Herbrand Models

Logic Lecture 2

Recall our discussion of the "simplest model" for a set of sentences.

Wouldn't it be nice to have a unique simplest model?

Example: Models

$\forall X(\forall Y((mother(X) \wedge child_of(Y,X)) \rightarrow loves(X,Y)))$

mother(mary)

child_of(tom,mary)

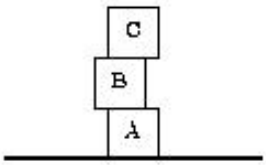


Figure 1.2: An alternative structure

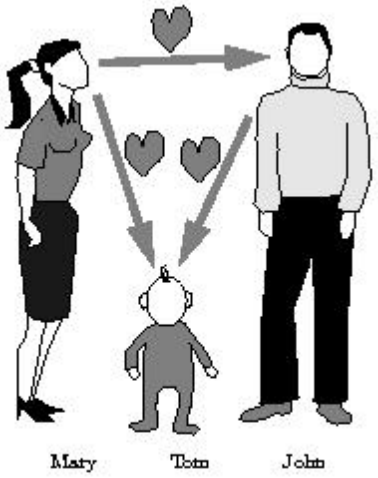
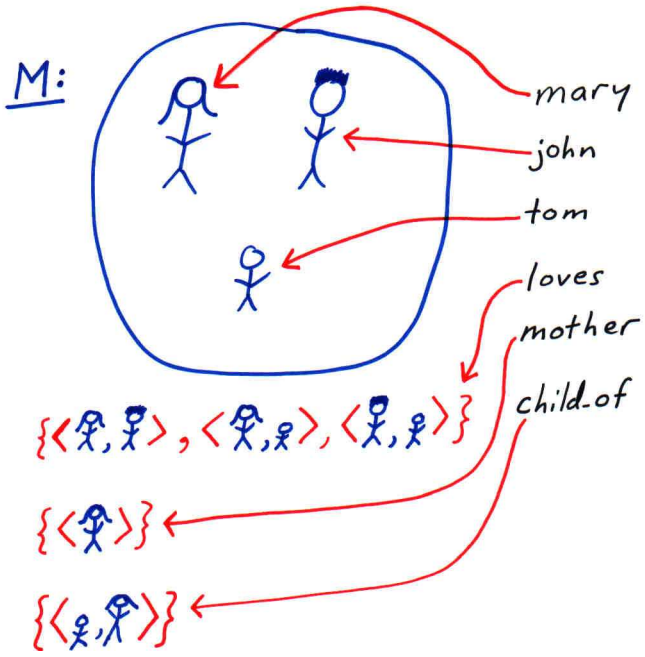
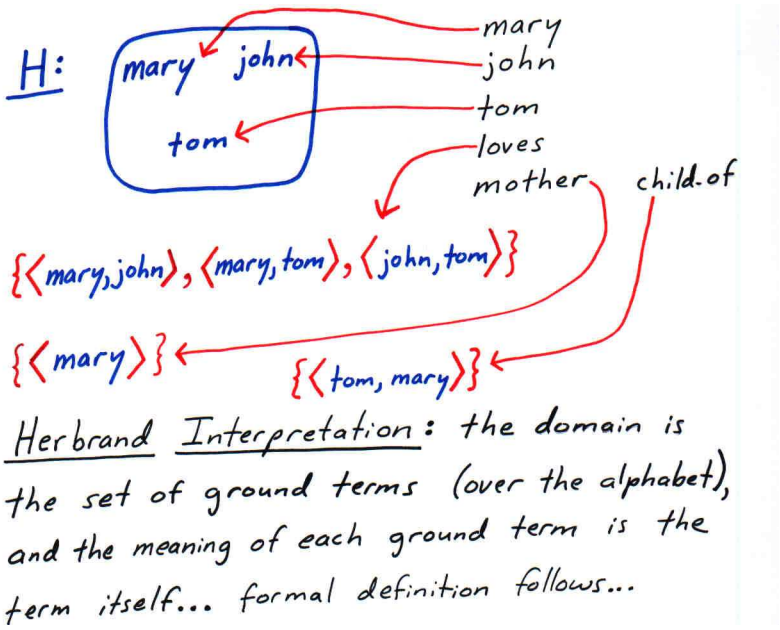


Figure 1.1: A family structure



Problem…

- Difficult to compare two interpretations with different domains… e.g., one domain consists of apples and the other of oranges.
- Could map one domain to another. Can be tricky to define… most domains are infinite.
- Idea: for a given alphabet, pick a *canonical* domain and mapping. But how?



