Tractable Lineages on Treelike Instances: Limits and Extensions

Antoine Amarilli[1], Pierre Bourhis[2], Pierre Senellart[1,3]

June 29th, 2016

[1]Télécom ParisTech

[2]CNRS CRIStAL

[3]National University of Singapore

Tuple independent databases (TID)

- Probabilistic databases: model uncertainty about data
- Simplest model: tuple independent databases (TID)
  - A relational database I
  - A probability valuation π mapping each fact of I to $[0, 1]$
- Semantics of a TID (I, π): a probability distribution on $I' \subseteq I$:
  - Each fact $F \in I$ is either present or absent with probability $\pi(F)$
  - Assume independence across facts

Example: TID

S		
a	*a*	1
b	*v*	.5
b	*w*	.2

Example: TID

S		
a	a	1
b	v	.5
b	w	.2

This TID (I, π) represents the following probability distribution:

Example: TID

S		
a	a	1
b	v	.5
b	w	.2

This TID (I, π) represents the following probability distribution:

$.5 \times .2$

S	
a	a
b	v
b	w

Example: TID

S		
a	a	1
b	v	.5
b	w	.2

This TID (I, π) represents the following probability distribution:

$.5 \times .2$

S	
a	a
b	v
b	w

$.5 \times (1 - .2)$

S	
a	a
b	v

Example: TID

S		
a	a	1
b	v	.5
b	w	.2

This TID (I, π) represents the following probability distribution:

$.5 \times .2$

S	
a	a
b	v
b	w

$.5 \times (1 - .2)$

S	
a	a
b	v

$(1 - .5) \times .2$

S	
a	a
b	w

Example: TID

S		
a	a	1
b	v	.5
b	w	.2

This TID (I, π) represents the following probability distribution:

$.5 \times .2$

S	
a	a
b	v
b	w

$.5 \times (1 - .2)$

S	
a	a
b	v

$(1 - .5) \times .2$

S	
a	a
b	w

$(1 - .5) \times (1 - .2)$

S	
a	a

Probabilistic query evaluation (PQE)

Let us fix:

- Relational signature σ
- Class $\mathcal{I}$ of relational instances on σ (e.g., acyclic, treelike)
- Class $\mathcal{Q}$ of Boolean queries (e.g., CQs, acyclic CQs)

Probabilistic query evaluation (PQE)

Let us fix:

- Relational signature σ
- Class $\mathcal{I}$ of relational instances on σ (e.g., acyclic, treelike)
- Class $\mathcal{Q}$ of Boolean queries (e.g., CQs, acyclic CQs)

Probabilistic query evaluation (PQE) problem for $\mathcal{Q}$ and $\mathcal{I}$:

- Fix a query $q \in \mathcal{Q}$
- Read an instance $I \in \mathcal{I}$ and a probability valuation π
- Compute the probability that (I, π) satisfies q
- Data complexity: measured as a function of (I, π)

Example: PQE

Signature σ, class $\mathcal{Q}$ of conjunctive queries, class $\mathcal{I}$ of all instances.

Example: PQE

Signature σ, class $\mathcal{Q}$ of conjunctive queries, class $\mathcal{I}$ of all instances.

$$q : \exists x\, y\; R(x) \wedge S(x, y) \wedge T(y)$$

Example: PQE

Signature σ, class $\mathcal{Q}$ of conjunctive queries, class $\mathcal{I}$ of all instances.

$$q : \exists x\, y\; R(x) \wedge S(x,y) \wedge T(y)$$

R	
a	1
b	.4
c	.6

Example: PQE

Signature σ, class $\mathcal{Q}$ of conjunctive queries, class $\mathcal{I}$ of all instances.

$$q : \exists x\, y\; R(x) \land S(x, y) \land T(y)$$

R	
a	1
b	.4
c	.6

S		
a	a	1
b	v	.5
b	w	.2

Example: PQE

Signature σ, class $\mathcal{Q}$ of conjunctive queries, class $\mathcal{I}$ of all instances.

$$q : \exists x\, y\ R(x) \wedge S(x, y) \wedge T(y)$$

R	
a	1
b	.4
c	.6

S		
a	*a*	1
b	*v*	.5
b	*w*	.2

T	
v	.3
w	.7
b	1

Example: PQE

Signature σ, class $\mathcal{Q}$ of conjunctive queries, class $\mathcal{I}$ of all instances.

