Locally Consistent Transformations and Query Answering in Data Exchange

Marcelo Arenas	Pablo Barceló	Ronald Fagin	Leonid Libkin
PUC Chile	U. of Toronto	IBM Almaden	U. of Toronto

Data exchange settings

Data Exchange Setting: $(\mathbf{S}, \mathbf{T}, \Sigma_{st})$

$\mathbf{S}$: Source schema.

$\mathbf{T}$: Target schema.

Σ_{st}: Set of source-to-target dependencies.

- Source-to-target dependency: FO sentence of the form

$$\forall \bar{x} \, (\varphi_{\mathbf{S}}(\bar{x}) \rightarrow \exists \bar{y} \, \psi_{\mathbf{T}}(\bar{x}, \bar{y})).$$

- $\varphi_{\mathbf{S}}(\bar{x})$: FO formula over $\mathbf{S}$.

- $\psi_{\mathbf{T}}(\bar{x}, \bar{y})$: conjunction of FO atomic formulas over $\mathbf{T}$.

Data exchange settings: Example

$$\mathbf{S} = \langle Employee(\cdot) \rangle$$

$$\mathbf{T} = \langle Dept(\cdot, \cdot) \rangle$$

$$\Sigma_{st} = \{\forall x\, (Employee(x) \rightarrow \exists y\, Dept(x, y))\}.$$

Data exchange problem

Given a source instance I, find a target instance J such that (I, J) satisfies Σ_{st}.

- J is called a solution for I.

Example: Possible solutions for $I = \{Employee(peter)\}$:

- $J_1 = \{Dept(peter, 1)\}$.
- $J_2 = \{Dept(peter, 1), Dept(peter, 2)\}$.
- $J_3 = \{Dept(peter, 1), Dept(john, 1)\}$.
- $J_4 = \{Dept(peter, X)\}$.
- $J_5 = \{Dept(peter, X), Dept(peter, Y)\}$.

Query answering

Q: Query over the target schema.

- What does it mean to answer Q?

$$\underline{\textit{certain}}(Q, I) \;=\; \bigcap_{J \text{ is a solution for } I} Q(J)$$

Example:

- $\underline{\textit{certain}}(\exists y \; Dept(x, y), \; I) = \{peter\}$.
- $\underline{\textit{certain}}(Dept(x, y), \; I) = \emptyset$.

Query rewriting

How can we compute $\underline{\textit{certain}}(Q, I)$?

- Naïve algorithm does not work: infinitely many solutions.

Approach proposed in [FKMP03]: **Query Rewriting**

Look for some specific $\mathcal{F} : \text{inst}(\mathbf{S}) \rightarrow \text{inst}(\mathbf{T})$, and find conditions under which $\underline{\textit{certain}}(Q, I) = Q'(\mathcal{F}(I))$ for every source instance I.

What is a good alternative for $\mathcal{F}$?

Outline

- Query rewriting over the canonical solution.
- Locality in data exchange.
  - Proving inexpressibility results.
- Query rewriting over the core.
  - Canonical solution versus core.
- Extensions.
  - Other semantics.
- Conclusions.

Outline

- Query rewriting over the canonical solution.
- Locality in data exchange.
  - Proving inexpressibility results.
- Query rewriting over the core.
  - Canonical solution versus core.
- Extensions.
  - Other semantics.
- Conclusions.

Canonical solution

Input: $(\mathbf{S}, \mathbf{T}, \Sigma_{st})$ and a source instance I

Output: Canonical solution J for I

Algorithm:

> for every $\forall \bar{x}\, (\varphi_{\mathbf{S}}(\bar{x}) \rightarrow \exists y\, \psi_{\mathbf{T}}(\bar{x}, \bar{y})) \in \Sigma_{st}$ do
>
> > for every $\bar{a}$ such that I satisfies $\varphi_{\mathbf{S}}(\bar{a})$ do
> >
> > > create a fresh tuple of null values $\overline{X}$
> > >
> > > insert $\psi_{\mathbf{T}}(\bar{a}, \overline{X})$ into J

Canonical solution: Example

$\Sigma_{st} = \{\forall x \, (Employee(x) \rightarrow \exists y \, Dept(x, y))\}$ and
$I = \{Employee(peter), Employee(john)\}$.

- For $a = peter$ do

  Create a fresh null value X

  Insert $Dept(peter, X)$ into J

- For $a = john$ do

  Create a fresh null value Y

  Insert $Dept(john, Y)$ into J

Canonical solution:

$$\{Dept(peter, X),\ Dept(john, Y)\}$$

Query rewriting over the canonical solution

$\mathcal{F}_{\text{can}}(I)$: canonical solution for I.

- Can be computed in polynomial time (data complexity).

Theorem [FKMP03]: For every data exchange setting and union of conjunctive queries Q, there exists Q' such that for every source instance I, *certain*$(Q, I) = Q'(\mathcal{F}_{\text{can}}(I))$.

- $C(x)$: holds whenever x is a constant.
- $Q'(x_1, \ldots, x_m) = C(x_1) \wedge \cdots \wedge C(x_m) \wedge Q(x_1, \ldots, x_m)$.

Query rewriting over the canonical solution

Can the theorem be extended to other classes of queries?

Theorem [FKMP03]: There exists a data exchange setting and a conjunctive query Q with one inequality such that Q is not FO-rewritable over $\mathcal{F}_{\text{can}}$.

- For every FO query Q', there exists an instance I such that *certain*$(Q, I) \neq Q'(\mathcal{F}_{\text{can}}(I))$.

We would like to study the query rewriting problem.

- We need some tools: How can we prove that a query is not FO-rewritable?

Query rewriting: Some facts

The problem of deciding whether an FO formula is FO-rewritable over $\mathcal{F}_{\mathrm{can}}$ is undecidable.

There exists other classes of queries that are FO-rewritable over the canonical solution.

- Every boolean query Q whose asymptotic probability is 0 is FO-rewritable: <u>*certain*</u>$(Q, I) = \mathit{false}$.

Outline

- Query rewriting over the canonical solution.
- Locality in data exchange.
  - Proving inexpressibility results.
- Query rewriting over the core.
  - Canonical solution versus core.
- Extensions.
  - Other semantics.
- Conclusions.

Locality in data exchange: Notation

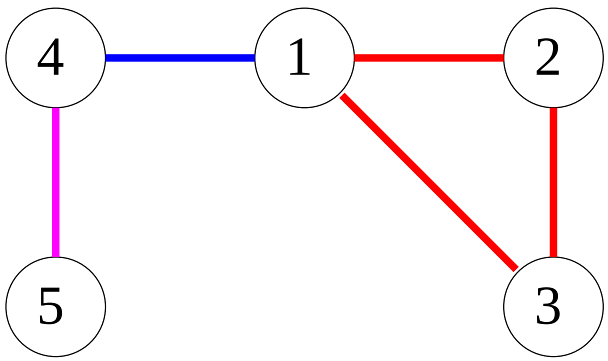
I: source instance.

Gaifman graph $\mathcal{G}(I)$ of I:

- adom(I) is the set of nodes of $\mathcal{G}(I)$.
- There exists an edge between a and b iff a and b belong to the same tuple of a relation in I.

Example: $I(R) = \{(1, 2, 3)\}$ and $I(T) = \{(1, 4), (4, 5)\}$.

$\mathcal{G}(I)$:



Locality in data exchange: Notation

$d_I(a,b)$: distance between a and b in $\mathcal{G}(I)$.

$d_I(\bar{a},b)$: minimum value of $d_I(a,b)$, where a is in $\bar{a}$.

$N_d^I(\bar{a})$: restriction of I to the elements at distance at most d from $\bar{a}$.

- Example: $\text{adom}(N_2^I(5)) = \{1,4,5\}$, $N_2^I(5)(R) = \emptyset$ and $N_2^I(5)(T) = \{(1,4),(4,5)\}$.