Let $\mathcal{A}$ be an alphabet containing at least one constant symbol. The set $U_{\mathcal{A}}$ of all ground terms constructed from functors and constants in $\mathcal{A}$ is called the Herbrand universe of $\mathcal{A}$. The set $B_{\mathcal{A}}$ of all ground, atomic formulas over $\mathcal{A}$ is called the Herbrand base of $\mathcal{A}$.

A Herbrand interpretation for alphabet $\mathcal{A}$ is an interpretation $\mathcal{I} = (\mathcal{D}, \mathcal{I})$ where:

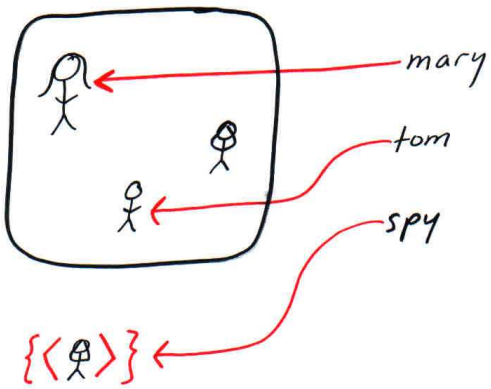
- $\mathcal{D} = U_{\mathcal{A}}$
- for any n-ary function symbol f $[\![f]\!]^{\mathcal{I}}(t_1, \ldots, t_n) = f(t_1, \ldots, t_n)$, where $t_1, \ldots, t_n$ are any ground terms
- for any n-ary predicate symbol p $[\![p]\!]^{\mathcal{I}}$ is an n-ary relation over $U_{\mathcal{A}}$ (subset of the set of all n-tuples over $U_{\mathcal{A}}$).

Some Notes

- Typically, we are given a theory (set of sentences) *T* and wish to speak of Herbrand interpretations relative to *T*.
- In this case we take the alphabet *A* to be the symbols in *T*.
- If *T* has no constants, we introduce one.

Notes (continued)

- Valuations with respect to a Herbrand interpretation may be thought of as grounding substitutions.
- We'd like know it's sufficient to consider only Herbrand interpretations… just ignore all others…



$\exists X \; spy(X)$
$\neg spy(mary)$
$\neg spy(tom)$

Herbrand Model Lemma

- Let *T* be a theory (set of sentences) in *Skolem Normal Form*.
- *T* has a model iff it has a Herbrand model.

Prenex Form

Move all quantifiers to the front of each sentence.

Example

$$\forall X \forall Y (gc(X,Y) \vee \neg \exists Z (c(X,Z) \wedge c(Z,Y)))$$
$$\forall X \forall Y (gc(X,Y) \vee \forall Z \neg (c(X,Z) \wedge c(Z,Y)))$$
$$\forall X \forall Y \forall Z (gc(X,Y) \vee \neg (c(X,Z) \wedge c(Z,Y)))$$