$$q : \exists x\, y\; R(x) \wedge S(x, y) \wedge T(y)$$

R	
a	1
b	.4
c	.6

S		
a	a	1
b	v	.5
b	w	.2

T	
v	.3
w	.7
b	1

Example: PQE

Signature σ, class $\mathcal{Q}$ of conjunctive queries, class $\mathcal{I}$ of all instances.

$$q : \exists x\, y\; R(x) \wedge S(x, y) \wedge T(y)$$

R	
a	1
b	.4
c	.6

S		
a	a	1
b	v	.5
b	w	.2

T	
v	.3
w	.7
b	1

Example: PQE

Signature σ, class $\mathcal{Q}$ of conjunctive queries, class $\mathcal{I}$ of all instances.

$$q : \exists x\, y\ R(x) \wedge S(x,y) \wedge T(y)$$

R	
a	1
b	.4
c	.6

S		
a	a	1
b	v	.5
b	w	.2

T	
v	.3
w	.7
b	1

- The query is true iff $R(b)$ is here and one of:

Example: PQE

Signature σ, class $\mathcal{Q}$ of conjunctive queries, class $\mathcal{I}$ of all instances.

$$q : \exists x\, y\ R(x) \wedge S(x, y) \wedge T(y)$$

R	
a	1
b	.4
c	.6

S		
a	a	1
b	v	.5
b	w	.2

T	
v	.3
w	.7
b	1

- The query is true iff $R(b)$ is here and one of:
  - $S(b, v)$ and $T(v)$ are here

Example: PQE

Signature σ, class $\mathcal{Q}$ of conjunctive queries, class $\mathcal{I}$ of all instances.

$$q : \exists x\, y\; R(x) \wedge S(x, y) \wedge T(y)$$

R	
a	1
b	.4
c	.6

S		
a	a	1
b	v	.5
b	w	.2

T	
v	.3
w	.7
b	1

- The query is true iff $R(b)$ is here and one of:
  - $S(b, v)$ and $T(v)$ are here
  - $S(b, w)$ and $T(w)$ are here

Example: PQE

Signature σ, class $\mathcal{Q}$ of conjunctive queries, class $\mathcal{I}$ of all instances.

$$q : \exists x\, y\ R(x) \wedge S(x, y) \wedge T(y)$$

R	
a	1
b	.4
c	.6

S		
a	a	1
b	v	.5
b	w	.2

T	
v	.3
w	.7
b	1

- The query is true iff $R(b)$ is here and one of:
  - $S(b, v)$ and $T(v)$ are here
  - $S(b, w)$ and $T(w)$ are here

$\rightarrow$ Probability:

Example: PQE

Signature σ, class $\mathcal{Q}$ of conjunctive queries, class $\mathcal{I}$ of all instances.

$$q : \exists x\, y\; R(x) \wedge S(x, y) \wedge T(y)$$

R	
a	1
b	.4
c	.6

S		
a	a	1
b	v	.5
b	w	.2

T	
v	.3
w	.7
b	1

- The query is true iff $R(b)$ is here and one of:
    - $S(b, v)$ and $T(v)$ are here
    - $S(b, w)$ and $T(w)$ are here

$\rightarrow$ Probability: $.4 \times$

Example: PQE

Signature σ, class $\mathcal{Q}$ of conjunctive queries, class $\mathcal{I}$ of all instances.

$$q : \exists x\, y\; R(x) \wedge S(x, y) \wedge T(y)$$

R			S				T	
a	1		a	a	1		v	.3
b	.4		b	v	.5		w	.7
c	.6		b	w	.2		b	1

- The query is true iff $R(b)$ is here and one of:
  - $S(b, v)$ and $T(v)$ are here
  - $S(b, w)$ and $T(w)$ are here

$\rightarrow$ Probability: $.4 \times (1 -$

Example: PQE

Signature σ, class $\mathcal{Q}$ of conjunctive queries, class $\mathcal{I}$ of all instances.