$N_d^I(\bar{a}) \cong N_d^I(\bar{b})$: members of $\bar{a}$ and $\bar{b}$ are treated as distinguished elements.

- $\bar{a} = (a_1,\ldots,a_m)$ and $\bar{b} = (b_1,\ldots,b_m)$.
- There is an isomorphism $f : N_d^I(\bar{a}) \rightarrow N_d^I(\bar{b})$ such that $f(a_i) = b_i$ ($1 \leq i \leq m$).

Locality in data exchange: Definition

Given: $(\mathbf{S}, \mathbf{T}, \Sigma_{st})$ and m-ary query Q over $\mathbf{T}$.

Definition: Q is **locally source-dependent** if there is $d \geq 0$ such that for every instance I of $\mathbf{S}$ and m-tuples $\bar{a}, \bar{b}$ in I,

$$N_d^I(\bar{a}) \cong N_d^I(\bar{b}) \quad \Longrightarrow \quad \begin{array}{c} \bar{a} \in \underline{\textit{certain}}(Q, I) \\ \text{iff} \\ \bar{b} \in \underline{\textit{certain}}(Q, I) \end{array}$$

Locality in data exchange: Main theorem

Theorem: If Q is FO-rewritable over the canonical solution, then Q is locally source-dependent.

This theorem can be used to prove inexpressibility results.

- If a query is not locally source-dependent, then it is not FO-rewritable.

Example: Proving inexpressibility

Data exchange setting:

$$\mathbf{S} = \langle G(\cdot,\cdot),\ R(\cdot),\ S(\cdot)\rangle$$

$$\mathbf{T} = \langle G'(\cdot,\cdot),\ R'(\cdot),\ S'(\cdot)\rangle$$

$$\begin{aligned}\Sigma_{st} = \ & \forall x \forall y\,(G(x,y) \rightarrow G'(x,y)),\\ & \forall x\,(R(x) \rightarrow R'(x)),\\ & \forall x\,(S(x) \rightarrow S'(x)).\end{aligned}$$

Query:

$$Q(x) = R'(x) \ \vee \ S'(x) \wedge \exists y \exists z (R'(y) \wedge G'(y,z) \wedge \neg R'(z))$$

Example: Proving inexpressibility

Assume that Q is FO-rewritable over the canonical solution.

Then there exists $d \geq 0$ such that

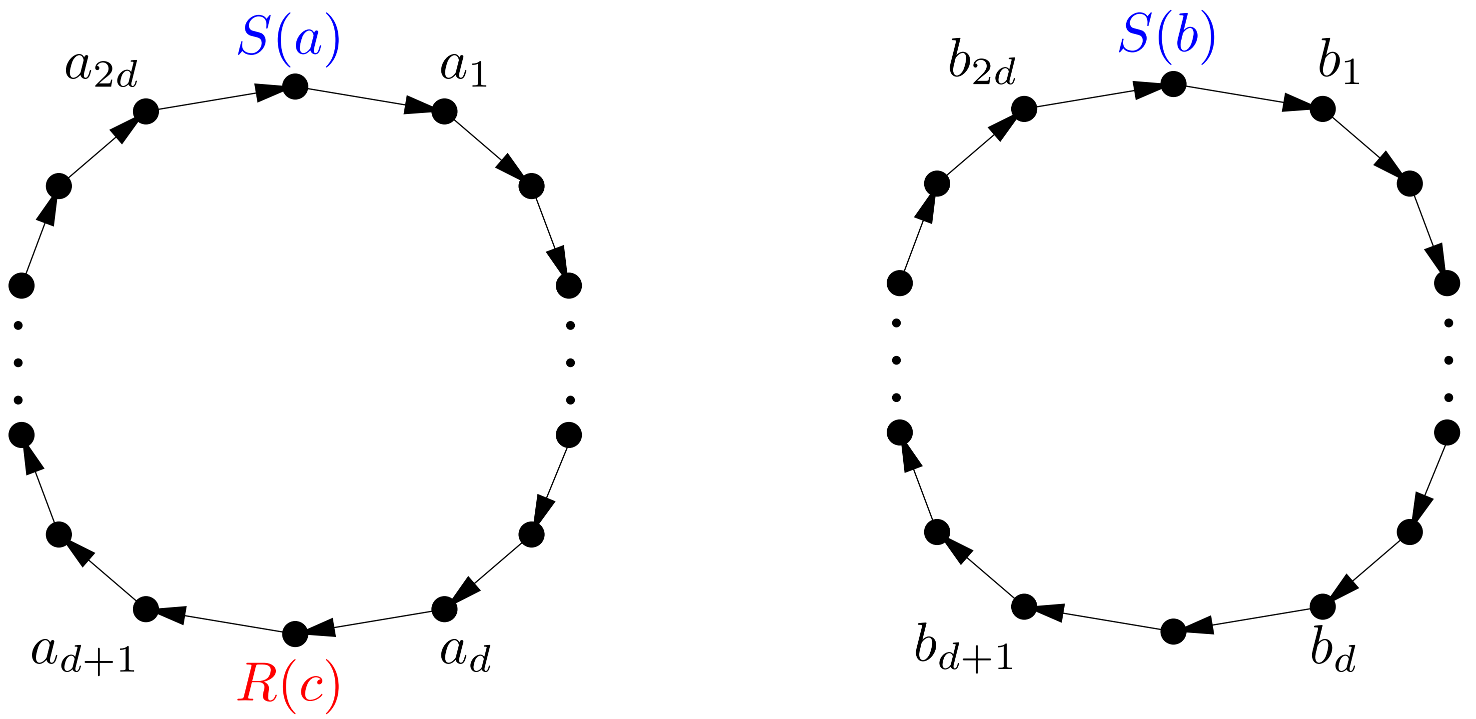
$$N_d^I(a) \cong N_d^I(b) \quad \Longrightarrow \quad a \in \underline{certain}(Q, I) \text{ iff } b \in \underline{certain}(Q, I).$$

Contradiction: find a source instance I such that

$$N_d^I(a) \cong N_d^I(b), \;\; a \in \underline{certain}(Q, I) \;\text{ and }\; b \notin \underline{certain}(Q, I).$$

Example: Defining instance I

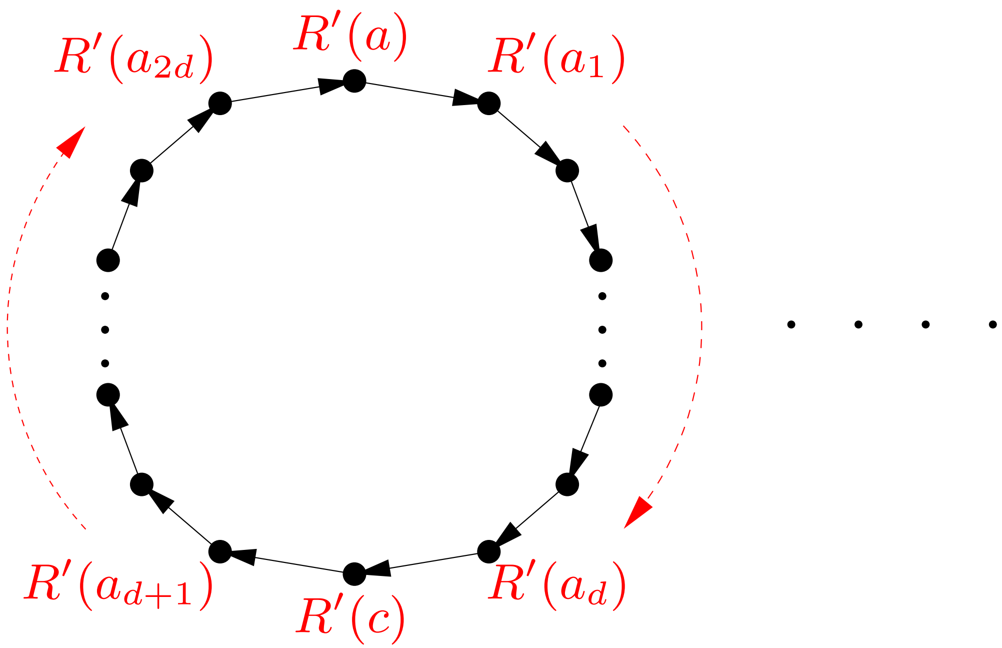
I:



