<is web>

Information Systems & Semantic Web
University of Koblenz ▪ Landau, Germany

Advanced Data Modeling

5: Semantics

Steffen Staab
with
Simon Schenk

UNIVERSITÄT KOBLENZ · LANDAU

Overview

- Logic as query language.
- Grounding.
- Minimal Herbrand models.
- Completion.

Logic as query language

Given:

- first-order formula $A[x_1, \ldots, x_n]$
- Herbrand interpretation I

This first-order formula can be considered as a definition of a relation R_A on T^n_Σ as follows:

$$(t_1, \ldots, t_n) \in R_A \quad := \quad I \models A[t_1, \ldots, t_n]$$

Clauses as definitions

We say that a clause

$p(t_1, \ldots , t_m)$:- $L_1 , \ldots , L_n$

defines the relation symbol *p*.

Let C be a set of clauses and *p* be a relation symbol. We call the definition of *p* in C the set of all clauses in C that define *p*.

Principles of semantics

- A deductive database is a **set of clauses.**
- This set of clauses is regarded as a **collection of definitions** of relations.
- The **semantics** defines the meaning of this definitions by associating with them an **interpretation**, or a class of interpretations.
- **Query answering** is based on the semantics.

Two key assumptions

- the **unique name assumption**:
  each name denotes a unique object.

- the **closed world assumption**:
  - a negative statement $\neg A$ holds if the corresponding positive one A does not hold.

Minimal Herbrand Models

Let *I* be a Herbrand model of a set of formulas *S*.

We call *I* a minimal Herbrand model of *S* if it is minimal w.r.t. the subset relation, i.e. for every Herbrand model *I´* of *S* of the same signature we have $I´ \subseteq I$.

I is called the least Herbrand model of *S* if for every Herbrand model *I´* of *S* of the same signature we have $I \subseteq I´$.

Does every set of formulas S have a least Herbrand model?

Steffen Staab
staab@uni-koblenz.de

Instance and variant

Let E, E´ be a pair of terms or formulas.

- E´ is an **instance** of E, denoted E < E´, if there exists a substitution θ such that Eθ = E´.
- **ground instance**: instance that is ground,
- E´ is a **variant** of E if E´ is an instance of E and E is an instance of E´.

Examples

- P(x,a) is instance of P(x,y) because of P(x,y)[y|b]
- P(b,a) is a ground instance
- P(x,y) and P(u,v) are variants of each other, because of
  - [x|u, y|v] and
  - [u|x, v|y]

Grounding

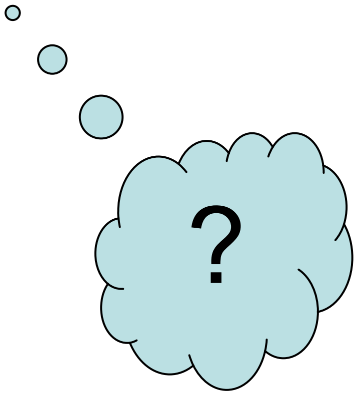
Let C be a set of clauses and Σ be any signature containing all symbols used in C. The **grounding of C w.r.t. Σ,** denoted **C^*** is the set of all ground instances of the signature Σ of clauses in C.

Lemma. Let I be a Herbrand interpretation and C be a set of clauses. Then $I \models C$ if and only if $I \models C^*$.

Proof

Logic with equality

- Additional atomic formulas **s = t**, where s, t are terms.
- Abbreviation: $x \neq y \ := \ \neg(x = y)$.
- Unlike other relations, the semantics of s = t is **predefined** in all Herbrand interpretations:
  $I \models s = t$ if s coincides with t.



Example valid formulas

$f(x_1, \ldots, x_n) = f(y_1, \ldots y_n) \rightarrow x_1 = y_1 \wedge \ldots \wedge x_n = y_n$

$f(x_1, \ldots, x_n) \neq g(y_1, \ldots, y_n)$

$f(x_1, \ldots, x_n) \neq c$

$d \neq c$

$A[t] \leftrightarrow \forall x(x = t \rightarrow A[x])$

Semantics of definitions

Consider a definition of a relation *r*

$$r(\bar{t}_1) :- G_1$$

$$\ldots$$

$$r(\bar{t}_m) :- G_m$$

What is the meaning of this definition?

Replace every clause by an equivalent one such that the arguments of

r are $x_1, \ldots, x_n$:

Given:

$r(t_1, \ldots, t_n)$:- G

Replace by:

$r(x_1, \ldots, x_n)$:- $x_1 = t_1 \wedge \ldots \wedge x_n = t_n \wedge G$

Completion. Step 2.

If there are variables $y_1, \ldots, y_k$ occurring in a body but not in the head, apply $\exists$ to these variables, i.e.,

Given

$r(x_1, \ldots, x_n)$:- G

Modify to

$r(x_1, \ldots, x_n)$:- $\exists\, y_1 \ldots \exists\, y_k$ G

Completion. Step 3.

If there are several definitions, replace them by one

Given

$r(x_1, \ldots ,x_n) \text{ :- } G_1$

$\ldots$

$r(x_1, \ldots , x_n) \text{ :- } G_m$

Replace by

$r(x_1, \ldots , x_n) \text{ :- } G_1 \vee \ldots \vee G_m$

Completion. Step 4.

Replace :- by $\leftrightarrow$:

Given

$r(x_1, \ldots, x_n) \text{ :- } G_1 \vee \ldots \vee G_m$

Replace by

$r(x_1, \ldots, x_n) \leftrightarrow G_1 \vee \ldots \vee G_m$

The formula

$r(x_1, \ldots, x_n) \leftrightarrow G_1 \vee \ldots \vee G_m$

is called the **completed definition** of the original set of clauses.

Properties

- All steps **preserve Herbrand models**, except for the last one.
- Gives a **unique semantics to non-recursive definitions**;
- On non-recursive definitions is **equivalent to first-order logic**.