Advanced Data Modeling

Steffen Staab
with
Simon Schenk

Overview

- First-order logic. Syntax and semantics.
- Herbrand interpretations;
- Clauses and goals;
- Datalog.

First-order signature

First-order signature Σ consists of

- *con* — the set of constants of Σ;
- *fun* — the set of function symbols of Σ;
- *rel* — the set of relation symbols of Σ.

Term of Σ with variables in X:

1. Constant $c \in con$;
2. Variable $v \in X$;
3. If $f \in fun$ is a function symbol of arity n and $t_1, \ldots, t_n$ are terms, then $f(t_1, \ldots, t_n)$ is a term.

- A term is ground if it has no variables
- *var*(*t*) — the set of variables of *t*

Abstract notation:

- *a*, *b*, *c*, *d*, *e* for constants;
- *x*, *y*, *z*, *u*, *v*, *w* for variables;
- *f*, *g*, *h* for function symbols;
- *p*, *q* for relation symbols,

Example: *f*(*x*, *g*(*y*)).

Concrete notation: teletype font for everything.
Variable names start with upper-case letters.

Example: likes(john, Anybody).

Formulas

- Atomic formulas, or atoms $p(t_1, \ldots, t_n)$.
- $(A_1 \wedge \ldots \wedge A_n)$ and $(A_1 \vee \ldots \vee A_n)$
- $(A \rightarrow B)$ and $(A \leftrightarrow B)$
- $\neg A$
- $\forall v\ A$ and $\exists v\ A$

Substitutions

- Substitution θ *:* is any mapping from the set *V* of variables to the set of terms such that there is only a finite number of variables $v \in V$ with $\theta(v) \neq v$.
- Domain $dom(\theta)$, range $ran(\theta)$ and variable range $vran(\theta)$:
  - $dom(\theta) = \{v \mid v \neq \theta(v)\}$,
  - $ran(\theta) = \{ t \mid \exists v \in dom(\theta)(\theta(v) = t)\}$,
  - $vran(\theta) = var(ran(\theta))$.
- Notation: $\{ x_1 \mapsto t_1, \ldots , x_n \mapsto t_n \}$
- empty substitution {}

Application of a substitution θ to a term *t*:

- $x\theta = \theta(x)$
- $c\theta = c$
- $f(t_1, \dots, t_n)\theta = f(t_1\theta, \dots, t_n\theta)$

Herbrand interpretation

A Herbrand interpretation of a signature Σ is any set of ground atoms of this signature.

Truth in Herbrand Interpretations

1. If A is atomic, then $I \models A$ if $A \in I$
2. $I \models B_1 \wedge \ldots \wedge B_n$ if $I \models B_i$ for all i
3. $I \models B_1 \vee \ldots \vee B_n$ if $I \models B_i$ for some i
4. $I \models B_1 \rightarrow B_2$ if either $I \models B_2$ or $I \not\models B_1$
5. $I \models \neg B$ if $I \not\models B$
6. $I \models \forall xB$ if $I \models B\{x \mapsto t\}$ for all ground terms t of the signature Σ
7. $I \models \exists xB$ if $I \models B\{x \mapsto t\}$ for some ground term t of the signature Σ

Literals

- Literal is either an atom or the negation $\neg A$ of an atom A.
- Positive literal: atom
- Negative literal: negation of an atom
- Complimentary literals: A and $\neg A$
- Notation: L

Clause

Clause: (or normal clause) formula $L_1 \wedge \ldots \wedge L_n \rightarrow A$,
where

- $n \geq 0$, each L_i is a literal and A is an atom.

- Notation: $A \text{ :- } L_1 \wedge \ldots \wedge L_n$ or $A \text{ :- } L_1, \ldots, L_n$
- Head: the atom A.
- Body: The conjunction $L_1 \wedge \ldots \wedge L_n$
- Definite clause: all L_i are positive
- Fact: clause with empty body

Classification

Clause	Class
lives(Person, sweden) :- sells(Person, wine, Shop), not open(Shop,saturday)	normal
spy(Person) :- russian(Person)	definite
spy(bond)	fact

Goal

- Goal (also normal goal) is any conjunction of literals $L_1 \wedge \ldots \wedge L_n$
- Definite goal: all L_i are positive
- Empty goal □: when $n = 0$