Example: $a \in \underline{certain}(Q, I)$

If J does not satisfy $S'(a) \wedge \exists y \exists z (R'(y) \wedge G'(y, z) \wedge \neg R'(z))$:

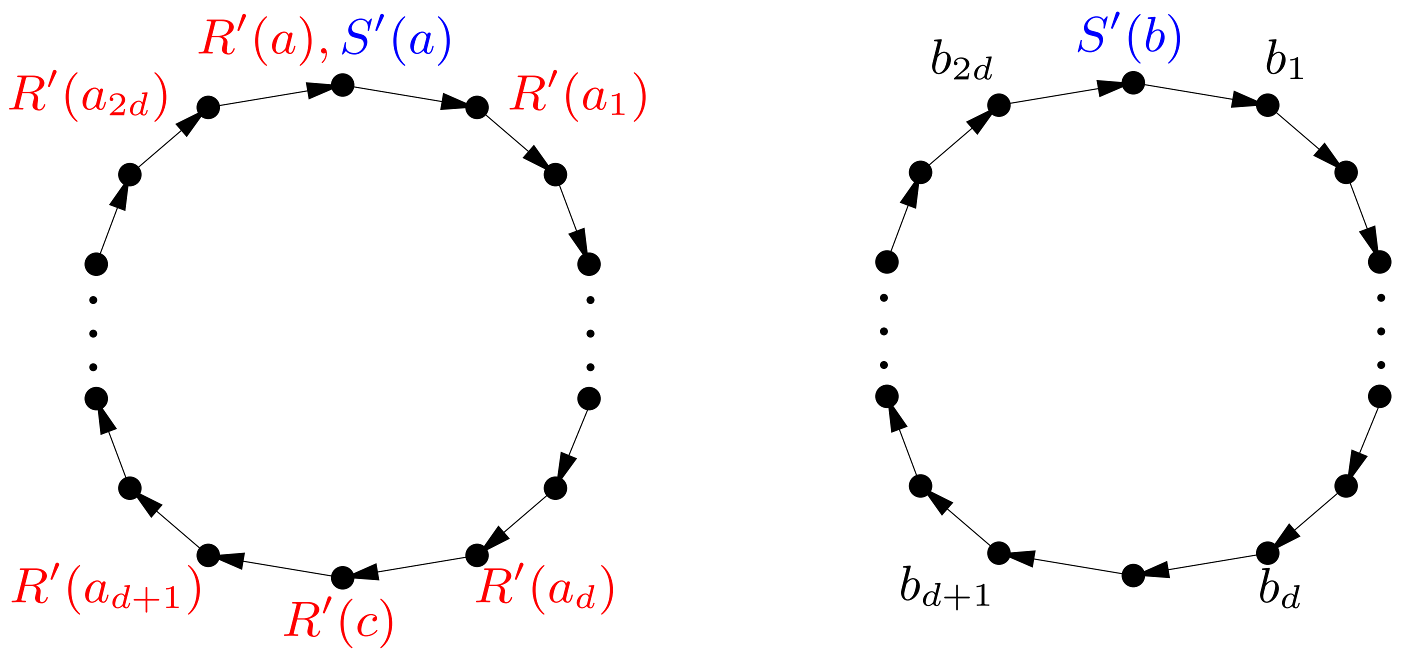
J:



Then: J satisfies $R'(a)$.

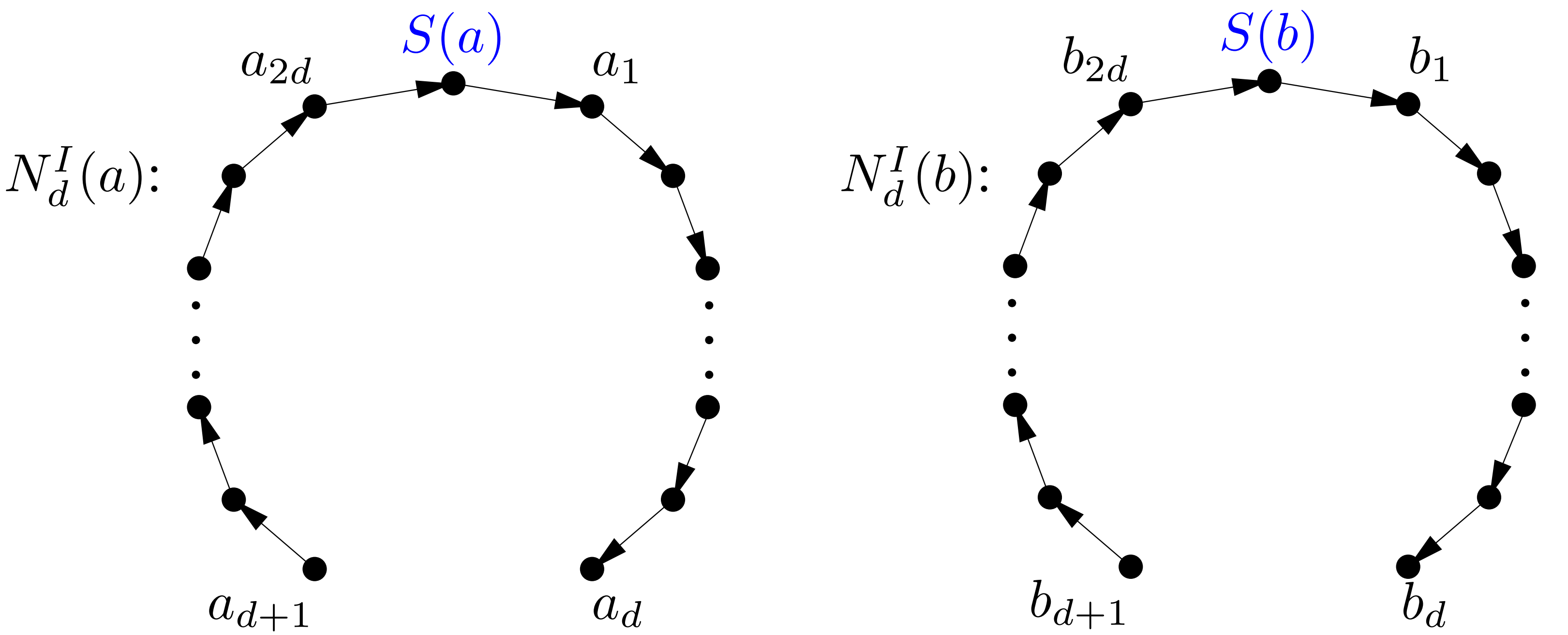
Example: $b \notin \underline{certain}(Q, I)$

J:



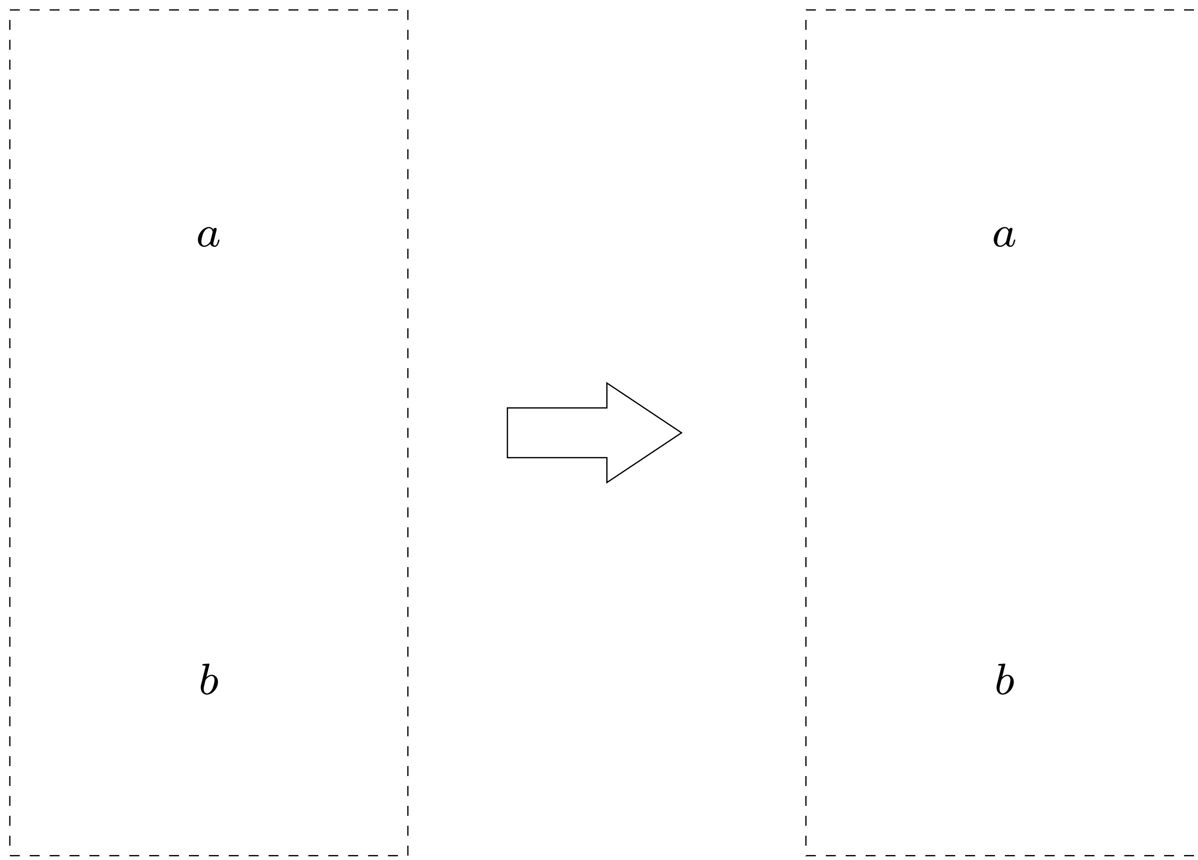
J does not satisfy $R'(b) \ \vee \ S'(b) \wedge \exists y \exists z (R'(y) \wedge G'(y, z) \wedge \neg R'(z))$.

Example: Getting a contradiction



Conclusion: Q is **not** FO-rewritable over the canonical solution.

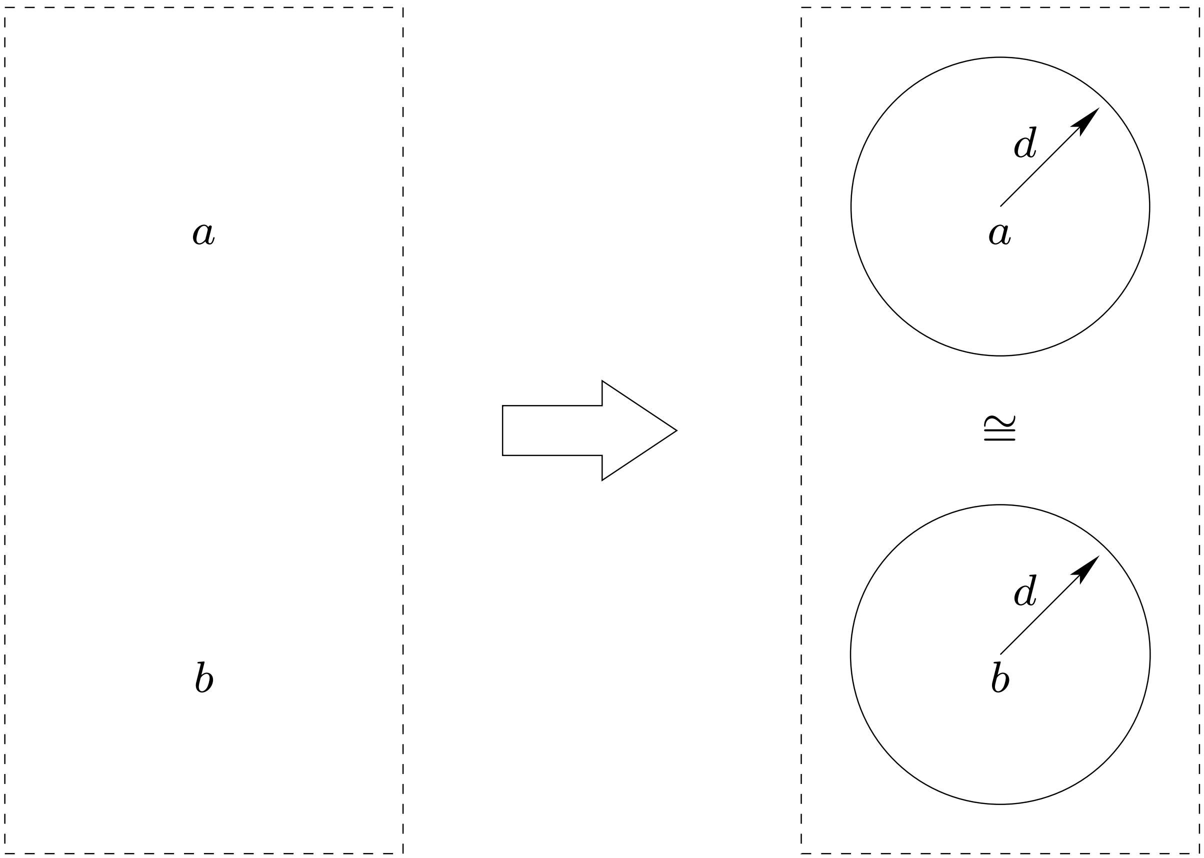
How do we prove the theorem?



