Probabilistic Databases under Updates

Boolean Query Evaluation
and Ranked Enumeration

Christoph Berkholz, Maximilian Merz

Talk at AlMoTh 2022, based on the PODS 2021 paper

1 Probabilistic Databases under Updates

Boolean Query Evaluation
and Ranked Enumeration

Christoph Berkholz, Maximilian Merz

Talk at AlMoTh 2022, based on the PODS 2021 paper

1 Probabilistic Databases 2 under Updates

Boolean Query Evaluation
and Ranked Enumeration

Christoph Berkholz, Maximilian Merz

Talk at AlMoTh 2022, based on the PODS 2021 paper

1 Probabilistic Databases 2 under Updates

3 Boolean Query Evaluation and Ranked Enumeration

Christoph Berkholz, Maximilian Merz

Talk at AlMoTh 2022, based on the PODS 2021 paper

1 Probabilistic Databases 2 under Updates

3 Boolean Query Evaluation 4 and Ranked Enumeration

Christoph Berkholz, Maximilian Merz

Talk at AlMoTh 2022, based on the PODS 2021 paper

Probabilistic Databases under Updates

Boolean Query Evaluation and Ranked Enumeration

Christoph Berkholz, Maximilian Merz

Talk at AlMoTh 2022, based on the PODS 2021 paper

R	A
	a
	b

S	A	B
	a	b
	a	d

?

R	A
	a
	b

S	A	B
	a	b
	a	d

?

R	A
	a
	b

S	A	B
	a	b
	a	d

?

R	A
	a
	b

S	A	B
	a	b
	a	d

R	A
	a

S	A	B
	a	b
	a	d

Probabilistic Databases (PDBs)

?

R	A
	a
	b

S	A	B
	a	b
	a	d

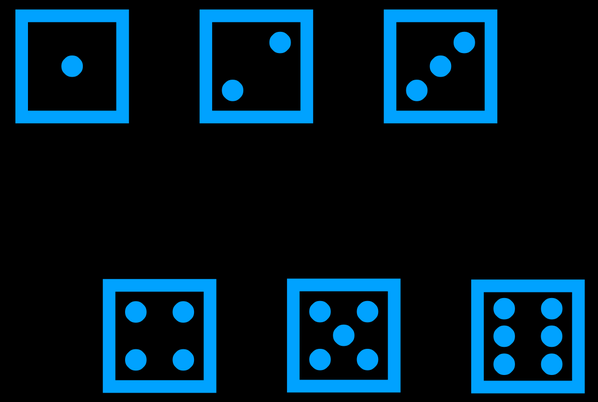
R	A
	a

S	A	B
	a	b
	a	d

Probabilistic Databases (PDBs)

Probabilistic Databases (PDBs)

Dice roll



Probability Space

(“possible worlds”)

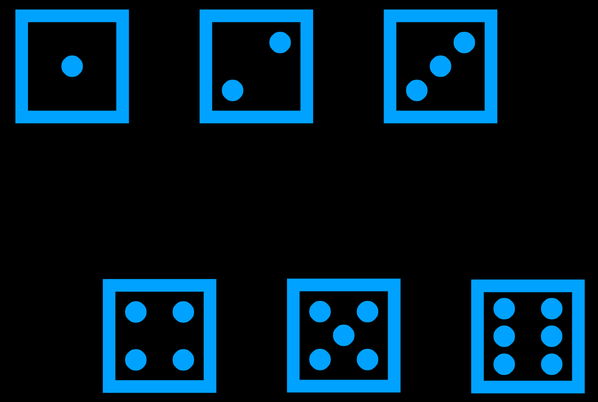
Probabilistic Databases (PDBs)

Dice roll

PDB

Probability Space

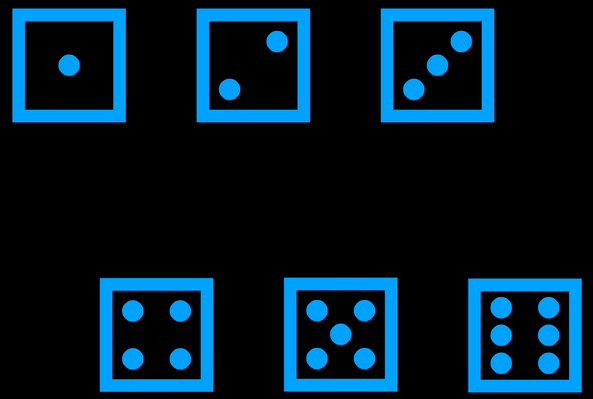
(“possible worlds”)



Probabilistic Databases (PDBs)

Dice roll

PDB



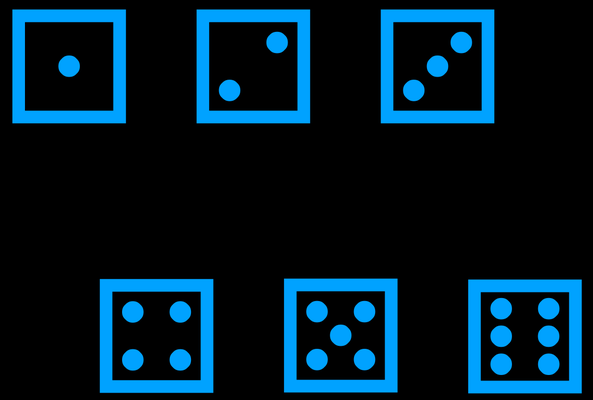
R	A
	a
	b

Probability Space

(“possible worlds”)