$$q : \exists x\, y\; R(x) \wedge S(x, y) \wedge T(y)$$

R	
a	1
b	.4
c	.6

S		
a	a	1
b	v	.5
b	w	.2

T	
v	.3
w	.7
b	1

- The query is true iff $R(b)$ is here and one of:
  - $S(b, v)$ and $T(v)$ are here
  - $S(b, w)$ and $T(w)$ are here

$\rightarrow$ Probability: $.4 \times (1 - (1 - .5 \times .3)$

Example: PQE

Signature σ, class $\mathcal{Q}$ of conjunctive queries, class $\mathcal{I}$ of all instances.

$$q : \exists x\, y\; R(x) \wedge S(x, y) \wedge T(y)$$

R	
a	1
b	.4
c	.6

S		
a	a	1
b	v	.5
b	w	.2

T	
v	.3
w	.7
b	1

- The query is true iff $R(b)$ is here and one of:
  - $S(b, v)$ and $T(v)$ are here
  - $S(b, w)$ and $T(w)$ are here

→ Probability: $.4 \times (1 - (1 - .5 \times .3) \times (1 - .2 \times .7))$

Example: PQE

Signature σ, class $\mathcal{Q}$ of conjunctive queries, class $\mathcal{I}$ of all instances.

$$q : \exists x\, y\; R(x) \wedge S(x, y) \wedge T(y)$$

R	
a	1
b	.4
c	.6

S		
a	a	1
b	v	.5
b	w	.2

T	
v	.3
w	.7
b	1

- The query is true iff $R(b)$ is here and one of:
  - $S(b, v)$ and $T(v)$ are here
  - $S(b, w)$ and $T(w)$ are here

$\rightarrow$ Probability: $.4 \times (1 - (1 - .5 \times .3) \times (1 - .2 \times .7)) = .1076$

Complexity of probabilistic query evaluation (PQE)

Question: what is the complexity of PQE
depending on the class $\mathcal{Q}$ of queries and class $\mathcal{I}$ of instances?

Complexity of probabilistic query evaluation (PQE)

Question: what is the complexity of PQE
depending on the class $\mathcal{Q}$ of queries and class $\mathcal{I}$ of instances?

- Existing dichotomy result on $\mathcal{Q}$ [Dalvi and Suciu, 2012]
  - $\mathcal{Q}$ are unions of CQs, $\mathcal{I}$ is all instances
  - There is a class $\mathcal{S} \subseteq \mathcal{Q}$ of safe queries

Complexity of probabilistic query evaluation (PQE)

Question: what is the complexity of PQE
depending on the class $\mathcal{Q}$ of queries and class $\mathcal{I}$ of instances?

- Existing dichotomy result on $\mathcal{Q}$ [Dalvi and Suciu, 2012]
  - $\mathcal{Q}$ are unions of CQs, $\mathcal{I}$ is all instances
  - There is a class $\mathcal{S} \subseteq \mathcal{Q}$ of safe queries

$\rightarrow$ PQE is PTIME for any $q \in \mathcal{S}$ on all instances

Complexity of probabilistic query evaluation (PQE)

Question: what is the complexity of PQE
depending on the class $\mathcal{Q}$ of queries and class $\mathcal{I}$ of instances?

- Existing dichotomy result on $\mathcal{Q}$ [Dalvi and Suciu, 2012]
  - $\mathcal{Q}$ are unions of CQs, $\mathcal{I}$ is all instances
  - There is a class $\mathcal{S} \subseteq \mathcal{Q}$ of safe queries

  $\rightarrow$ PQE is PTIME for any $q \in \mathcal{S}$ on all instances

  $\rightarrow$ PQE is #P-hard for any $q \in \mathcal{Q} \backslash \mathcal{S}$ on all instances

Complexity of probabilistic query evaluation (PQE)

Question: what is the complexity of PQE
depending on the class $\mathcal{Q}$ of queries and class $\mathcal{I}$ of instances?

- Existing dichotomy result on $\mathcal{Q}$ [Dalvi and Suciu, 2012]
  - $\mathcal{Q}$ are unions of CQs, $\mathcal{I}$ is all instances
  - There is a class $\mathcal{S} \subseteq \mathcal{Q}$ of safe queries
  - $\rightarrow$ PQE is PTIME for any $q \in \mathcal{S}$ on all instances
  - $\rightarrow$ PQE is #P-hard for any $q \in \mathcal{Q} \backslash \mathcal{S}$ on all instances
  - $q : \exists x\, y\, R(x) \wedge S(x, y) \wedge T(y)$ is unsafe!