Skolem Normal Form

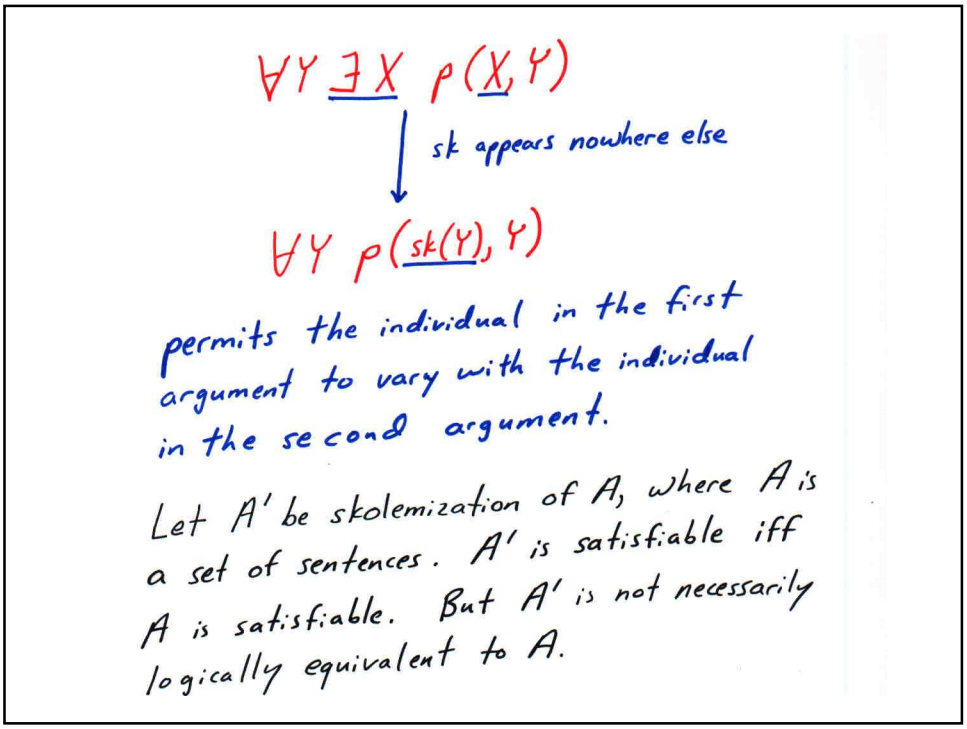
Get rid of all existential quantifiers.

$$\exists X \forall Y \; p(X,Y)$$

↓ sk appears nowhere else

$$\forall Y \; p(sk, Y)$$

Introduces "skolem functions" such as sk (here sk has arity 0).



Skolemization

- Process is applied to one sentence at a time and applied only to the entire sentence (so outermost quantifier first). Each sentence initially has empty vector of free variables.
- Replace ∀X A(X) with A(X), and add X to vector of free variables.
- Replace ∃X A(X) with A(x(V)) where x is a new function symbol and V is the current vector of free variables.

Herbrand Model Lemma

- Let *T* be a theory (set of sentences) in Skolem Normal Form.
- *T* has a model iff it has a Herbrand model.

Now recall our goal of identifying a unique simplest model.

Example: Models

$\forall X(\forall Y((mother(X) \wedge child_of(Y,X)) \rightarrow loves(X,Y)))$

mother(mary)

child_of(tom,mary)

