrand's Theorem, φ_E has a Herbrand model $M = (\{c, f(c), f(f(c)), \ldots, \}, E^M = \{(t, t') \in H(\varphi_E) \mid$ the number of f's in both t, t' have the same parity$\})$
- $M \models \varphi_E$

Herbrand's Method

Given a FO sentence φ, is it satisfiable? Wlg, assume that φ is equality-free and is in Skolem normal form.

- Let $\varphi = \forall x_1 \ldots \forall x_n \psi(x_1, \ldots, x_n)$
- Let $H(\varphi)$ be the Herbrand universe of φ
- Let $E(\varphi) = \{\psi(t_1, \ldots, t_n) \mid t_1, \ldots, t_n \in H(\varphi)\}$ be the set obtained by substituting terms from $H(\varphi)$ for the variables $x_1, \ldots, x_n$ in φ
- φ is satisfiable iff $E(\varphi)$ is satisfiable

Herbrand's Method

- Assume φ is satisfiable. Then $\mathcal{A} \models \forall x_1, \ldots, x_n \psi(x_1, \ldots, x_n)$
- Then $\mathcal{A} \models \psi(t_1, \ldots, t_n)$ where $t_1, \ldots, t_n \in H(\varphi)$
- Then $\mathcal{A} \models \varphi_i$ for all $\varphi_i \in E(\varphi)$
- Hence, $E(\varphi)$ is satisfiable.

Herbrand's Method

- Assume $E(\varphi)$ is satisfiable. $E(\varphi)$ is a set of equality-free sentences.
- By Herbrand's Theorem, there is a Herbrand model M for $E(\varphi)$.
- The Herbrand signature for $E(\varphi)$ is the same as the Herbrand signature of φ.
- The universe of M is $H(\varphi)$
- For $t_1, \ldots, t_n \in H(\varphi)$, $M \models \psi(t_1, \ldots, t_n)$
- Then $M \models \forall x_1 \ldots x_n \psi(x_1, \ldots, x_n)$
- Then $M \models \varphi$ and φ is satisfiable.
- φ is unsatisfiable iff $E(\varphi)$ is unsatisfiable.

Checking Unsatisfiability of φ

- $E(\varphi) = \{\varphi_1, \varphi_2, \dots\}$ is a set of quantifier free sentences, so it can be seen as a set of propositional logic formulae
- Since φ is in Skolem normal form, each formula $\varphi_i \in E(\varphi)$ is in CNF
- We know that $E(\varphi)$ is unsatisfiable iff $\emptyset \in Res^*(E(\varphi))$
- That is, there is some finite subset $F = \{\varphi_1, \dots, \varphi_m\} \subseteq E(\varphi)$ such that $\emptyset \in Res^*(F)$
- So if $\emptyset \in Res^*(\{\varphi_1, \dots, \varphi_m\}$ for some finite m, we conclude φ is unsatisfiable

Checking Satisfiability of φ

- If $\emptyset \notin \mathit{Res}^*(\{\varphi_1, \ldots, \varphi_m\}$, then we look at $\mathit{Res}^*(\{\varphi_1, \ldots, \varphi_m, \varphi_{m+1}\}$
- If $\emptyset \notin \mathit{Res}^*(\{\varphi_1, \ldots, \varphi_{m+1}\}$, then we look at $\mathit{Res}^*(\{\varphi_1, \ldots, \varphi_{m+1}, \varphi_{m+2}\}$

  $\vdots$
- If φ is satisfiable, then this procedure will continue.

Wrapping Up

- We have a method to show that a FOL formula φ is unsatisfiable
- First, write φ in equality free Skolem normal form
- Check if $\emptyset \in \mathit{Res}^*(E(\varphi))$, this may take some time
- There is a more systematic resolution for FOL which we do not cover (this also uses Herbrand Theory)
- We also do not cover a direct undecidability proof for the satisfiability of FOL (at least now)