Instance-based dichotomy

$\rightarrow$ We show a dichotomy result on $\mathcal{I}$ instead:

- $\mathcal{Q}$ are Boolean monadic second-order queries (includes all UCQs)

Instance-based dichotomy

$\rightarrow$ We show a dichotomy result on $\mathcal{I}$ instead:

- $\mathcal{Q}$ are Boolean monadic second-order queries (includes all UCQs)

$\rightarrow$ PQE is in linear time for $\mathcal{Q}$ on any bounded-treewidth class $\mathcal{I}$ up to arithmetic costs [Same authors, ICALP'15]

Instance-based dichotomy

→ We show a dichotomy result on $\mathcal{I}$ instead:

- $\mathcal{Q}$ are Boolean monadic second-order queries (includes all UCQs)

→ PQE is in linear time for $\mathcal{Q}$ on any bounded-treewidth class $\mathcal{I}$ up to arithmetic costs [Same authors, ICALP'15]

→ **This talk:** PQE is intractable for $\mathcal{Q}$ on any unbounded-treewidth class $\mathcal{I}$ (under some assumptions)

Instance-based dichotomy

$\rightarrow$ We show a dichotomy result on $\mathcal{I}$ instead:

- $\mathcal{Q}$ are Boolean monadic second-order queries (includes all UCQs)

$\rightarrow$ PQE is in linear time for $\mathcal{Q}$ on any bounded-treewidth class $\mathcal{I}$ up to arithmetic costs [Same authors, ICALP'15]

$\rightarrow$ **This talk:** PQE is intractable for $\mathcal{Q}$ on any unbounded-treewidth class $\mathcal{I}$ (under some assumptions)

$\rightarrow$ Treewidth measures how much data is close to a tree

- Trees have treewidth 1
- Cycles have treewidth 2
- k-cliques and $(k-1)$-grids have treewidth $k-1$

Main result: hardness of PQE

Theorem

For any arity-2 signature σ, there is a first-order query q such that for any constructible unbounded-treewidth class $\mathcal{I}$, the PQE problem for $\mathcal{Q} = \{q\}$ and $\mathcal{I}$ is #P-hard under RP reductions

- Arity-2: Relations have arity ≤ 2 (and one has arity 2), i.e., graphs
- Unbounded-treewidth: for all $k \in \mathbb{N}$, there is $I_k \in \mathcal{I}$ of treewidth $\geq k$
- Constructible: given k, we can compute such an instance I_k in PTIME
- #P-hard under RP reductions: reduce in PTIME with high probability from the problem of counting accepting paths of a PTIME machine

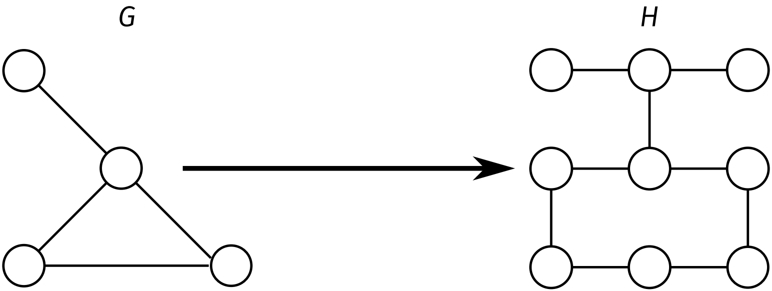
Table of contents

Idea: Topological minors

- G is a topological minor of H if:

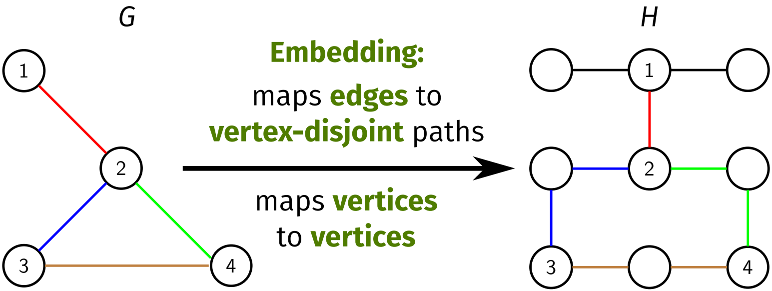
Idea: Topological minors

- G is a topological minor of H if:



Idea: Topological minors

- G is a topological minor of H if:



Topological minor extraction results

Theorem [Robertson and Seymour, 1986]
For any planar graph G of degree ≤ 3,
for any graph H of sufficiently high treewidth,
G is a topological minor of H.