Probabilistic Databases (PDBs)

Dice roll



PDB

R	A
	a
	b

R	A
	a

Probability Space

(“possible worlds”)

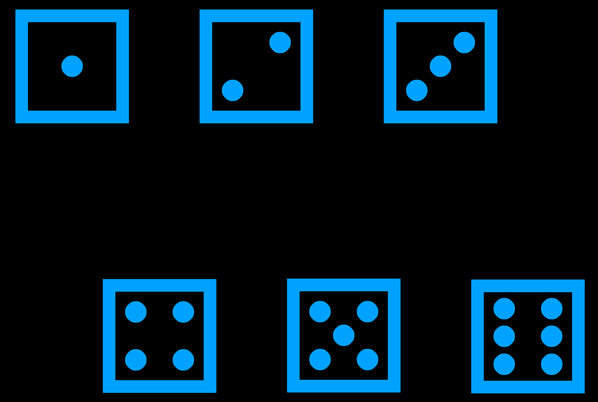
Probabilistic Databases (PDBs)

Dice roll

PDB

Probability Space

(“possible worlds”)



R	A
	a
	b

R	A
	a

S	A	B
	a	b
	a	d

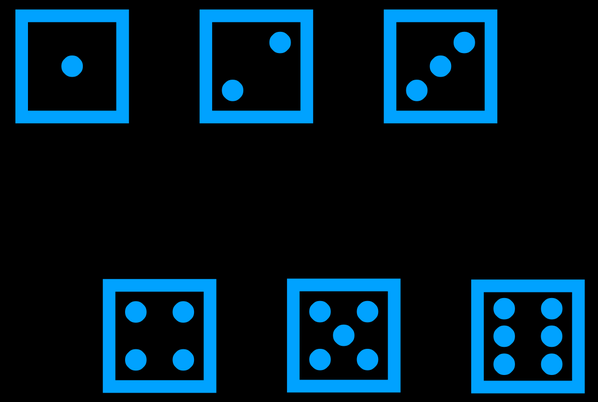
Probabilistic Databases (PDBs)

Dice roll

PDB

Probability Space

(“possible worlds”)



R	A
	a
	b

R	A
	a

S	A	B
	a	b
	a	d

R	A
	a

S	A	B
	a	b
	a	d

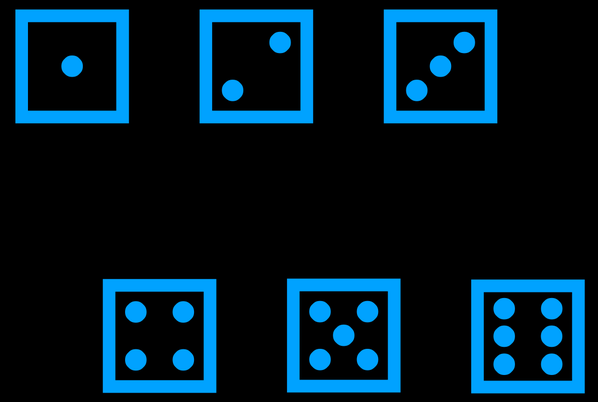
Probabilistic Databases (PDBs)

Dice roll

PDB

Probability Space

(“possible worlds”)



R	A
	a
	b

R	A
	a

. . .

S	A	B
	a	b
	a	d

R	A
	a

S	A	B
	a	b
	a	d

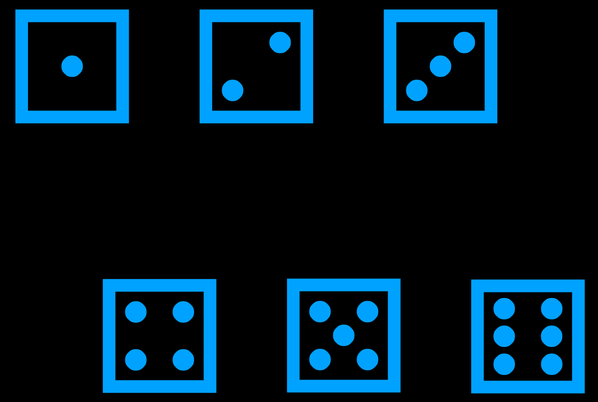
Probabilistic Databases (PDBs)

Dice roll

PDB

Probability Space

(“possible worlds”)



R	A
	a
	b

R	A
	a

. . .

S	A	B
	a	b
	a	d

R	A
	a

S	A	B
	a	b
	a	d

storing the set of possible worlds explicitly
is not practical!

?