source

canonical solution

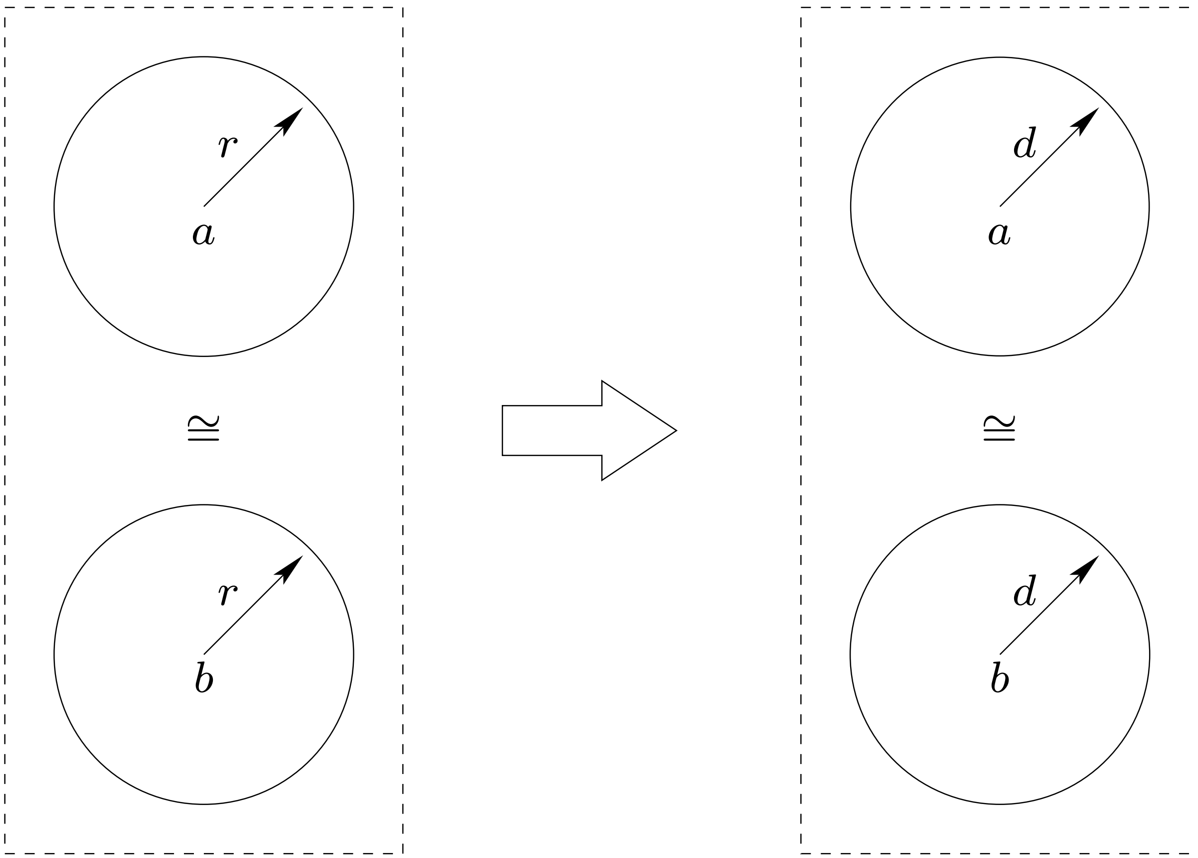
How do we prove the theorem?



source

canonical solution

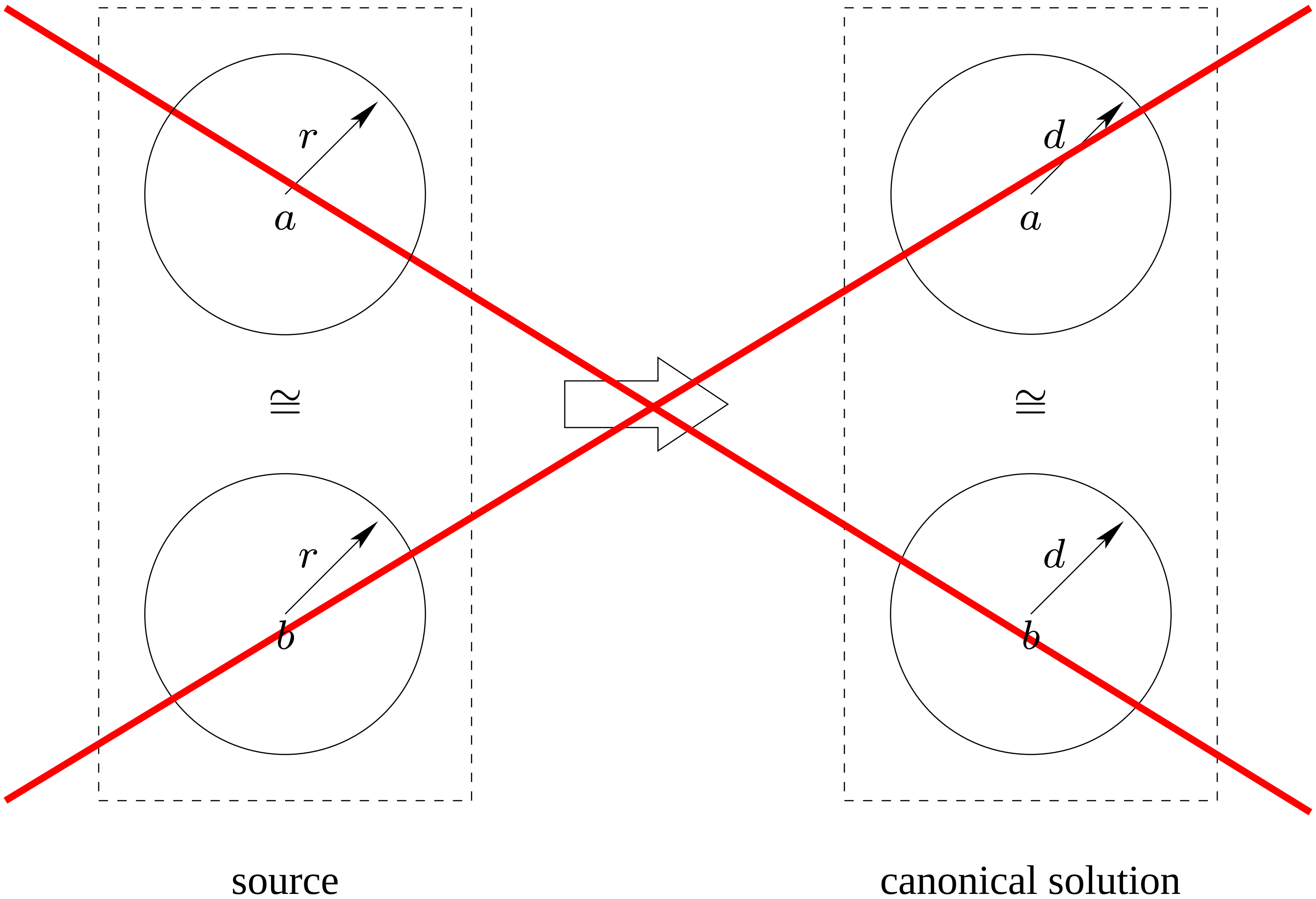
How do we prove the theorem?