Topological minor extraction results

Theorem [Robertson and Seymour, 1986]

For any planar graph G of degree ≤ 3,
for any graph H of sufficiently high treewidth,
G is a topological minor of H.

More recently:

Theorem [Chekuri and Chuzhoy, 2014]

There is a certain constant $c \in \mathbb{N}$ such that
for any planar graph G of degree ≤ 3 and graph H of treewidth $\geq |G|^c$,
we can embed G as a topological minor of H in PTIME with high proba

Proof sketch

- Choose a problem from which to reduce:
  - Must be #P-hard on planar degree-3 graphs
  - Must be encodable to an FO query q

$\rightarrow$ We use the problem of counting graph matchings

Proof sketch

- Choose a problem from which to reduce:
    - Must be #P-hard on planar degree-3 graphs
    - Must be encodable to an FO query q

  $\rightarrow$ We use the problem of counting graph matchings

- Given an input graph G, compute $k := |G|^c$

Proof sketch

- Choose a problem from which to reduce:
  - Must be #P-hard on planar degree-3 graphs
  - Must be encodable to an FO query q

  → We use the problem of counting graph matchings

- Given an input graph G, compute $k := |G|^c$

- Compute in PTIME an instance I_k of $\mathcal{I}$ of treewidth $\geq k$

Proof sketch

- Choose a problem from which to reduce:
  - Must be #P-hard on planar degree-3 graphs
  - Must be encodable to an FO query q

  $\rightarrow$ We use the problem of counting graph matchings
- Given an input graph G, compute $k := |G|^c$
- Compute in PTIME an instance I_k of $\mathcal{I}$ of treewidth $\geq k$
- Compute in randomized PTIME an embedding of G in I_k

Proof sketch

- Choose a problem from which to reduce:
  - Must be #P-hard on planar degree-3 graphs
  - Must be encodable to an FO query q
- $\rightarrow$ We use the problem of counting graph matchings
- Given an input graph G, compute $k := |G|^c$
- Compute in PTIME an instance I_k of $\mathcal{I}$ of treewidth $\geq k$
- Compute in randomized PTIME an embedding of G in I_k
- Construct a probability valuation π of I_k such that:
  - Unneccessary edges of I_k are removed
  - PQE for q gives the answer to the hard problem
- $\rightarrow$ Easy for MSO but trickier for FO

Table of contents

Non-probabilistic evaluation

Apply [Chekuri and Chuzhoy, 2014] to method of [Ganian et al., 2014] for MSO non-probabilistic query evaluation (QE):

Theorem

For any arity-2 signature σ and level Σ_i^P of the polynomial hierarchy, there is a MSO query q_i such that, for any constructible, subinstance-closed, unbounded-tw class $\mathcal{I}$, the QE problem for $\mathcal{Q} = \{q_i\}$ and $\mathcal{I}$ is Σ_i^P-hard under RP reductions

Non-probabilistic evaluation

Apply [Chekuri and Chuzhoy, 2014] to method of [Ganian et al., 2014] for MSO non-probabilistic query evaluation (QE):

Theorem

For any arity-2 signature σ and level Σ_i^P of the polynomial hierarchy, there is a MSO query q_i such that, for any constructible, subinstance-closed, unbounded-tw class $\mathcal{I}$, the QE problem for $\mathcal{Q} = \{q_i\}$ and $\mathcal{I}$ is Σ_i^P-hard under RP reductions

Variant: we also show a result like in [Ganian et al., 2014], with:

- weaker constructibility requirement on $\mathcal{I}$:
  - $\rightarrow$ $\mathcal{I}$ is densely unbounded poly-logarithmically
- stronger complexity hypothesis:
  - $\rightarrow$ PH does not admit $2^{o(n)}$-sized circuits

Lower bound on OBDD representations

OBDD for a query q on instance I:

ordered decision diagram on the facts of I to decide whether q holds

Lower bound on OBDD representations

OBDD for a query q on instance I:
ordered decision diagram on the facts of I to decide whether q holds

$$q : \exists x\, y\; R(x) \wedge S(x, y) \wedge T(y)$$

R	
a	r_1
b	r_2
c	r_3

S		
a	a	s_1
b	v	s_2
b	w	s_3

T	
v	t_1
w	t_2
b	t_3

Lower bound on OBDD representations

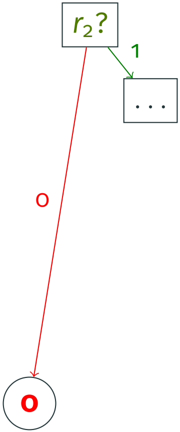
OBDD for a query q on instance I:
ordered decision diagram on the facts of I to decide whether q holds

$$q : \exists x\, y\; R(x) \wedge S(x, y) \wedge T(y)$$

R	
a	r_1
b	r_2
c	r_3

S		
a	a	s_1
b	v	s_2
b	w	s_3

T	
v	t_1
w	t_2
b	t_3



Lower bound on OBDD representations

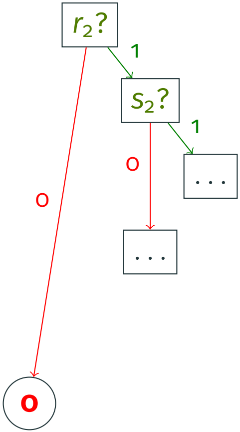
OBDD for a query q on instance I:
ordered decision diagram on the facts of I to decide whether q holds

$$q : \exists x\, y\; R(x) \wedge S(x, y) \wedge T(y)$$

R	
a	r_1
b	r_2
c	r_3

S		
a	a	s_1
b	v	s_2
b	w	s_3

T	
v	t_1
w	t_2
b	t_3



Lower bound on OBDD representations

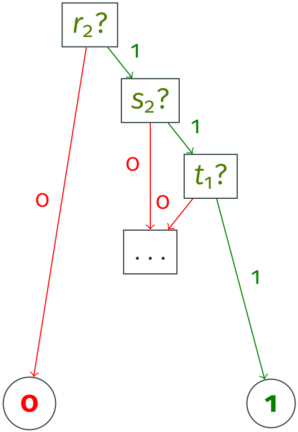
OBDD for a query q on instance I:
ordered decision diagram on the facts of I to decide whether q holds

$$q : \exists x\, y\; R(x) \wedge S(x, y) \wedge T(y)$$

R	
a	r_1
b	r_2
c	r_3

S		
a	a	s_1
b	v	s_2
b	w	s_3

T	
v	t_1
w	t_2
b	t_3



Lower bound on OBDD representations

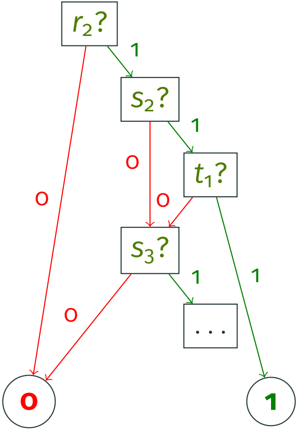
OBDD for a query q on instance I:
ordered decision diagram on the facts of I to decide whether q holds

$$q : \exists x\, y\; R(x) \wedge S(x, y) \wedge T(y)$$

R	
a	r_1
b	r_2
c	r_3

S		
a	a	s_1
b	v	s_2
b	w	s_3

T	
v	t_1
w	t_2
b	t_3



Lower bound on OBDD representations

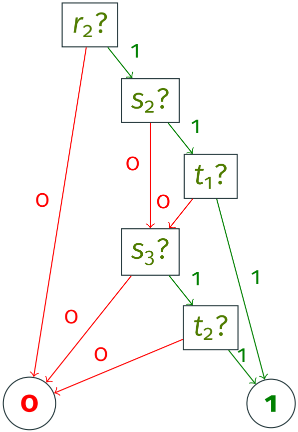
OBDD for a query q on instance I:
ordered decision diagram on the facts of I to decide whether q holds

$$q : \exists x\, y\; R(x) \land S(x, y) \land T(y)$$

R	
a	r_1
b	r_2
c	r_3

S		
a	a	s_1
b	v	s_2
b	w	s_3

T	
v	t_1
w	t_2
b	t_3



Lower bound on OBDD representations (result)