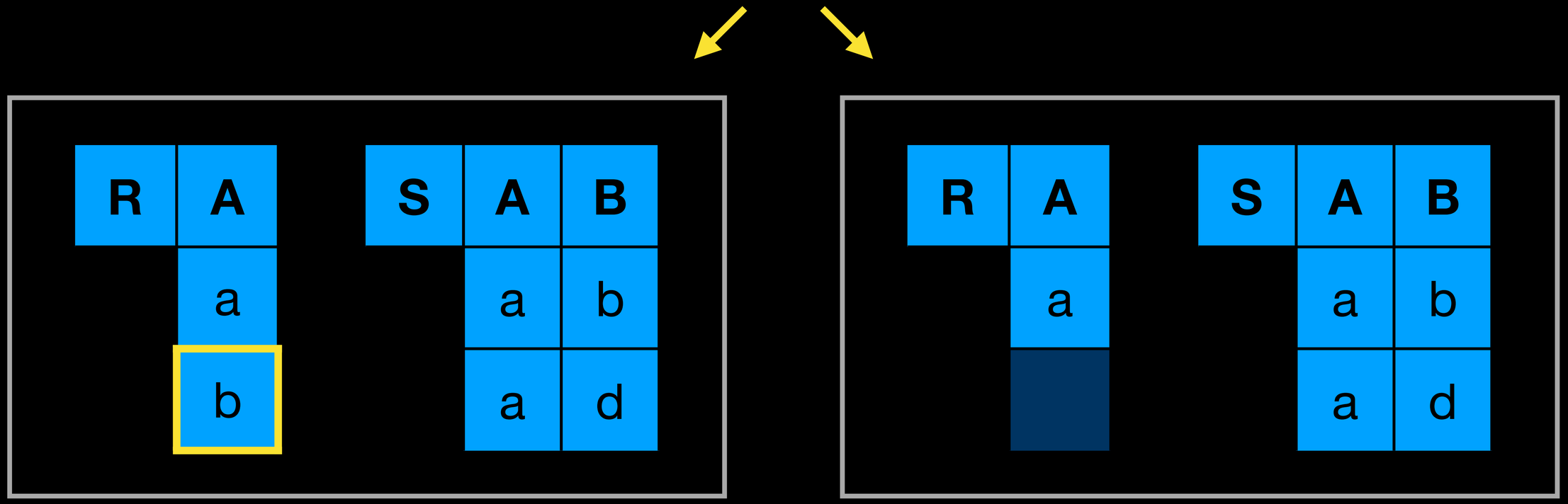
R	A
	a
	b

S	A	B
	a	b
	a	d

R	A
	a

S	A	B
	a	b
	a	d

Tuple-Independent PDB



Tuple-Independent PDB

Assumption:
Every tuple is an independent probabilistic event

R	A
	a
	b

S	A	B
	a	b
	a	d

R	A
	a

S	A	B
	a	b
	a	d

Tuple-Independent PDB

Assumption:
Every tuple is an independent probabilistic event

R	A	p
	a	1
	b	⅓

S	A	B	p
	a	b	1
	a	d	1

R	A
	a
	b

S	A	B
	a	b
	a	d

R	A
	a

S	A	B
	a	b
	a	d

Tuple-Independent PDB

Assumption:
Every tuple is an independent probabilistic event

R	A	p
	a	1
	b	⅓

S	A	B	p
	a	b	1
	a	d	1

$$P(w_1) = \frac{1}{3}$$

R	A
	a
	b

S	A	B
	a	b
	a	d

R	A
	a

S	A	B
	a	b
	a	d

Tuple-Independent PDB

Assumption:
Every tuple is an independent probabilistic event

R	A	p
	a	1
	b	⅓

S	A	B	p
	a	b	1
	a	d	1

$P(w_1) = \frac{1}{3}$