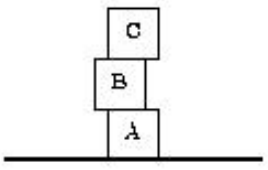


Figure 1.2: An alternative structure

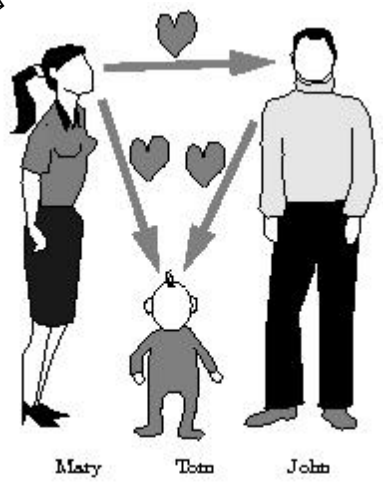
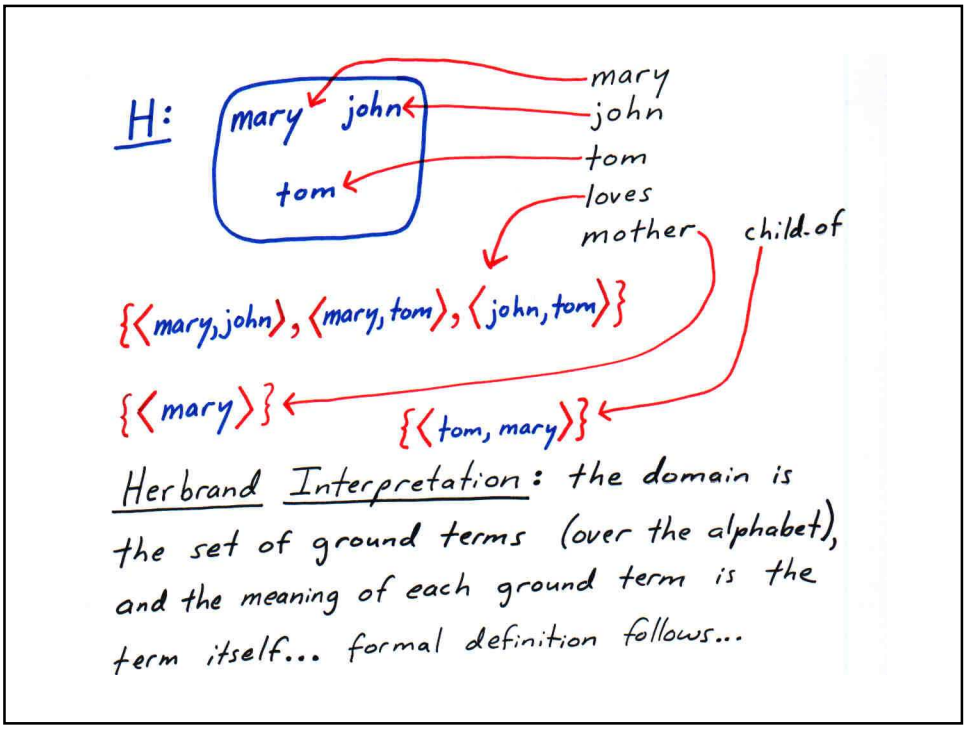
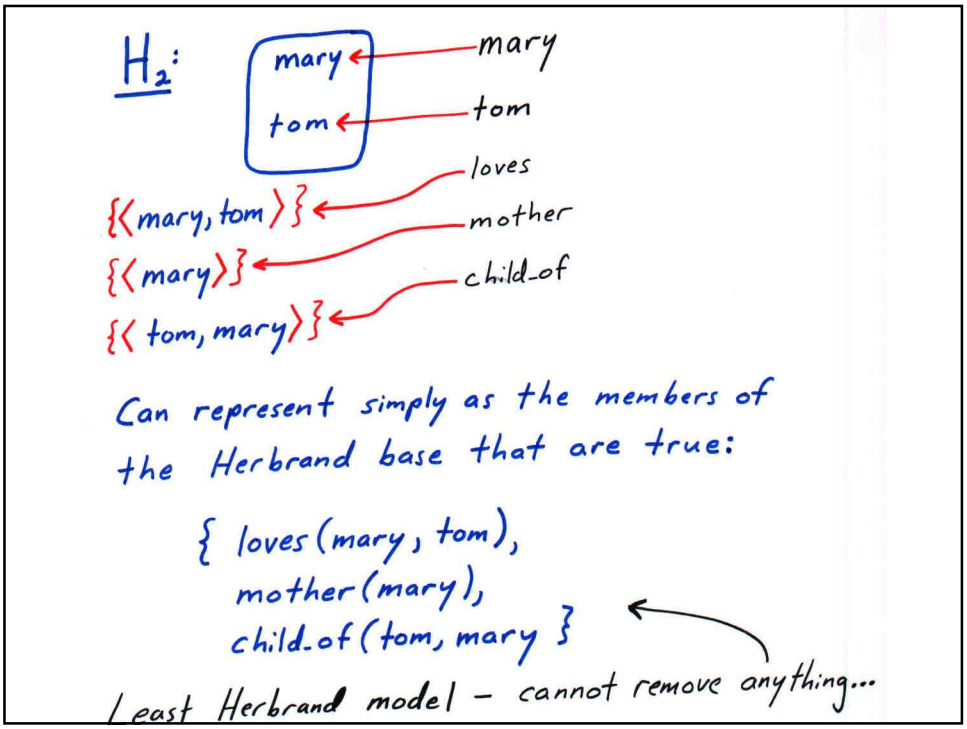


Figure 1.1: A family structure



Herbrand Interpretation: the domain is the set of ground terms (over the alphabet), and the meaning of each ground term is the term itself... formal definition follows...



Can represent simply as the members of the Herbrand base that are true:

{ loves(mary, tom),
mother(mary),
child_of(tom, mary }

least Herbrand model – cannot remove anything...

Minimal Herbrand model: Herbrand model such that any subset is not a model.

odd (s(0))

odd (s(s(X))) ← odd (X).

model {odd(0), odd(s(0)), odd(s(s(0))), ... }

∪

⋮

∪

model {odd(s(0)), odd(s(s(s(0)))), ... }

∪

not model

Least Herbrand model: Herbrand model such that every other Herbrand model is a superset.