- If we find an OBDD of q on I in PTIME, then PQE of q on I also is

$\rightarrow$ Show inexistence of poly-size OBDDs (rather than PQE hardness)

Lower bound on OBDD representations (result)

- If we find an OBDD of q on I in PTIME, then PQE of q on I also is

$\rightarrow$ Show inexistence of poly-size OBDDs (rather than PQE hardness)

Theorem

For any arity-2 signature σ, there is a UCQ with inequalities q s.t. for any $\mathcal{I}$ of treewidth densely unbounded poly-logarithmically, q has no OBDDs of polynomial size on instances of $\mathcal{I}$

Lower bound on OBDD representations (result)

- If we find an OBDD of q on I in PTIME, then PQE of q on I also is

$\rightarrow$ Show inexistence of poly-size OBDDs (rather than PQE hardness)

Theorem

For any arity-2 signature σ, there is a UCQ with inequalities q s.t. for any $\mathcal{I}$ of treewidth densely unbounded poly-logarithmically, q has no OBDDs of polynomial size on instances of $\mathcal{I}$

$\rightarrow$ Meta-dichotomy on the connected UCQ$^{\neq}$ for which this holds

Also in paper

- Hardness results for MSO match counting:
  - $\rightarrow$ Count how many subsets X are such that I satisfies $q(X)$
- Connect the tractability result of MSO on treelike TID to the study of tractable lineages:
  - $\rightarrow$ d-DNNFs, OBDDs, formulae, etc.
- Connect the same result to tractability of safe queries:
  - Inversion-free queries: subclass of safe queries, tractable OBDDs
  - $\rightarrow$ We can always rewrite their input instances to treelike instances in a lineage-preserving way (hence, probability-preserving)

Table of contents

Summary of our dichotomy

Upper. PQE for MSO on treelike instances
has linear data complexity up to arithmetic costs

Summary of our dichotomy

Upper. PQE for MSO on treelike instances
has linear data complexity up to arithmetic costs

Lower. PQE for FO on any constructible, arity-2, unbounded-tw
instance family is #P-hard under RP reductions

Summary of our dichotomy

Upper. PQE for MSO on treelike instances
has linear data complexity up to arithmetic costs

Lower. PQE for FO on any constructible, arity-2, unbounded-tw
instance family is #P-hard under RP reductions

$\rightarrow$ Bounded treewidth is the right notion for tractability of PQE

Future work

- Can we show #P-hardness under usual P reductions?
  - $\rightarrow$ Depends on [Chekuri and Chuzhoy, 2014]

Future work

- Can we show #P-hardness under usual P reductions?
  - $\rightarrow$ Depends on [Chekuri and Chuzhoy, 2014]

- Can we extend the result to arbitrary arity signatures?

Future work

- Can we show #P-hardness under usual P reductions?
  - $\rightarrow$ Depends on [Chekuri and Chuzhoy, 2014]
- Can we extend the result to arbitrary arity signatures?
- Can we extend the result to weaker query languages like $\mathrm{UCQ}^{\neq}$?

Conjecture

PQE is hard on any constructible unbounded-tw family for:

$$q : (E(x,y) \vee E(y,x)) \wedge (E(y,z) \vee E(z,y)) \wedge x \neq z$$

$\rightarrow$ This query is alerady hard in terms of OBDDs

Future work

- Can we show #P-hardness under usual P reductions?
  → Depends on [Chekuri and Chuzhoy, 2014]

- Can we extend the result to arbitrary arity signatures?

- Can we extend the result to weaker query languages like $\mathrm{UCQ}^{\neq}$?