$P(w_2) = \frac{2}{3}$

R	A
	a
	b

S	A	B
	a	b
	a	d

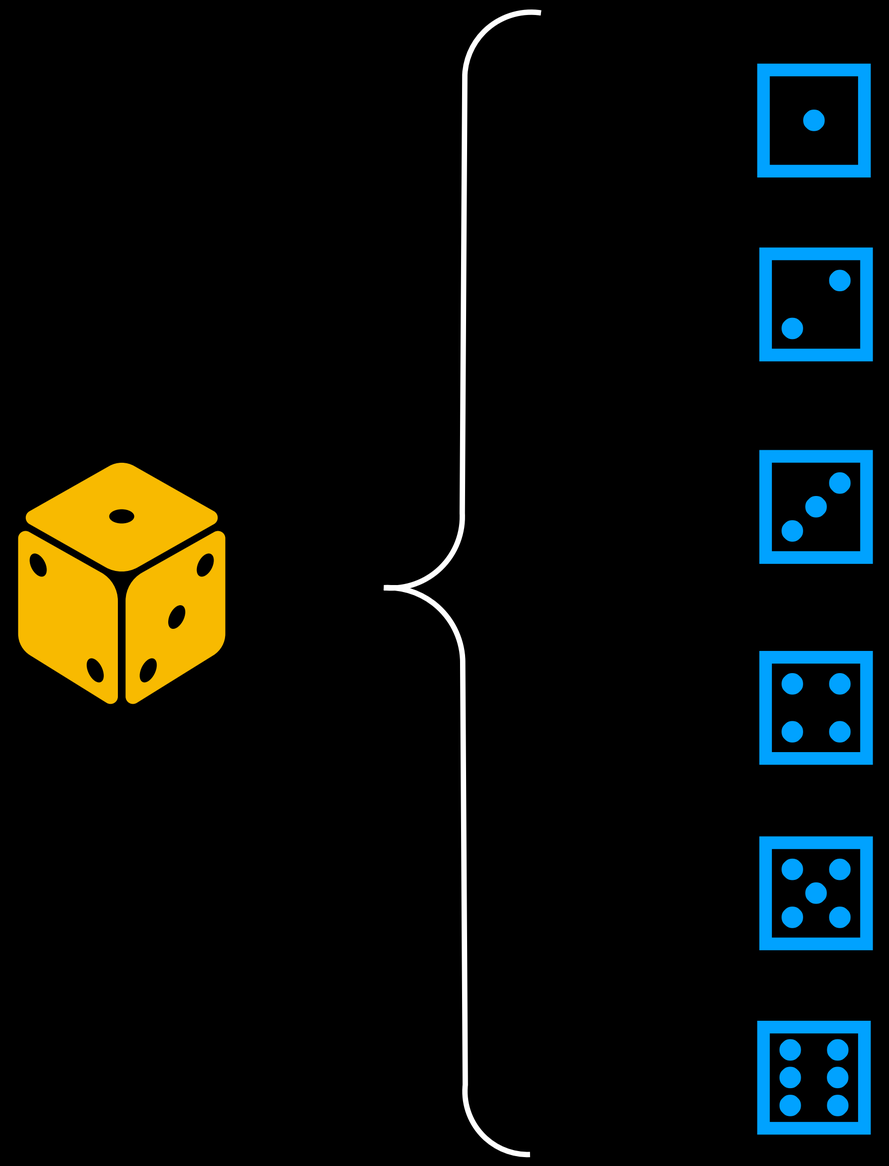
R	A
	a

S	A	B
	a	b
	a	d

Dice roll

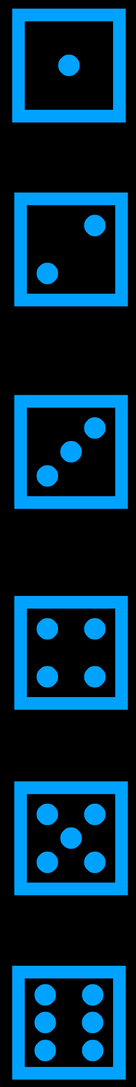
Dice roll

possible worlds



Dice roll

possible worlds



P = ⅙

P = ⅙

P = ⅙

P = ⅙

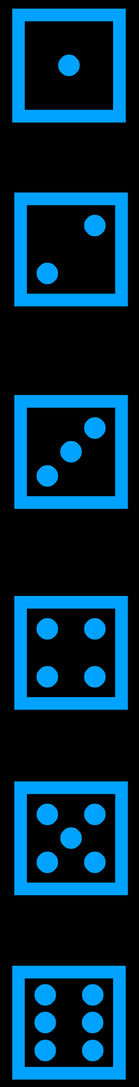
P = ⅙

P = ⅙

Dice roll

possible worlds

query evaluation
“Does the dice show a prime?”



P = ⅙

P = ⅙

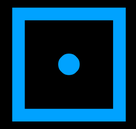
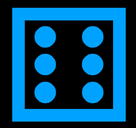
P = ⅙

P = ⅙

P = ⅙

P = ⅙

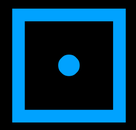
Dice roll

possible worlds		query evaluation "Does the dice show a prime?"
	P = 1/6	✗
	P = 1/6	✓
	P = 1/6	✓
	P = 1/6	✗
	P = 1/6	✓
	P = 1/6	✗

Dice roll

possible worlds

query evaluation
“Does the dice show a prime?”



