Last time: Skolemization

For any \( \varphi \in L(\Sigma) \), we produce \( \varphi_S \in L(\Sigma') \) which is universal and in RPF-CNF form and is equi-satisfiable with \( \varphi \), i.e., \( \varphi \in \text{FO-SAT} \iff \varphi_S \in \text{FO-SAT} \).

\[
\varphi_s \equiv \forall x_1 \cdots x_k \bigwedge_{i=1}^{a} \bigvee_{j=1}^{b} \ell_{ij}
\]

Resolution for FO:

To show \( \Gamma \vdash \alpha \), we show that \( \Gamma \cup \{\neg \alpha\} \vdash \square \).

Let \( \varphi \overset{\text{def}}{=} \Gamma \cup \{\neg \alpha\} \). (Assume that \( \Gamma \) is a finite set of formulas. Otherwise we would do this repeatedly for increasing subsets of \( \Gamma \) until we prove \( \square \).)

Put \( \varphi \) into RPF-CNF form and then Skolemize, getting \( \varphi_S \equiv \forall x_1 \cdots x_k (C_1 \land C_2 \land \cdots C_t) \).

We then try to use resolution to show that \( \{C_1, \ldots, C_t\} \vdash \square \), where we remember that all the variables are universally quantified, so at any step, for any clause, \( C_i \), we may substitute any term, \( t \) for any variable, \( x_j \), getting the new clause \( C_i[x_j/t] \).

Goal for today and next time: we will show that FO resolution is complete. Thus we will prove,

**Gödel’s Completeness Theorem** [1929] There is a complete proof system for FO, i.e., for all \( \varphi \in L(\Sigma) \) and \( \Gamma \subseteq L(\Sigma) \), \( \Gamma \models \varphi \iff \Gamma \vdash \varphi \). (Gödel proved in his Ph.D. thesis that the proof system given in Principia Mathematica by Russell and Whitehead is complete. We will prove that Resolution is complete.)

Herbrand Theory:

We will assume that our vocabulary, \( \Sigma \), always includes at least one constant symbol, \( c \). For example, let \( \Sigma = (R_1^{a_1}, \ldots, R_s^{a_s}; c_2, \ldots, c_k, f_1^{r_1}, \ldots, f_t^{r_t}) \). recall that we defined \( \text{term}(\Sigma) \) by induction. A term is closed if it has no variables. Since every vocabulary has the constant symbol, \( c, c \in \text{closedTerm}(\Sigma) \), so \( \text{closedTerm}(\Sigma) \neq \emptyset \).

Example: \( \Sigma_0 = (R^1; c) \). \( \text{closedTerm}(\Sigma_0) = \{c\} \).

Example: \( \Sigma_{\text{#thy}} = (\leq^2; 0, 1, +^2[\text{infix}], *^2[\text{infix}]) \). \( \text{closedTerm}(\Sigma_{\text{#thy}}) = \{0, 1, 0 + 0, 0 + 1, (0 * 0), (1 + 1) * (1 + 1), (1 + 1) * (1 + 1) + 1, \ldots\} \).

Recall that each term \( t \in \text{term}(\Sigma) \) is a sequence of symbols. Each structure \( A \in \text{STRUC}[\Sigma] \) interprets the term \( t \) as an element of its universe: \( t^A \in |A| \).

Def. The Herbrand Universe for \( (\Sigma) \overset{\text{def}}{=} \text{closedTerm}(\Sigma) \).

Def. A Herbrand Structure \( \mathcal{H} \in \text{STRUC}[\Sigma] \) has the Herbrand Universe as its universe, i.e., \( |\mathcal{H}| = \text{closedTerm}(\Sigma) \). Furthermore, for every function symbol \( f^r \in \Sigma \), and every \( t_1, \ldots, t_r \in |\mathcal{H}|, \quad f^H(t_1, \ldots, t_r) = f(t_1, \ldots, t_r) \).

**Proposition.** Let \( \mathcal{H} \) be a Herbrand Structure of vocabulary \( \Sigma \) and let \( t \) be any closed term of \( \Sigma \). Then \( t^H = t \).

**Proof:** By induction on \( t \). \( \square \)

Example: \( \Sigma_0 \) has exactly two Herbrand structures: \( \mathcal{H}_0, \mathcal{H}_1 \) where \( \mathcal{H}_0 = (\{c\}, R^{H_0} = \emptyset); \quad \mathcal{H}_1 = (\{c\}, R^{H_1} = \{c\}) \).
Note that $\mathcal{H}_0 \models \forall x \neg R(x)$ and $\mathcal{H}_1 \models \forall x R(x)$. Let $\beta = \exists x \forall y (R(x) \land \neg R(y))$. Let $\mathcal{A}_2 = \{\{0, 1\}, r^{\beta_2} = \{1\}\}$. Note that $\mathcal{A}_2 \models \beta$. However, $\mathcal{H}_0 \models \neg \beta$ and $\mathcal{H}_1 \models \neg \beta$. Thus, all the Herbrand structures of vocabulary $\Sigma_0$ satisfy $\neg \beta$. Thus, $\beta$ is satisfiable, but has no Herbrand model with vocabulary $\Sigma_0$.