  Conjecture

  PQE is hard on any constructible unbounded-tw family for:

  $$q : (E(x,y) \lor E(y,x)) \land (E(y,z) \lor E(z,y)) \land x \neq z$$

  → This query is alerady hard in terms of OBDDs

Thanks for your attention!

References I

Amarilli, A., Bourhis, P., and Senellart, P. (2015).
Provenance circuits for trees and treelike instances.
In *Proc. ICALP*, LNCS, pages 56–68.

Chaudhuri, S. and Vardi, M. Y. (1992).
On the equivalence of recursive and nonrecursive Datalog programs.
In *PODS*.

Chekuri, C. and Chuzhoy, J. (2014).
Polynomial bounds for the grid-minor theorem.
In *STOC*.

References II

Dalvi, N. and Suciu, D. (2012).
The dichotomy of probabilistic inference for unions of conjunctive queries.
JACM, 59(6):30.

Ganian, R., Hliněný, P., Langer, A., Obdržálek, J., Rossmanith, P., and Sikdar, S. (2014).
Lower bounds on the complexity of MSO1 model-checking.
JCSS, 1(80).

Robertson, N. and Seymour, P. D. (1986).
Graph minors. V. Excluding a planar graph.
J. Comb. Theory, Ser. B, 41(1).

Encoding treelike instances [Chaudhuri and Vardi, 1992]

Instance:

N	
a	*b*
b	*c*
c	*d*
d	*e*
e	*f*

S	
a	*c*
b	*e*

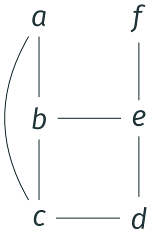
Encoding treelike instances [Chaudhuri and Vardi, 1992]

Instance:

N	
a	*b*
b	*c*
c	*d*
d	*e*
e	*f*

S	
a	*c*
b	*e*

Gaifman graph:



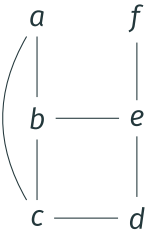
Encoding treelike instances [Chaudhuri and Vardi, 1992]

Instance:

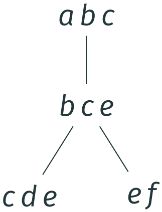
N	
a	*b*
b	*c*
c	*d*
d	*e*
e	*f*

S	
a	*c*
b	*e*

Gaifman graph:



Tree decomp.:



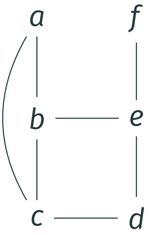
Encoding treelike instances [Chaudhuri and Vardi, 1992]

Instance:

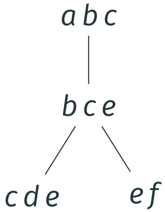
N	
a	b
b	c
c	d
d	e
e	f

S	
a	c
b	e

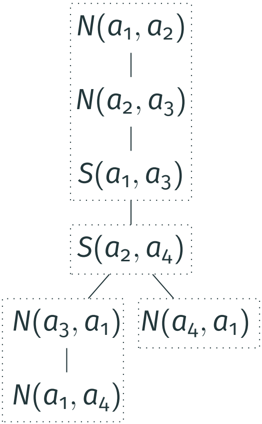
Gaifman graph:



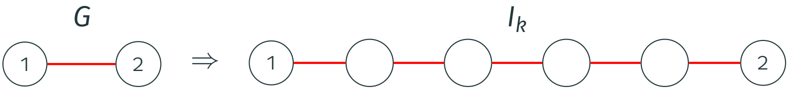
Tree decomp.:



Tree encoding:

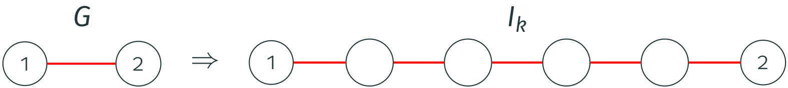


Defining the query



- In the embedding, edges of G can become long paths in I_k
- q must answer the hard problem on G despite subdivisions
  - → Easy in MSO but tricky in FO!

Defining the query



- In the embedding, edges of G can become long paths in I_k
- q must answer the hard problem on G despite subdivisions
  - → Easy in MSO but tricky in FO!

→ Our q restricts to a subset of the worlds of known weight and gives the right answer up to renormalization