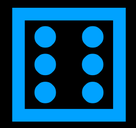
P = ⅙ ✗

P = ⅙ ✔

P = ⅙ ✔

P = ⅙ ✗

P = ⅙ ✔



P = ⅙ ✗

P(prime)
= ⅙+⅙+⅙
= ½

tuple-independent
PDB

R	A	p
	a	½
	b	⅓

tuple-independent
PDB

possible worlds

R	A	p
	a	½
	b	⅓

tuple-independent
PDB

R	A	p
	a	½
	b	⅓

possible worlds

R	A
	a
	b

$P(w_1) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$

tuple-independent
PDB

R	A	p
	a	½
	b	⅓

possible worlds

R	A
	a
	b

$P(w_1) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$

R	A
	a

$P(w_2) = \frac{1}{2} \cdot \underline{\frac{2}{3}} = \frac{1}{3}$

tuple-independent
PDB

R	A	p
	a	½
	b	⅓

possible worlds

R	A
	a
	b

$P(w_1) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$

R	A
	a

$P(w_2) = \frac{1}{2} \cdot \underline{\frac{2}{3}} = \frac{1}{3}$

R	A
	b

$P(w_3) = \underline{\frac{1}{2}} \cdot \frac{1}{3} = \frac{1}{6}$

tuple-independent
PDB

R	A	p
	a	½
	b	⅓

possible worlds

R	A
	a
	b

$P(w_1) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$

R	A
	a

$P(w_2) = \frac{1}{2} \cdot \underline{\frac{2}{3}} = \frac{1}{3}$

R	A
	b

$P(w_3) = \underline{\frac{1}{2}} \cdot \frac{1}{3} = \frac{1}{6}$

R	A

$P(w_4) = \underline{\frac{1}{2}} \cdot \underline{\frac{2}{3}} = \frac{1}{3}$

tuple-independent
PDB

R	A	p
	a	½
	b	⅓

possible worlds

R	A
	a
	b

$P(w_1) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$

R	A
	a

$P(w_2) = \frac{1}{2} \cdot \underline{\frac{2}{3}} = \frac{1}{3}$

R	A
	b

$P(w_3) = \underline{\frac{1}{2}} \cdot \frac{1}{3} = \frac{1}{6}$

R	A

$P(w_4) = \underline{\frac{1}{2}} \cdot \underline{\frac{2}{3}} = \frac{1}{3}$

query evaluation

$$\varphi = \exists x\, Rx$$

tuple-independent
PDB

R	A	p
	a	½
	b	⅓

possible worlds

R	A
	a
	b

$P(w_1) = ½ \cdot ⅓ = ⅙$

R	A
	a

$P(w_2) = ½ \cdot \underline{⅔} = ⅓$

R	A
	b

$P(w_3) = \underline{½} \cdot ⅓ = ⅙$

R	A

$P(w_4) = \underline{½} \cdot \underline{⅔} = ⅓$

query evaluation

$\varphi = \exists x\, Rx$

$\varphi(w_1) = ✔$

tuple-independent
PDB

R	A	p
	a	½
	b	⅓

possible worlds

R	A
	a
	b

$P(w_1) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$

R	A
	a

$P(w_2) = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$

R	A
	b

$P(w_3) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$

R	A

$P(w_4) = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$

query evaluation

$\varphi = \exists x\, Rx$

$\varphi(w_1) = \checkmark$

$\varphi(w_2) = \checkmark$

tuple-independent
PDB

R	A	p
	a	½
	b	⅓

possible worlds

query evaluation

$\varphi = \exists x\, Rx$

R	A
	a
	b

$P(w_1) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$

$\varphi(w_1) = \checkmark$

R	A
	a

$P(w_2) = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$

$\varphi(w_2) = \checkmark$

R	A
	b

$P(w_3) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$

$\varphi(w_3) = \checkmark$

R	A

$P(w_4) = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$

tuple-independent
PDB

R	A	p
	a	½
	b	⅓

possible worlds

query evaluation

$\varphi = \exists x\, Rx$

R	A
	a
	b

$P(w_1) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$

$\varphi(w_1) =$ ✔

R	A
	a

$P(w_2) = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$

$\varphi(w_2) =$ ✔

R	A
	b

$P(w_3) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$

$\varphi(w_3) =$ ✔

R	A

$P(w_4) = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$

$\varphi(w_4) =$ ✘

tuple-independent
PDB

R	A	p
	a	½
	b	⅓

possible worlds

R	A
	a
	b

$P(w_1) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$

R	A
	a

$P(w_2) = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$

R	A
	b

$P(w_3) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$

R	A

$P(w_4) = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$

query evaluation

$\varphi = \exists x\, Rx$

$\varphi(w_1) = \checkmark$

$\varphi(w_2) = \checkmark$

$\varphi(w_3) = \checkmark$

$\varphi(w_4) = \times$

$P(\varphi)$

tuple-independent
PDB

R	A	p
	a	½
	b	⅓

possible worlds

query evaluation

$\varphi = \exists x\, Rx$

R	A
	a
	b

$P(w_1) = ½ \cdot ⅓ = ⅙$ $\quad \varphi(w_1) = ✔$

R	A
	a

$P(w_2) = ½ \cdot ⅔ = ⅓$ $\quad \varphi(w_2) = ✔$

R	A
	b

$P(w_3) = ½ \cdot ⅓ = ⅙$ $\quad \varphi(w_3) = ✔$

R	A

$P(w_4) = ½ \cdot ⅔ = ⅓$ $\quad \varphi(w_4) = ✘$

$P(\varphi)$

$= ⅙ + ⅓ + ⅙$

$= ⅔$

1 Probabilistic Databases 2 under Updates

3 Boolean Query Evaluation 4 and Ranked Enumeration

Christoph Berkholz, Maximilian Merz

Talk at AlMoTh 2022, based on the PODS 2021 paper

✔ Probabilistic Databases
2 under Updates

3 Boolean Query Evaluation
4 and Ranked Enumeration

Christoph Berkholz, Maximilian Merz

Talk at AlMoTh 2022, based on the PODS 2021 paper

Probabilistic Databases under Updates

3 Boolean Query Evaluation
4 and Ranked Enumeration

Christoph Berkholz, Maximilian Merz

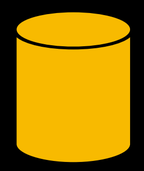
Talk at AlMoTh 2022, based on the PODS 2021 paper

