

# CQA: Query Rewriting

Jan Chomicki

University at Buffalo and Warsaw University

March 22, 2007

## Query rewriting

Constraints in **clausal form** (disjunctions of literals).

### Residues

- associated with single literals  $p(\bar{x})$  or  $\neg p(\bar{x})$  (one of each for every database relation  $p$ )
- for each literal and each constraint that contains a complementary literal (after renaming), the **local residue** is obtained by removing the complementary literal and the quantifiers for its associated variables
- for each literal, **global residue** = conjunction of local residues.

### Functional dependencies

$$\begin{aligned} &(\forall x, y, z, y', z')(\neg E(x, y, z) \\ &\quad \vee \neg E(x, y', z') \vee y = y') \\ &(\forall x, y, z, y', z')(\neg E(x, y, z) \\ &\quad \vee \neg E(x, y', z') \vee z = z') \end{aligned}$$

### Query

$$E(x, y, z)$$

### Local residues

$$\begin{aligned} &(\forall y', z')(\neg E(x, y', z') \vee y = y') \\ &(\forall y', z')(\neg E(x, y', z') \vee z = z') \end{aligned}$$

## Constructing the rewritten query

### Literal expansion

For every literal in the original query, construct the expanded version as the conjunction of this literal and its global residue.

### Iteration

The expansion step is iterated by replacing the literals in the residue by their expanded versions, until no changes occur.

### Query expansion

Replace the literals in the query by their final expanded versions.

### Functional dependencies

$$(\forall x, y, z, y', z')(\neg E(x, y, z) \vee \neg E(x, y', z') \vee y = y')$$

$$(\forall x, y, z, y', z')(\neg E(x, y, z) \vee \neg E(x, y', z') \vee z = z')$$

### Query

$$E(x, y, z)$$

### Rewritten query

$$E(x, y, z) \wedge (\forall y', z')(\neg E(x, y', z') \vee y = y') \wedge (\forall y', z')(\neg E(x, y', z') \vee z = z')$$

## Iteration

### Integrity constraints

$$(\forall x)(\neg P(x) \vee R(x))$$

$$\forall x(\neg R(x) \vee S(x))$$

Literal	Residue	First expansion	Second (final) expansion
$R(x)$	$S(x)$	$R(x) \wedge S(x)$	$R(x) \wedge S(x)$
$P(x)$	$R(x)$	$P(x) \wedge R(x)$	$P(x) \wedge R(x) \wedge S(x)$
$\neg R(x)$	$\neg P(x)$	$\neg R(x) \wedge \neg P(x)$	$\neg R(x) \wedge \neg P(x)$
$\neg S(x)$	$\neg R(x)$	$\neg S(x) \wedge \neg R(x)$	$\neg S(x) \wedge \neg R(x) \wedge \neg P(x)$

### Scope of query rewriting

- queries involving conjunctions of literals (*relational algebra*:  $\sigma, \times, -$ ) and binary universal integrity constraints [ABC99]
- existentially-quantified conjunctions ( $\pi, \sigma, \times$ ) and single-key dependencies (under certain syntactic restrictions) [FM05]



M. Arenas, L. Bertossi, and J. Chomicki.

Consistent Query Answers in Inconsistent Databases.

In *ACM Symposium on Principles of Database Systems (PODS)*, pages 68–79, 1999.



A. Fuxman and R. J. Miller.

First-Order Query Rewriting for Inconsistent Databases.

In *International Conference on Database Theory (ICDT)*, pages 337–351. Springer, LNCS 3363, 2005.

Full version to appear in *JCSS*.