Lecture 14

Consequences of the Herbrand Universe Construction

February 24, 1998
Lecturer: Richard Ladner
Notes: Sumeet Sobti

14.1. Herbrand Universe Construction: An Example

Theorem-provers almost invariably use the Herbrand universe construction. The line of action is the following. Suppose we need to show that $\Gamma \models \alpha$. Now it suffices to show that $\Gamma \cup \{\neg\alpha\}$ is not satisfiable. For this, we construct $H_{\Gamma \cup \{\neg\alpha\}}$ (in fact, we construct only a part of it) and look for propositional unsatisfiability there. Lets we look at a simple example of theorem-proving by Herbrand universe construction.

Suppose we want to prove the following statement. If s is a 1-1 onto function and t is the left inverse of s, then s is the left inverse of t. (Note that t is the left inverse of s if and only if $\forall x, t(s(x)) = x$.)

The above statement can be formulated as $\Gamma \models \alpha$, where

$$\Gamma = \{\forall x \forall y(s(x) = s(y) \Longrightarrow x = y), \forall y \exists x(s(x) = y), \forall x(t(s(x)) = x)\}$$

$$\alpha = \forall x(s(t(x)) = x)$$

The first two formulae in Γ say that s is a 1-1 onto function. The third one says that t is the left inverse of s. The formula α says that s is the left inverse of t.

Thus,

$$\Gamma \cup \{\neg\alpha\} = \{\forall x \forall y(s(x) = s(y) \Rightarrow x = y), \forall y \exists x(s(x) = y), \forall x(t(s(x)) = x), \exists x(\neg(s(t(x)) = x))\}$$

Let $\Delta = \Gamma \cup \{\neg\alpha\}$. Now, we construct as much of H_Δ as we need. Clearly, Δ is already in prenex form. For *Skolemization*, we introduce a unary function g for the second sentence and a 0-ary function (i.e. a constant) c for the fourth sentence. Also adding the equality axioms, we get the following sentences in E_Δ.

1. $\forall x \forall y(s(x) = s(y) \Rightarrow x = y)$

2. $\forall y(s(g(y)) = y)$

3. $\forall x(t(s(x)) = x)$

4. $\neg(s(t(c)) = c)$

5. $\forall x(x = x)$

6. $\forall x \forall y(x = y \Rightarrow y = x)$

7. $\forall x \forall y \forall z(x = y \wedge y = z \Rightarrow x = z)$

8. $\forall x \forall y(x = y \Rightarrow (s(x) = s(y) \wedge t(x) = t(y) \wedge g(x) = g(y)))$

Now since we have already introduced a constant symbol c, we can construct the set of all terms as

$$T_\Delta = \{c, g(c), t(c), s(c), g(s(c)), \ldots\}$$

Now lets enumerate some formulas of H_Δ.

a. $s(g(c)) = c$ $\{$*Using 2* $\}$

b. $t(s(g(c))) = t(c)$ $\{$*Using 8 and a* $\}$

c. $t(s(g(c))) = g(c)$ $\{$*Using 3*$\}$

d. $t(c) = g(c)$ $\{$*Using 6, 7, b and c*$\}$

e. $s(t(c)) = s(g(c))$ $\{$*Using 8 and d*$\}$

f. $s(t(c)) = c$ $\{$*Using 7, a and e*$\}$

g. $\neg(s(t(c)) = c)$ $\{$*Using 4*$\}$

h. F $\{$*Using f and g*$\}$

Hence we conclude that $\Delta = \Gamma \cup \{\neg\alpha\}$ is not satisfiable.

Of course, for mechanization, we need strategies for making substitutions. Unification is one method of selecting substitutions of terms for variables. In order to combine the Herbrand construction with the propositional resolution procedure it is necessary to look for good substitutions where the result for example yields a literal in one clause and its complement in another.

14.2. Another Consequence of Herbrand Construction

Theorem 14.1: If Γ is a set of sentences (on a finite language) that has infinitely many finite models, then it has an infinite model.

Proof: Define formula $\alpha_i = \exists x_1 \dots \exists x_i (\bigvee_{j \neq k} \neg(x_j = x_k))$. Thus any model for α_i will have at least i elements. Let $\Delta = \Gamma \cup \{\alpha_i \mid i \geq 1\}$.