Updates to the Database

Updates to the Database

init

- fix db schema
- fix finite domain
- fix FO query

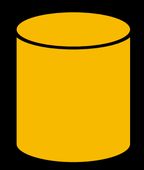


time

Updates to the Database

init

- fix db schema
- fix finite domain
- fix FO query



time

Updates to the Database

init

- fix db schema
- fix finite domain
- fix FO query

R	A	p
	a	¾
	b	⅓

S	A	B	p
	a	b	⅛

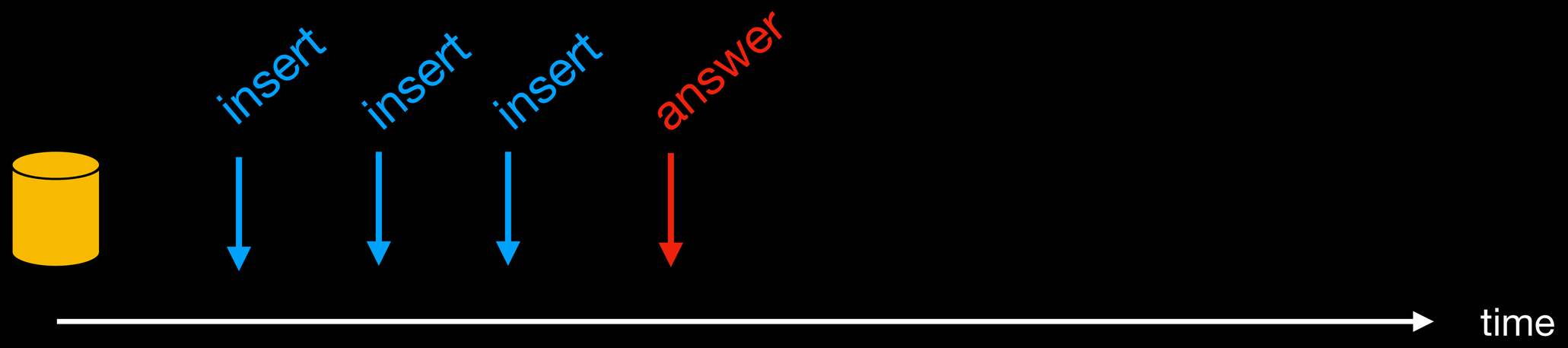
Updates to the Database

init

- fix db schema
- fix finite domain
- fix FO query

R	A	p
	a	¾
	b	⅓

S	A	B	p
	a	b	⅛



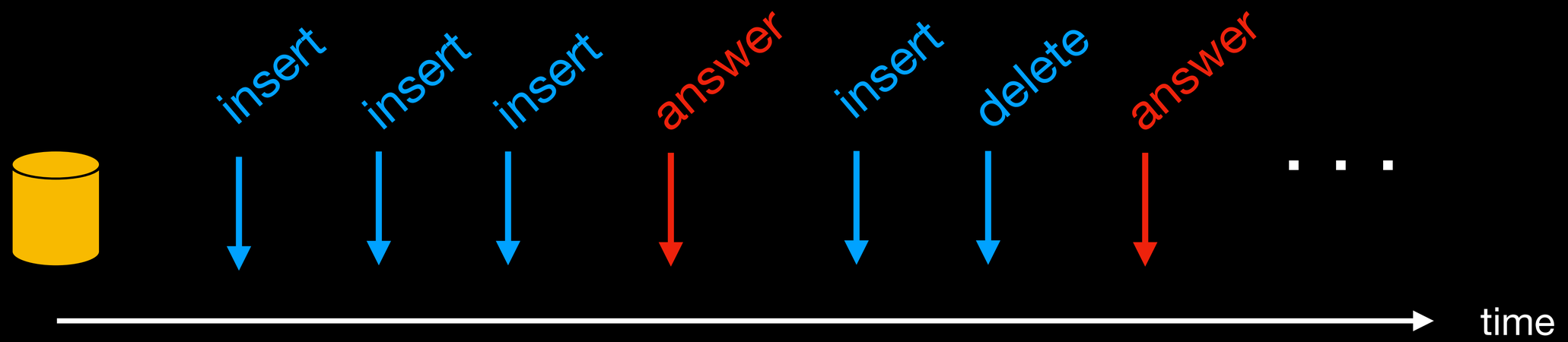
Updates to the Database

init

- fix db schema
- fix finite domain
- fix FO query

R	A	p
b	⅓	

S	A	B	p
	a	b	⅛
	a	d	1



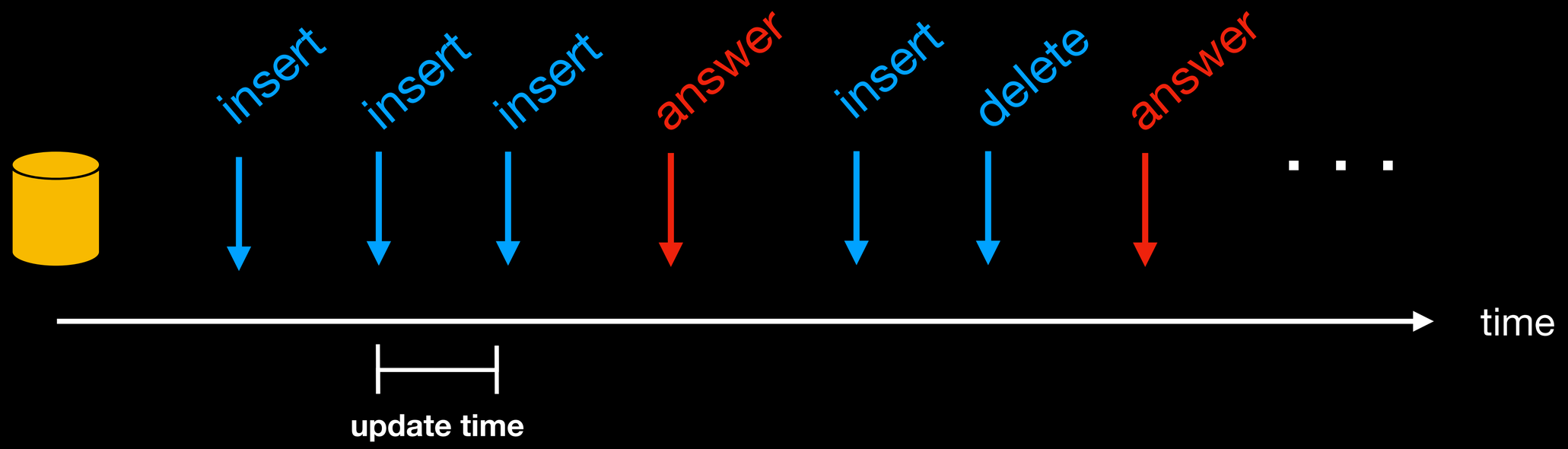
Updates to the Database

init

- fix db schema
- fix finite domain
