Advanced Data Modeling

Steffen Staab

with

Simon Schenk
Overview

- First-order logic. Syntax and semantics.
- Herbrand interpretations;
- Clauses and goals;
- Datalog.
First-order signature $\Sigma$ consists of
- $con$ — the set of constants of $\Sigma$;
- $fun$ — the set of function symbols of $\Sigma$;
- $rel$ — the set of relation symbols of $\Sigma$. 
Terms

Term of $\Sigma$ 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
- $\text{var}(t)$ — the set of variables of $t$
Abstract and concrete notation

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 $p(t_1, \ldots, t_n)$.
- $(A_1 \land \ldots \land A_n)$ and $(A_1 \lor \ldots \lor A_n)$
- $(A \rightarrow B)$ and $(A \leftrightarrow B)$
- $\neg A$
- $\forall v A$ and $\exists v A$
Substitutions

- Substitution $\theta :$ 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 $\text{dom}(\theta)$, range $\text{ran}(\theta)$ and variable range $\text{vran}(\theta)$:
  - $\text{dom}(\theta) = \{ v \mid v \neq \theta(v) \}$,
  - $\text{ran}(\theta) = \{ t \mid \exists v \in \text{dom}(\theta)(\theta(v) = t) \}$,
  - $\text{vran}(\theta) = \text{var}(\text{ran}(\theta))$.
- Notation: $\{ x_1 \mapsto t_1, \ldots, x_n \mapsto t_n \}$
- empty substitution $\{\}$
Application of substitution $\theta$ to a term $t$:

- $x\theta = \theta(x)$
- $c\theta = c$
- $f(t_1, \ldots, t_n)\theta = f(t_1\theta, \ldots, t_n\theta)$
A Herbrand interpretation of a signature $\sum$ 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. \( I \models B_1 \land \ldots \land B_n \) if \( I \models B_i \) for all \( i \)
3. \( I \models B_1 \lor \ldots \lor B_n \) 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. \( I \models \forall x B \) if \( I \models B\{x \mapsto t\} \) for all ground terms \( t \) of the signature \( \Sigma \)
7. \( I \models \exists x B \) if \( I \models B\{x \mapsto t\} \) for some ground term \( t \) of the signature \( \Sigma \)
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: (or normal clause) formula $L_1 \land \ldots \land L_n \rightarrow A$, where

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

- Notation: $A :- L_1 \land \ldots \land L_n$ or $A :- L_1, \ldots, L_n$
- Head: the atom $A$.
- Body: The conjunction $L_1 \land \ldots \land L_n$
- Definite clause: all $L_i$ are positive
- Fact: clause with empty body
<table>
<thead>
<tr>
<th>Clause</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>lives(Person, sweden) :- sells(Person, wine, Shop), not open(Shop,saturday)</td>
<td>normal</td>
</tr>
<tr>
<td>spy(Person) :- russian(Person)</td>
<td>definite</td>
</tr>
<tr>
<td>spy(bond)</td>
<td>fact</td>
</tr>
</tbody>
</table>
Goal

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