Claim: Every finite subset of Δ is satisfiable.
Proof: Suppose Φ is a finite subset of Δ. Let k be the maximum index such that $\alpha_k \in \Phi$. Now, if Γ doesn't belong to Φ, then any model of size at least k will satisfy Φ. If $\Gamma \in \Phi$, then any model M that satifsies Γ and has size at least k will satisfy Φ. Existence of such an M is ensured by the fact that Γ has infinitely many finite models. Thus, in any case, we see that Φ is satisfiable.

By the compactness theorem, we conclude that Δ is satisfiable. Now since $\forall i \geq 1 : \alpha_i \in \Delta$, Δ can't be satisfied by a finite model. Thus, Δ and hence Γ is satisfied by an infinite model. □

14.3. Graphs as Models

A directed graph $G = (V, E)$ can be seen as a model with universe V and a binary relation E defined on elements of V. A sentence ϕ may or may not be satisfied by G. Thus, ϕ defines a class of graphs namely, the graphs which satisfy ϕ. In other words, ϕ defines a *property* on graphs. For example,

- $\phi = \forall x(\neg E(x, x))$ represents the property of *not having any self-loops.*
- $\phi = \forall x \forall y(E(x, y) \Rightarrow \neg E(y, x))$ represents the *anti-symmetry* property.
- $\phi = \forall x \forall y(\neg(x = y) \Rightarrow E(x, y))$ is only satisfied by *cliques.*

Now the question comes "Can we describe all properties of graphs as first-order sentences?" And the answer is "no" as proved in the next theorem.

Theorem 14.2: The property of being *strongly-connected* is not a first-order property of directed graphs.

Proof: Assume that sentence Ψ_{SC} represents the property of being *strongly-connected.* Define sentences Ψ_{SL}, Ψ_{IN} and Ψ_{OUT} as follows.

- Let $\Psi_{SL} = \forall x(\neg E(x, x))$. Thus, Ψ_{SL} states the property of having *no self-loops.*
- Let $\Psi_{OUT} = \forall x \forall y \forall z(E(x, y) \wedge E(x, z) \Rightarrow y = z)$. Thus, Ψ_{OUT} states the property that each vertex in the graph has outdegree 1.
- Let $\Psi_{IN} = \forall x \forall y \forall z(E(y, x) \wedge E(z, x) \Rightarrow y = z)$. Thus, Ψ_{IN} states the property that each vertex in the graph has indegree 1.

Let $\Psi = \Psi_{SC} \wedge \Psi_{SL} \wedge \Psi_{OUT} \wedge \Psi_{IN}$. Thus Ψ describes the class of graphs that are strongly connected, have no self loops and have all vertices of indegree and outdegree 1. This is exactly the class of *cycle* graphs (of finite sizes). Clearly, there are infinitely many *cycle* graphs. Thus, by the previous theorem, there must be an infinite graph satisfying Ψ. But this is impossible. Hence we have a contradiction and the desired result follows. □

14.4. Valid First-Order Sentences

The set of valid propositional sentences is computable. What about the set of valid first-order sentences? As we see below, this is recursively enumerable, but not computable.

Note that $CN(\phi) = \{\alpha \mid \phi \models \alpha\}$ is exactly the set of valid first-order sentences. Now since ϕ is recursively enumerable, so is $CN(\phi)$. The following theorem shows that the set of valid first-order sentences is not computable.

Theorem 14.3: The set of valid first-order sentences is not computable.

Proof: A sentence α is satisfiable if and only if $\neg\alpha$ is not valid. Thus computability of the set of valid first-order sentences implies computability of the set of satisfiable first-order sentences.

Suppose $T = (S, \{a\}, \Gamma, \delta, s)$ is a Turing machine that accepts a non-computable subset of $\{a\}^*$. Let a^n be any input to T. Assume that if T does not accept x, then it never halts. Below we give an algorithmic construction of a sentence α_n such that α_n is satisfiable if and only if T does not accept x. Thus, by the observation made in the previous paragraph, the desired result will follow.

We represent a configuration of T on a^n as a string $\#\beta q\gamma$ where q is the current state, β is the tape contents to the left of the head and γ is the tape contents to the right of the head (including the symbol under the head). Let $\Delta = S \cup \Gamma \cup \{\#\}$ and $k = |\Delta|$. Thus, all symbols in any configuration belong to Δ. As we have seen before, there is a function $f : \Delta^4 \to \Delta$ such that if $\sigma_0, \sigma_1, \sigma_2$ and σ_3 are symbols in $i^{th}, (i+1)^{th}, (i+2)^{th}$ and $(i+3)^{th}$ positions of some configuration respectively, then the $(i+1)^{th}$ symbol in the succeeding configuration is $f(\sigma_0, \sigma_1, \sigma_2, \sigma_3)$.