- fix FO query

R	A	p
b	⅓	

S	A	B	p
	a	b	⅛
	a	d	1



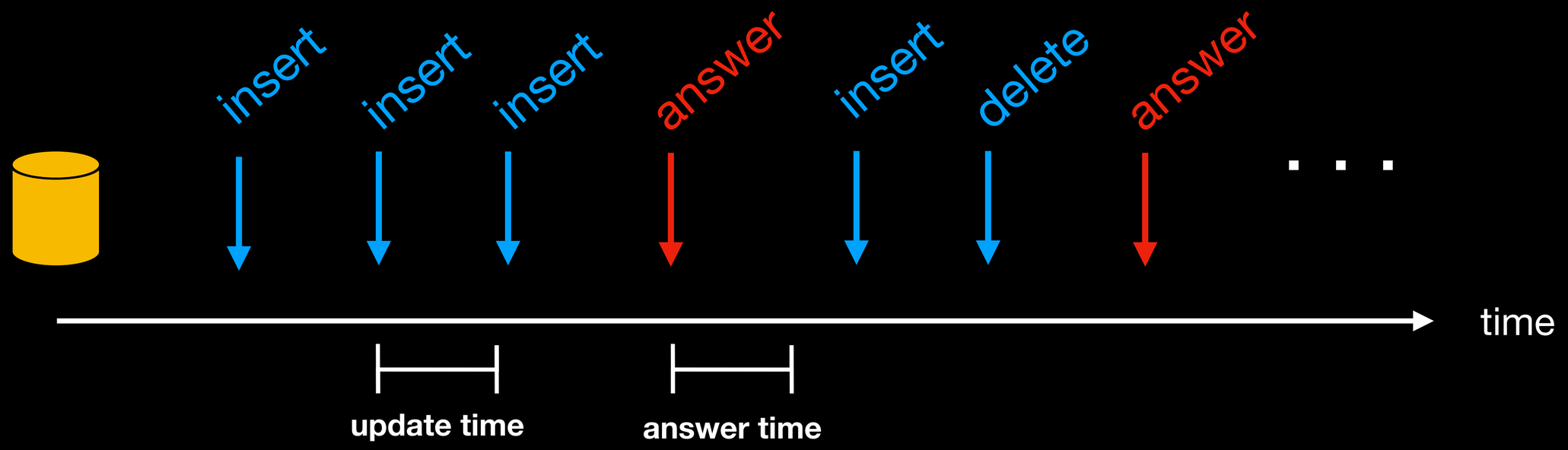
Updates to the Database

init

- fix db schema
- fix finite domain
- fix FO query

R	A	p
b	⅓	

S	A	B	p
	a	b	⅛
	a	d	1



Probabilistic Databases under Updates

3 Boolean Query Evaluation
4 and Ranked Enumeration

Christoph Berkholz, Maximilian Merz

Talk at AlMoTh 2022, based on the PODS 2021 paper

Probabilistic Databases under Updates

3 Boolean Query Evaluation
4 and Ranked Enumeration

Christoph Berkholz, Maximilian Merz

Talk at AlMoTh 2022, based on the PODS 2021 paper

Probabilistic Databases under Updates

Boolean Query Evaluation
and Ranked Enumeration

Christoph Berkholz, Maximilian Merz

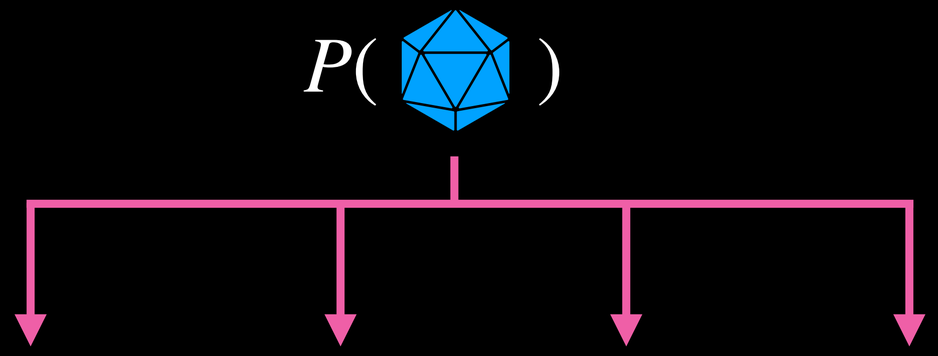
Talk at AlMoTh 2022, based on the PODS 2021 paper