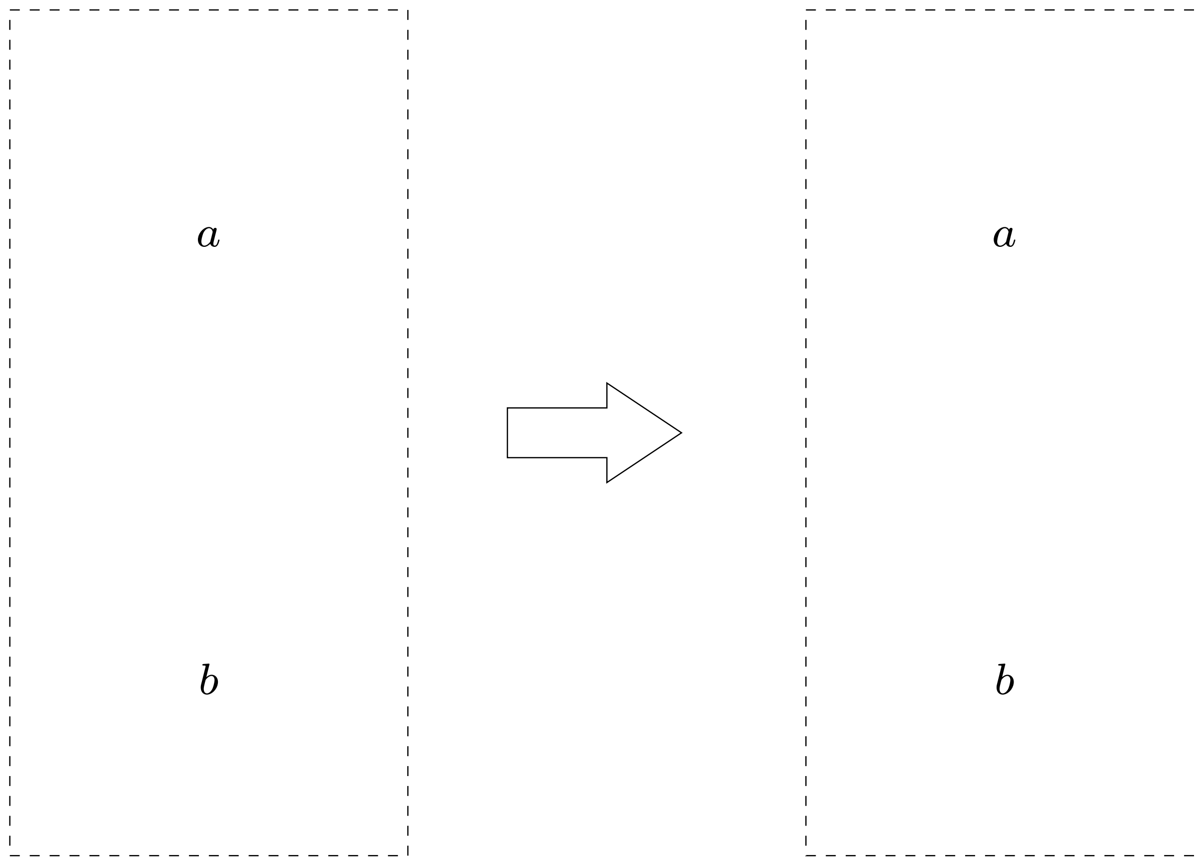
source

canonical solution

How do we prove the theorem?

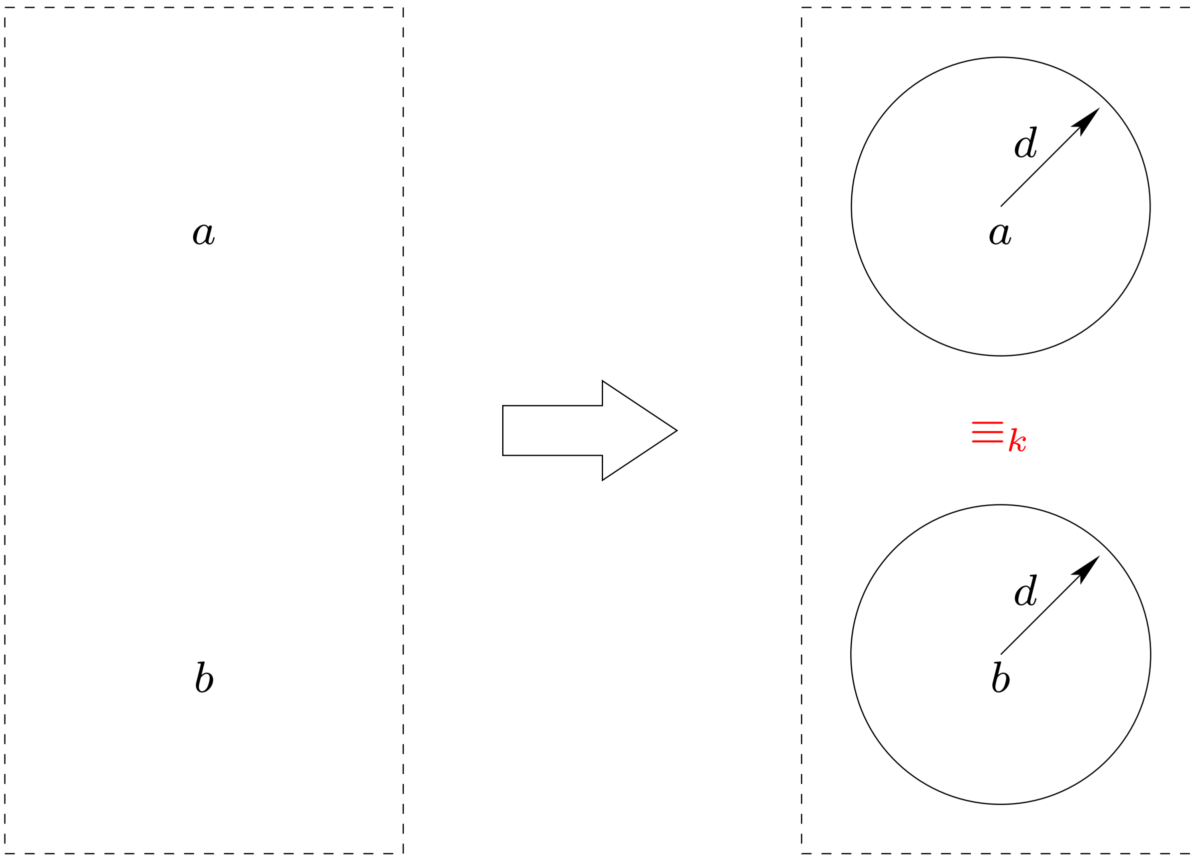


How do we prove the theorem?



source canonical solution

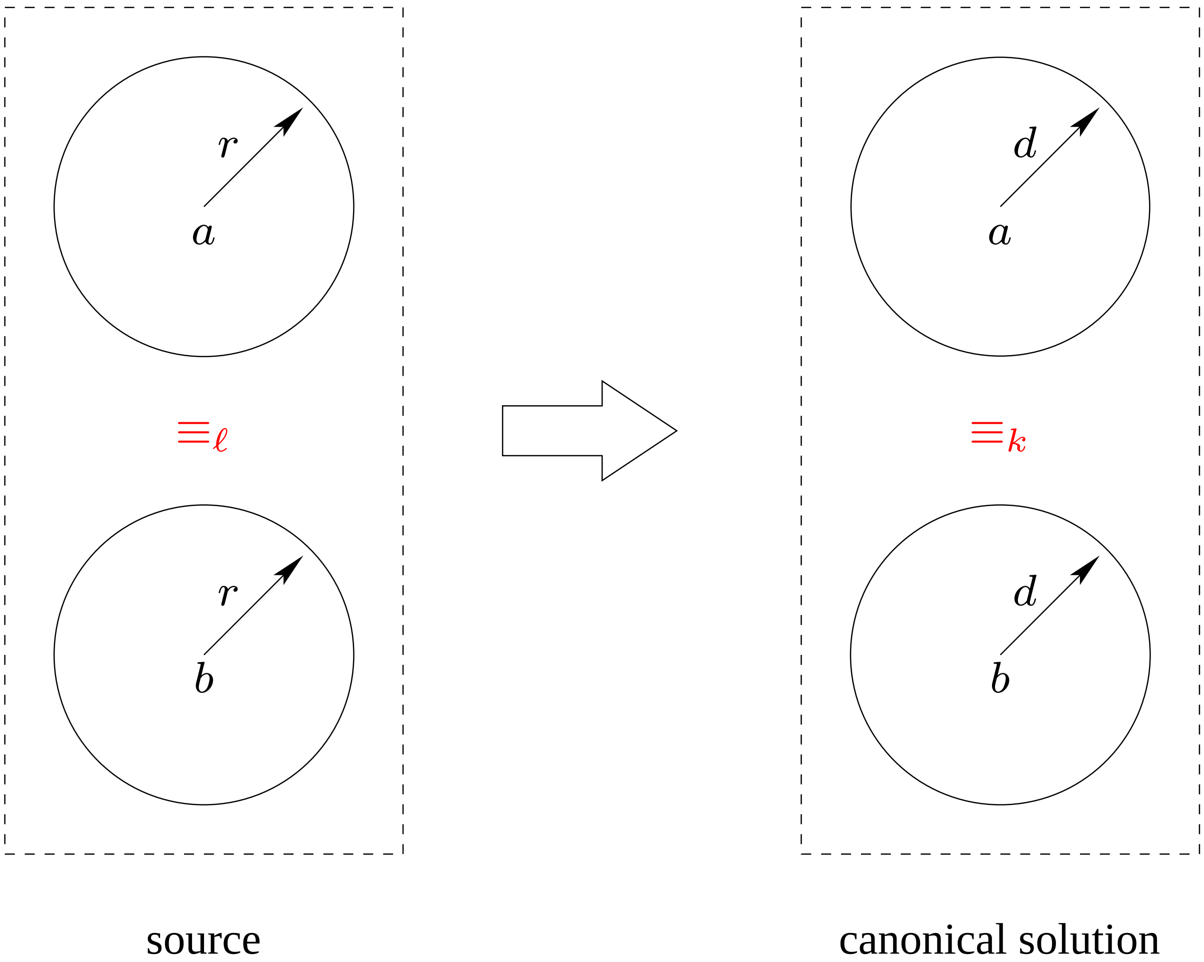
How do we prove the theorem?



source

canonical solution

How do we prove the theorem?



