Logic and Databases

Enrico Franconi

(ILS-2014)

Queries

Since a query can be an arbitrary first-order formula, its answer may depend on the domain, which we do not know in advance or may vary from system to system.
For example:

- the query $Q(x) = \neg Student(x)$ over the database $Student(a), Student(b)$, with domain $\{a, b, c\}$ has the answer $\{x = c\}$,
- the same query with domain $\{a, b, c, d\}$ has the answer $\{x = c, x = d\}$.

Therefore, the notion of *domain independent* queries has been introduced in relational databases.

Domain Independence

A formula $\mathcal{Q}(\mathbb{X})$ is *domain independent with respect to the integrity constraints $\mathcal{IC}$*

if and only if

for every two models $\mathcal{I}$ and $\mathcal{J}$ of $\mathcal{IC}$ (i.e., $\mathcal{I} = \langle |\mathcal{I}|, \cdot^{\mathcal{I}} \rangle$ and $\mathcal{J} = \langle |\mathcal{J}|, \cdot^{\mathcal{J}} \rangle$) which agree on the interpretation of the predicates and constants (i.e. $\cdot^{\mathcal{I}} = \cdot^{\mathcal{J}}$), and for every assignment $v : \mathbb{X} \mapsto |\mathcal{I}| \cup |\mathcal{J}|$, we have:

$$rng(v) \subseteq |\mathcal{I}| \quad \text{and} \quad \mathcal{I} \models \mathcal{Q}(\mathbb{X})[v]$$

if and only if

$$rng(v) \subseteq |\mathcal{J}| \quad \text{and} \quad \mathcal{J} \models \mathcal{Q}(\mathbb{X})[v].$$

Examples

- $Q_1(x) = \neg A(x) \wedge B(x)$
- $Q_2(x) = \exists x.(A(x) \vee B(a))$
- $Q_3(x) = \neg A(x)$
- $Q_4(x) = \forall x.\, A(x)$

Check whether they are:

- domain independent,
- safe range.