**Note:** For a larger vocabulary, $\Sigma_1 = \Sigma_0 \cup \{f^1\}$, $\beta$ has a Herbrand model of vocabulary $\Sigma_1$.

closedTerms($\Sigma_1$) = \{c, f(c), f^2(c), f^3(c), \ldots\}.

Let $\mathcal{H}_{even}$ be a Herbrand Structure of vocabulary $\Sigma_1$, such that $R^{\mathcal{H}_{even}} = \{c, f^2(c), f^4(c), \ldots\}$. Then $\mathcal{H}_{even} \models \beta$.

**Herbrand’s Theorem** Let $\Sigma$ be a vocabulary s.t. closedTerm($\Sigma$) $\neq \emptyset$. Let $\varphi \in L(\Sigma)$ be a universal formula, $\varphi = \forall x_1 \cdots x_k(\alpha)$ where $\alpha$ is quantifier free. Then $\varphi \in \text{FO-SAT}$ iff $\varphi$ has a Herbrand model.

**Proof:** $\Leftarrow$: If $\varphi$ has a Herbrand model, $\mathcal{H} \models \varphi$, then $\varphi$ has a model, so it is satisfiable.

$\Rightarrow$: Assume that $\varphi \in \text{FO-SAT}$ and let $\mathcal{A} \models \varphi$.

**Goal:** define a Herbrand structure, $\mathcal{H}$, s.t. $\mathcal{H} \models \varphi$.

Two parts of building $\mathcal{H}$ are trivial, i.e., we must have that $|\mathcal{H}| = \text{closedTerm}(\Sigma)$ and for all $f^r \in \Sigma$ and $t_1, \ldots, t_r \in \text{closedTerm}(\Sigma)$, $f^\mathcal{H}(t_1, \ldots, t_r) = f(t_1, \ldots, t_r)$.

What remains to be defined is $R^\mathcal{H}$ for each $R^a \in \Sigma$. For these definitions, we just ask $\mathcal{A}$:

$$R^\mathcal{H} \equiv \{(t_1, \ldots, t_a) \in |\mathcal{H}|^a \mid \mathcal{A} \models R(t_1, \ldots, t_a)\} \quad (*)$$

**Claim.** For all closed quantifier-free $\beta \in L(\Sigma)$, $\mathcal{H} \models \beta$ iff $\mathcal{A} \models \beta$.

**Proof:** By induction on $\beta$.

**base case:** $\beta = R(t_1, \ldots, t_a)$. By $(*)$, $\mathcal{H} \models \beta$ iff $\mathcal{A} \models \beta$.

**inductive case:** Assume true for $\beta_1$ and $\beta_2$.

$$\mathcal{H} \models \neg \beta_1 \Leftrightarrow \mathcal{H} \not\models \beta_1 \Leftrightarrow \mathcal{A} \not\models \beta_1 \Leftrightarrow \mathcal{A} \models \neg \beta_1$$

$$\mathcal{H} \models \beta_1 \lor \beta_2 \Leftrightarrow (\mathcal{H} \models \beta_1) \text{ or } (\mathcal{H} \models \beta_2) \Leftrightarrow (\mathcal{A} \models \beta_1) \text{ or } (\mathcal{A} \models \beta_2) \Leftrightarrow \mathcal{A} \models \beta_1 \lor \beta_2$$.

$\Box$

Finally we show that $\mathcal{H} \models \forall x_1 \cdots x_k(\alpha)$.

Since $\mathcal{A} \models \forall x_1 \cdots x_k(\alpha)$, it follows that for all closed terms, $t_1, \ldots, t_k$, $\mathcal{A}, x_1/t_1^A, \ldots, x_k/t_k^A \models \alpha$.

That is, for all closed terms, $t_1, \ldots, t_k$, $\mathcal{A} \models \alpha[x_1/t_1^A, \ldots, x_k/t_k^A]$.

Thus, by our Claim, for all closed terms, $t_1, \ldots, t_k$, $\mathcal{H} \models \alpha[x_1/t_1^A, \ldots, x_k/t_k^A]$.

That is, for all closed terms, $t_1, \ldots, t_k$, $\mathcal{H}, x_1/t_1, \ldots, x_k/t_k \models \alpha$.

Thus, $\mathcal{H} \models \forall x_1 \cdots x_k(\alpha)$. $\Box$