Outline

- Query rewriting over the canonical solution.
- Locality in data exchange.
  - Proving inexpressibility results.
- Query rewriting over the core.
  - Canonical solution versus core.
- Extensions.
  - Other semantics.
- Conclusions.

What about other transformations?

Core of canonical solution J: Substructure $J^\star$ of J such that there is a homomorphism from J to $J^\star$ and there is no homomorphism from J to a proper substructure of $J^\star$.

- Homomorphism $h : J \to J'$: mapping from $\text{adom}(J)$ to $\text{adom}(J')$ such that $h(c) = c$ for all constant c, and $\bar{t} \in J(R)$ implies $h(\bar{t}) \in J'(R)$.

Core is the smallest solution that is *homomorphically equivalent* to the canonical solution.

- It can be computed in polynomial time [FKP03].

Query rewriting over the core

$\mathcal{F}_{\text{core}}(I)$: core of the canonical solution for I.

Theorem [FKMP03]: For every data exchange setting and union conjunctive queries Q, there exists Q' such that for every source instance I, *certain*$(Q, I) = Q'(\mathcal{F}_{\text{core}}(I))$.

- Certain answers can be computed more efficiently by using the core.

Rewritability over the core: Can we use locality?

Canonical solution versus core: First attempt

Proposition: There exists a data exchange setting $\mathcal{A} = (\mathbf{S}, \mathbf{T}, \Sigma_{st})$ such that for every data exchange setting $\mathcal{B} = (\mathbf{S}, \mathbf{T}, \Gamma_{st})$, there exists instance I of $\mathbf{S}$ such that:

$$\mathcal{F}^{\mathcal{A}}_{\text{core}}(I) \quad \ncong \quad \mathcal{F}^{\mathcal{B}}_{\text{can}}(I).$$

We need a different approach ...

Expressiveness: Canonical solution versus core

Theorem: If Q is FO-rewritable over the core, then Q is also FO-rewritable over the canonical solution.

- There is a PTIME algorithm that, given a rewriting of Q over the core, finds a rewriting of Q over the canonical solution.

Corollary: If Q is FO-rewritable over the core, then Q is locally source-dependent.

Proof sketch

Assume $\varphi(\bar{x}) = \exists u \forall v\, \psi(\bar{x}, u, v)$ is a rewriting of Q over the core, where $\psi(\bar{x}, u, v)$ is quantifier-free.

- For every source instance I and tuple of constants $\bar{a}$: $\bar{a} \in$ *certain*(Q, I) iff $\mathcal{F}_{\text{core}}(I) \models \varphi(\bar{a})$.

Assume that:

$\alpha_1(x)$	:	holds if there is a core of $\mathcal{F}_{\text{can}}(I)$ containing null x.
$\alpha_2(x, y)$	:	holds if there is a core of $\mathcal{F}_{\text{can}}(I)$ containing nulls x and y.

Proof sketch

If $\alpha_1(x)$ and $\alpha_2(x, y)$ are FO-definable, then Q is FO-rewritable over the canonical solution:

$$\bar{a} \in \underline{\textit{certain}}(Q, I) \quad \text{iff} \quad \mathcal{F}_{\text{core}}(I) \models \exists u \forall v\, \varphi(\bar{a}, u, v)$$
$$\text{iff} \quad \mathcal{F}_{\text{can}}(I) \models \exists u\, (\alpha_1(u) \wedge \forall v\, (\alpha_2(u, v) \rightarrow \varphi(\bar{a}, u, v))).$$

How can we define $\alpha_1(x)$ and $\alpha_2(x, y)$ in FO?

- We show how to define $\alpha_1(x)$.

Proof sketch

Notation:

$\mathsf{nulls}(X, J)$	:	$\{Y \mid Y$ is a null of J and X, Y are in the same connected component of the graph induced from $\mathcal{G}(J)$ by the nulls of $J\}$
$\mathsf{block}(X, J)$	:	$\{t \mid t$ is a tuple in J containing a null in $\mathsf{nulls}(X, J)\}$

If J is a canonical solution: $|\mathsf{nulls}(X, J)|$ and $|\mathsf{block}(X, J)|$ are bounded.

Defining $\alpha_1(x)$

Lemma: Let J be the canonical solution for I and X a null value of J. There exists a core of J containing X iff for every pair of target structures J', J'' satisfying the following conditions:

- $J' \subseteq J$ and $|J'| \leq |\mathsf{block}(X, J)|$,
- there exists a homomorphism $h : \mathsf{block}(X, J) \rightarrow J'$ such that X is not a null of $h(\mathsf{block}(X, J))$,
- and $J' \subseteq J'' \subseteq \left(J' \cup \bigcup_{\{X \,|\, X \text{ is a null of } J'\}} \mathsf{block}(X, J)\right)$,

it is the case that there exists a homomorphism $h' : J'' \rightarrow J$ such that X is a null of $h'(J'')$.

Defining $\alpha_1(x)$

Lemma: Let J be the canonical solution for I and X a null value of J. There exists a core of J containing X iff for every pair of target structures J', J'' satisfying the following conditions:

- $J' \subseteq J$ and $|J'| \leq |\mathsf{block}(X, J)|$,
- there exists a homomorphism $h : \mathsf{block}(X, J) \rightarrow J'$ such that X is not a null of $h(\mathsf{block}(X, J))$,
- and $J' \subseteq J'' \subseteq \left(J' \cup \bigcup_{\{X \,|\, X \text{ is a null of } J'\}} \mathsf{block}(X, J)\right)$,

it is the case that there exists a homomorphism $h' : J'' \rightarrow J$ such that X is a null of $h'(J'')$.

Defining $\alpha_1(x)$

Lemma: Let J be the canonical solution for I and X a null value of J. There exists a core of J containing X iff for every pair of target structures J', J'' satisfying the following conditions:

- $J' \subseteq J$ and $|J'| \leq |\mathsf{block}(X, J)|$,
- there exists a homomorphism $h : \mathsf{block}(X, J) \rightarrow J'$ such that X is not a null of $h(\mathsf{block}(X, J))$,
- and $J' \subseteq J'' \subseteq \left(J' \cup \bigcup_{\{X \mid X \text{ is a null of } J'\}} \mathsf{block}(X, J) \right)$,

it is the case that there exists a homomorphism $h' : J'' \rightarrow J$ such that X is a null of $h'(J'')$.

Defining $\alpha_1(x)$

Lemma: Let J be the canonical solution for I and X a null value of J. There exists a core of J containing X iff for every pair of target structures J', J'' satisfying the following conditions:

- $J' \subseteq J$ and $|J'| \leq |\mathsf{block}(X, J)|$,
- there exists a homomorphism $h : \mathsf{block}(X, J) \to J'$ such that X is not a null of $h(\mathsf{block}(X, J))$,
- and $J' \subseteq J'' \subseteq \left(J' \cup \bigcup_{\{X \,|\, X \text{ is a null of } J'\}} \mathsf{block}(X, J) \right)$,

it is the case that there exists a homomorphism $h' : J'' \to J$ such that X is a null of $h'(J'')$.