Let $L = \{0, s, P\}$, where s is a unary function, P is a 3-ary relation and 0 is a constant.

Denote $s^i(0)$ by $\underline{i}$. $\underline{0}$ will denote 0 only. Lets fix an ordering of the symbols in Δ. If σ is the i^{th} symbol in that ordering, then we also denote $\underline{i}$ by $\underline{\sigma}$.

The sentence α_n that we construct is the conjunction of the following.

1. $\forall x(\neg(s(x) = 0))$
2. $\forall x \forall y(s(x) = s(y) \Rightarrow x = y)$
3. $\forall x \forall y \forall z(P(x, y, z) \Rightarrow (z = \underline{0} \vee z = \underline{1} \vee \ldots \vee z = \underline{k-1}))$
4. $\forall x \forall y \forall z \forall z'(P(x, y, z) \wedge P(x, y, z') \Rightarrow z = z')$
5. $\forall y(P(\underline{0}, y, \underline{\#}))$
6. $P(\underline{1}, \underline{0}, \underline{s}) \wedge \bigwedge_{i=2}^{n+1} P(\underline{i}, \underline{0}, \underline{a}) \wedge (\forall x((\bigwedge_{i=0}^{n+1} \neg(x = \underline{i})) \Rightarrow P(x, \underline{0}, \underline{\sqcup})))$
7. $\forall x \forall y(\bigwedge_{\sigma_0, \sigma_1, \sigma_2, \sigma_3 \in \Delta}(P(x, y, \underline{\sigma_0}) \wedge P(s(x), y, \underline{\sigma_1}) \wedge P(s(s(x)), y, \underline{\sigma_2}) \wedge P(s(s(s(x))), y, \underline{\sigma_3}) \Rightarrow P(s(x), s(y), \underline{f}(\sigma_0, \sigma_1, \sigma_2, \sigma_3))))$
8. $\forall x \forall y(\neg P(x, y, \underline{A}))$, where A is the accept state.

Claim: α_n is satisfiable $\Longleftrightarrow T$ does not accept a^n.
Proof: ($\Leftarrow$) Suppose T does not accept a^n. Define a model $M = (U^M, I^M)$ as follows. Let $U^M = \{0, 1, 2, \ldots\}$, the set of non-negative integers. I^M defines s^M to be the successor relation among the integers in U^M. Relation P^M is defined as: $P^M(x, y, z)$ if and only if "the symbol in position x of the y^{th} configuration of T (when run on a^n) is the z^{th} symbol of Δ (in the ordering fixed above)". Now it is easy to verify that all of the above sentences are satisfied by M. Sentence 1 states that 0 is not the successor of any other element of the universe. Sentence 2 states that each element has a unique successor. Sentence 3 states that in each position of each configuration of T, there is some symbol of Δ. Sentence 4 states that this symbol has got to be unique. Sentence 5 states that the first symbol of all configurations is $\#$. Sentence 6 states that the 0^{th} configuration is the initial configuration of T on a^n. Sentence 7 states that all the transitions between successive configurations are valid. Sentence 8 states that machine T never reaches the accept state. It is easy to see that M satisfies all of these. Thus, α_n is satisfiable.

($\Rightarrow$) Suppose $M = (U^M, I^M)$ is a satisfying model for α_n. Then, sentences 1 and 2 ensure that the elements $\underline{0}, \underline{1}, \underline{2}, \ldots$ in U^M are in a 1-1 correspondence with non-negative integers. (Remember that $\underline{i}$ denotes $s^i(0)$). Now since sentences in α_n only have universal quantifiers, we can restrict M to obtain model $M' = (U^{M'}, I^{M'})$, where $U^{M'} = \{\underline{0}, \underline{1}, \underline{2}, \ldots\}$ and $I^{M'}$ defines $s^{M'}$ and $P^{M'}$ to be restrictions of s^M and P^M respectively to $U^{M'}$. It is easy to see that M' also satisfies α_n. Sentences 3 through 7 ensure that $P^{M'}$ must now directly correspond to the *transition relation* of T on a^n. Sentence 8 ensures that T does not accept a^n. Hence the claim follows.

$\square$