FORMAL METHODS

LECTURE IX

MODEL CHECKING VS. PROOF THEORY

Alessandro Artale

Faculty of Computer Science – Free University of Bolzano

Room 2.03

artale@inf.unibz.it http://www.inf.unibz.it/∼artale/

Reading

Reiter, R. "Towards a Logical Reconstruction of Relational Database Theory". In *On Conceptual Modeling*. Springer, 1984.

DB and Logic: The Model View

- SQL is equivalent to *Relational Calculus* which is essentially a first order language.
- A Relational Calculus query is a formula of FOL that is *evaluated* with respect to a set of database facts.
  1. A DB can be viewed as a **first order interpretation**;
  2. The result of a query is the set of values that when substituted for the free variables of the query make the query true in the interpretation provided by the DB.

DB and Logic: The Proof View

1. A DB is viewed as a set of FOL formulas, i.e., a **first order Theory**
2. Queries are formulas to be *proven* given the DB as premises.
3. Reiter's Conclusions.
   (a) The Model and Proof paradigms can be reconciled;
   (b) The Proof view is reacher (Deductive DBs, DBs with incomplete information, etc...)

Relational First Order Language

Let R be a first order language over an alphabet Σ, then R is said a **relational language** if:

1. Σ has a finite number of constants and predicates;
2. Σ does not have function symbols;
3. One of the predicates in Σ is the *equality* binary predicate (we call R a FOL with equality);
4. Among the predicates there is a distinguished set of unary predicates called *Types* (capture the notion of attribute domains for relations).

Relational Interpretation

Let R be a *relational language* over an alphabet Σ, an interpretation $I = (\Delta, \cdot^I)$ is a **relational interpretation** for R if:

1. $\cdot^I$: *constants in* $\Sigma \mapsto \Delta$, is 1-1 and onto
2. $(=)^I = \{(d, d) \mid d \in \Delta\}$.

Relational Database

A **relational database** is a triple $DB = (\mathsf{R}, \mathcal{I}, IC)$ where:

1. R is a relational language;
2. *IC* is a set of formulas over R (Integrity Constraints) s.t. for all $P \in \Sigma$ (distinct from "=" and types) *IC* contains:

   $$\forall x_1, \ldots, x_n . P(x_1, \ldots, x_n) \rightarrow \tau_1(x_1) \wedge \ldots \wedge \tau_n(x_n)$$

   where τ_i are types (said the *domains* of P);
3. $\mathcal{I}$ is a relational interpretation for R satisfying *IC*.

Queries as Model Checking

Queries are defined w.r.t. a relational language R.

- Let $DB = (\mathsf{R}, \mathcal{I}, IC)$, then a query over DB is a formula $Q(x_1, \ldots, x_n)$ over R with $x_1, \ldots, x_n$ as the only free variables.
- The *answer set* of a query $Q(x_1, \ldots, x_n)$ is the set:

$$\{c_1, \ldots, c_n \in \Sigma \mid \mathcal{I} \models Q(c_1, \ldots, c_n)\}$$

Model Checking Vs. Query Answering. A tuple $(c_1, \ldots, c_n)$ belongs to the answer set of a query Q iff we can answer positively to the **Model Checking** problem:

$$\mathcal{I} \models Q(c_1, \ldots, c_n)$$

The Proof Theoretic View: Intro

- The *Model Checking* perspective on DBs can be reinterpreted in purely *Proof Theoretic* terms.
- **Main Idea:** Define a *First Order Theory*, Γ, called relational theory, and show an equivalence between such theories and relational interpretations, i.e.:

$$(\mathsf{R}, \mathcal{I}, IC) \equiv (\mathsf{R}, \Gamma, IC)$$

- Truth in the interpretation $\mathcal{I}$ will be reformulated in terms of provability in the theory Γ.

Relational Theories

Let R be a relational language with alphabet Σ and wff W.
$\Gamma \subseteq W$ is a relational theory of R iff:

- Domain Closure. If $c_1, \ldots, c_n$ are all of the constants in Σ, then Γ contains the axiom:

$$\forall x.(x = c_1 \vee \ldots \vee x = c_n)$$

- Unique Name Assumption. Γ contains the axiom:

$$\neg(c_i = c_j) \quad i, j = 1, \ldots, n \;\; i < j$$

- Atomic Assertions. Let $V \subseteq W$ a set of ground atomic formulas (equality is not considered here). Then:

$$V \subseteq \Gamma$$

Relational Theories (Cont.)

- Completion. Let $P \in \Sigma$ an m-ary predicate (different from "="), we define: $C_P = \{(c_1, \dots, c_m) \mid P(c_1, \dots, c_m) \in V\}$. Suppose that $C_P = \{(c_1^1, \dots, c_m^1), \dots (c_1^p, \dots, c_m^p)\}$. Then Γ contains the axiom for P:
$$\forall x_1, \dots, x_m . [P(x_1, \dots, x_m) \rightarrow (x_1 = c_1^1 \wedge \dots \wedge x_m = c_m^1) \vee \dots \vee (x_1 = c_1^p \wedge \dots \wedge x_m = c_m^p)]$$
If $C_P = \emptyset$ then the axiom is: $\forall x_1, \dots, x_m . \neg P(x_1, \dots, x_m)$
- Γ contains each of the following equality axioms:
  - Reflexivity. $\forall x.(x = x)$
  - Commutativity. $\forall x, y.(x = y) \rightarrow (y = x)$
  - Transitivity. $\forall x, y, z.(x = y) \wedge (y = z) \rightarrow (x = z)$
  - Leibnitz's principle of substitution. For each $P \in \Sigma$: $\forall x_1, \dots, x_m, y_1, \dots, y_m . [P(x_1, \dots, x_m) \wedge (x_1 = y_1) \wedge \dots \wedge (x_m = y_m) \rightarrow P(y_1, \dots, y_m)]$

Model Theory Vs Proof Theory

Theorem. Let R be a relational language. Then:

1. If Γ is a relational theory of R then Γ has a unique model which is a relational interpretation for R.
2. If $\mathcal{I}$ is a relational interpretation for R then there is a relational theory of R, Γ, such that $\mathcal{I}$ is the only model of Γ.

Corollary. Let Γ be a relational theory of a relational language R, and $\mathcal{I}$ be the model of Γ. Then, for any φ of R:

$$\mathcal{I} \models \varphi \quad \text{iff} \quad \Gamma \models \varphi$$