Defining $\alpha_1(x)$

Lemma: Let J be the canonical solution for I and X a null value of J. There exists a core of J containing X iff for every pair of target structures J', J'' satisfying the following conditions:

- $J' \subseteq J$ and $|J'| \leq |\mathsf{block}(X, J)|$,
- there exists a homomorphism $h : \mathsf{block}(X, J) \to J'$ such that X is not a null of $h(\mathsf{block}(X, J))$,
- and $J' \subseteq J'' \subseteq \left(J' \cup \bigcup_{\{X \,|\, X \text{ is a null of } J'\}} \mathsf{block}(X, J)\right)$,

it is the case that there exists a homomorphism $h' : J'' \to J$ such that X is a null of $h'(J'')$.

Defining $\alpha_1(x)$

Lemma: Let J be the canonical solution for I and X a null value of J. There exists a core of J containing X iff for every pair of target structures J', J'' satisfying the following conditions:

- $J' \subseteq J$ and $|J'| \leq |\mathsf{block}(X, J)|$,

Is this definable in FO?

- and $J' \subseteq J'' \subseteq \left(J' \cup \bigcup_{\{X \,|\, X \text{ is a null of } J'\}} \mathsf{block}(X, J) \right)$,

it is the case that there exists a homomorphism $h' : J'' \rightarrow J$ such that X is a null of $h'(J'')$.

Defining $\alpha_1(x)$

Lemma: Let J be the canonical solution for I and X a null value of J. There exists a core of J containing X iff for every pair of target structures J', J'' satisfying the following conditions:

- $J' \subseteq J$ and $|J'| \leq |\mathsf{block}(X, J)|$,
- there exists a homomorphism $h : \mathsf{block}(X, J) \rightarrow J'$ such that X is not a null of $h(\mathsf{block}(X, J))$,
- and $J' \subseteq J'' \subseteq \left(J' \cup \bigcup_{\{X \mid X \text{ is a null of } J'\}} \mathsf{block}(X, J) \right)$,

it is the case that there exists a homomorphism $h' : J'' \rightarrow J$ such that X is a null of $h'(J'')$.

Defining $\alpha_1(x)$

Lemma: Let J be the canonical solution for I and X a null value of J. There exists a core of J containing X iff for every pair of target structures J', J'' satisfying the following conditions:

- $J' \subseteq J$ and $|J'| \leq |\text{block}(X, J)|$,
- there exists a homomorphism $h : \text{block}(X, J) \to J'$ such that X is not a null of $h(\text{block}(X, J))$,
- and $J' \subseteq J'' \subseteq \left(J' \cup \bigcup_{\{X \,|\, X \text{ is a null of } J'\}} \text{block}(X, J) \right)$,

it is the case that there exists a homomorphism $h' : J'' \to J$ such that X is a null of $h'(J'')$.

Defining $\alpha_1(x)$

Lemma: Let J be the canonical solution for I and X a null value of J. There exists a core of J containing X iff for every pair of target structures J', J'' satisfying the following conditions:

- $J' \subseteq J$ and $|J'| \leq |\mathsf{block}(X, J)|$,
- there exists a homomorphism $h : \mathsf{block}(X, J) \to J'$ such that X is not a null of $h(\mathsf{block}(X, J))$,
- and $J' \subseteq J'' \subseteq \left(J' \cup \bigcup_{\{X \,|\, X \text{ is a null of } J'\}} \mathsf{block}(X, J) \right)$,

it is the case that there exists a homomorphism $h' : J'' \to J$ such that X is a null of $h'(J'')$.

Defining $\alpha_1(x)$

Lemma: Let J be the canonical solution for I and X a null value of J. There exists a core of J containing X iff for every pair of target structures J', J'' satisfying the following conditions:

- $J' \subseteq J$ and $|J'| \leq |\mathsf{block}(X, J)|$,
- there exists a homomorphism $h : \mathsf{block}(X, J) \rightarrow J'$ such that X is not a null of $h(\mathsf{block}(X, J))$,
- and $J' \subseteq J'' \subseteq \left(J' \cup \bigcup_{\{X \,|\, X \text{ is a null of } J'\}} \mathsf{block}(X, J)\right)$,

it is the case that there exists a homomorphism $h' : J'' \rightarrow J$ such that X is a null of $h'(J'')$.

Defining $\alpha_1(x)$

Lemma: Let J be the canonical solution for I and X a null value of J. There exists a core of J containing X iff for every pair of target structures J', J'' satisfying the following conditions:

- $J' \subseteq J$ and $|J'| \leq |\mathsf{block}(X, J)|$,
- there exists a homomorphism $h : \mathsf{block}(X, J) \rightarrow J'$ such that X is not a null of $h(\mathsf{block}(X, J))$,
- and $J' \subseteq J'' \subseteq \left(J' \cup \bigcup_{\{X \mid X \text{ is a null of } J'\}} \mathsf{block}(X, J) \right)$,

it is the case that there exists a homomorphism $h' : J'' \rightarrow J$ such that X is a null of $h'(J'')$.

Expressiveness: Canonical solution versus core

Theorem: There exists an FO query that is FO-rewritable over the canonical solution but not over the core.

Expressiveness point of view: Canonical solution is better than the core.

- Canonical solution contains more information than the core.

Outline

- Query rewriting over the canonical solution.
- Locality in data exchange.
  - Proving inexpressibility results.
- Query rewriting over the core.
  - Canonical solution versus core.
- Extensions.
  - Other semantics.
- Conclusions.

What about other semantics?

Usual certain answers semantics sometimes exhibit counterintuitive behavior.

Good solutions: Universal solutions.

- Homomorphically equivalent to the canonical solution.

May be more meaningful to consider semantics based on universal solutions:

$$\underline{\textit{u-certain}}(Q, I) \;=\; \bigcap_{J \text{ is a universal solution for } I} Q(J).$$

Query rewriting under the universal solutions semantics

Given query Q, we want to find Q' such that
u-certain$(Q, I) = Q'(\mathcal{F}(I))$ for every source instance I.

Theorem [FKP03]: For every data exchange setting and existential query Q, there exists Q' such that for every source instance I, *u-certain*$(Q, I) = Q'(\mathcal{F}_{\text{core}}(I))$.

Query rewriting under the universal solutions semantics

Definition: Q is locally source-dependent under the universal solution semantics if there is $d \geq 0$ such that:

$$N_d^I(\bar{a}) \cong N_d^I(\bar{b}) \quad \Longrightarrow \quad \begin{array}{c} \bar{a} \in \underline{\textit{u-certain}}(Q, I) \\ \text{iff} \\ \bar{b} \in \underline{\textit{u-certain}}(Q, I) \end{array}$$

Theorem: All the previous results hold for the universal solution semantics.

- If Q is FO-rewritable over the canonical solution (core) under the universal solutions semantics, then Q is locally source-dependent under the universal solutions semantics.

What about target constraints?

Locality is no longer valid.

tgd: Even with a single full tgd.

$$\forall x \forall y \forall z \, (R(x,y) \wedge R(y,z) \rightarrow R(x,z)).$$

egd: Even for key dependencies.

Except for GAV settings: $\forall \bar{x} \, (\varphi_{\mathbf{S}}(\bar{x}) \rightarrow T(\bar{x}))$.

Conclusions

- Common data exchange transformations map similar neighborhoods into similar neighborhoods.

- This propertity can be used to formulate a locality notion for the canonical solution and the core.

- Locality can be used to prove that a query is not FO-rewritable.
  - Holds for other semantics.

- Expressiveness point of view: Canonical solution is better than the core.