May not exist:

p(a) ∨ p(b)

{p(a), p(b)}

{p(a)} {p(b)}

Why no least Herbrand model?

- Disjunctive "positive" information… creates uncertainty. We can satisfy the disjunction by satisfying either disjunct – a choice.
- This is somewhat analogous to the uncertainty created by existential quantifiers.
- This uncertainty also causes inefficiencies in deduction (recall prop. SAT is NP-complete but SAT for Horn CNFs is linear-time solvable).

Clausal (CNF) Form

Set (conjunction) of clauses.

Clause: Universally quantified disjunction of literals.

Literal: atomic formula or its negation.

$$\forall X \forall Y \forall Z \, (\neg p(X,Y) \vee q(Y,Z))$$

Can convert any set of sentences A into CNF theory A' such that A is satisfiable iff A' is satisfiable.

Definite Clause

A clause is definite iff it has exactly one positive (unnegated) literal.

$$\forall X \forall Y \forall Z \, (p(X,Z) \vee \neg p(X,Y) \vee \neg p(Y,Z))$$

Logically equivalent to implication with the positive literal in the consequent (head) and the negative literals conjoined, unnegated in the antecedent (body).

$$\forall X \forall Y \forall Z \, (p(X,Z) \leftarrow (p(X,Y) \wedge p(Y,Z)))$$

Definite Program...

... is a finite set of definite clauses.

For notational simplicity, we drop the explicit universal quantifiers. We always write the heads on the left. A definite clause with an empty body is called a "fact."

Examples of Definite Programs

mother(mary)
child_of(tom,mary)
loves(X,Y) $\leftarrow$ *mother(x)* $\wedge$ *child_of(X,Y)*

odd(s(0))
odd(s(s(X))$\leftarrow$ *odd(X)*

Prolog Notation

mother(mary).
child_of(tom,mary).
loves(X,Y):- mother(x), child_of(X,Y).

odd(s(0)).
odd(s(s(X)):- odd(X).

Every definite program has a least Herbrand model. Assume not...

Then for two minimal models H_1 and H_2, $H_1 \cap H_2$ is not a model.

Then some ground instance of a clause in the program

$$A_0 \leftarrow A_1 \wedge \cdots \wedge A_m$$

is false in $H_1 \cap H_2$. Then $H_1 \cap H_2$ contains $A_1, \ldots, A_m$ but not A_0. Then both H_1 and H_2 contain $A_1, \ldots, A_m$ but one lacks A_0 and hence is not a model — contradiction. □

About Least Herbrand Models

- The least Herbrand model M_P of a program P is the set of all ground atomic logical consequences of the program.
- In general it is undecidable whether a ground atomic formula is in the least Herbrand model of a program (logically follows from the program). But if it follows, it can be eventually shown…

Alternative Characterization of Least Herbrand Model

- Let P be a definite program. T_P is a function on Herbrand interpretations defined as follows:

  $T_P(I) = \{A_0 \mid A_0 \leftarrow A_1,\ldots,A_m \in P_{gr}$ and $\{A_1,\ldots,A_m\} \subseteq I\}$

- The least interpretation I such that $T_P(I) = I$ is the least Herbrand model of P.

Contruction/Approximation of Least Herbrand Model

- $T_P \uparrow 0 = \varnothing$
- $T_P \uparrow (i+1) = T_P(T_P \uparrow i)$
- $T_P \uparrow \omega$ is the union of $T_P \uparrow i$ for all i from 0 to ∞

- The least Herbrand model M_P of P is the least fixpoint of T_P: the least Herbrand interpretation such that $T_P(M_P) = M_P$.
- $M_P = T_P \uparrow \omega$.

Example

- $odd(s(0))$.
- $odd(s(s(X)) \leftarrow odd(X)$.

- $T_P \uparrow 0 = \varnothing$
- $T_P \uparrow 1 = \{odd(s)\}$
- $T_P \uparrow 2 = \{odd(s(s(s(0))), odd(s)\}$
- $T_P \uparrow \omega = \{odd(s^n(0)) \mid n \in \{1,3,5,\ldots\}\}$