Dalvi & Suciu Algorithm

$P($  $)$

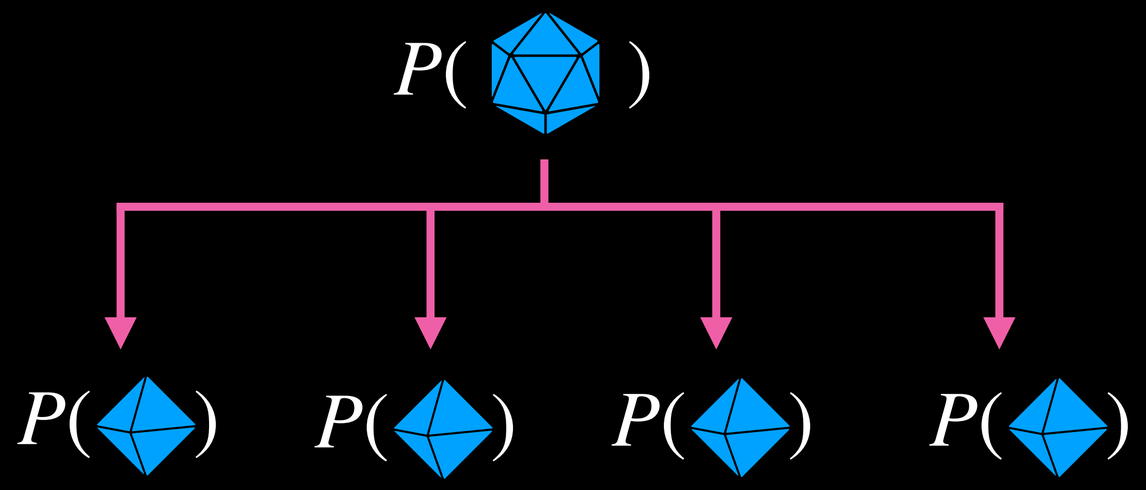
 = a large query expression

Dalvi & Suciu Algorithm



= a large query expression

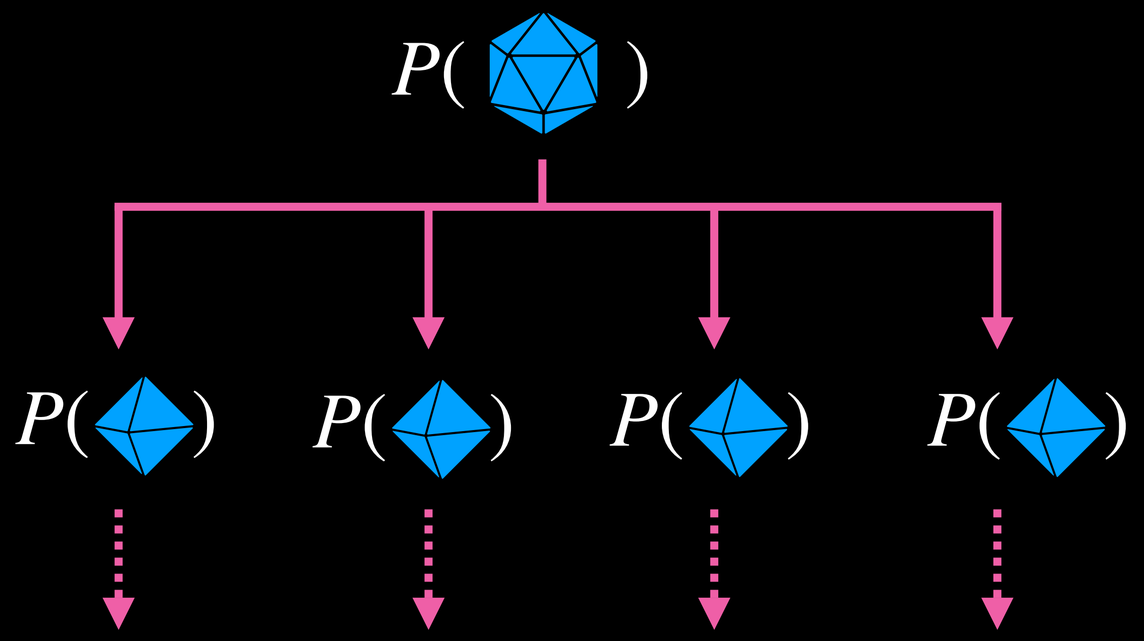
Dalvi & Suciu Algorithm




= a large query expression


= smaller queries

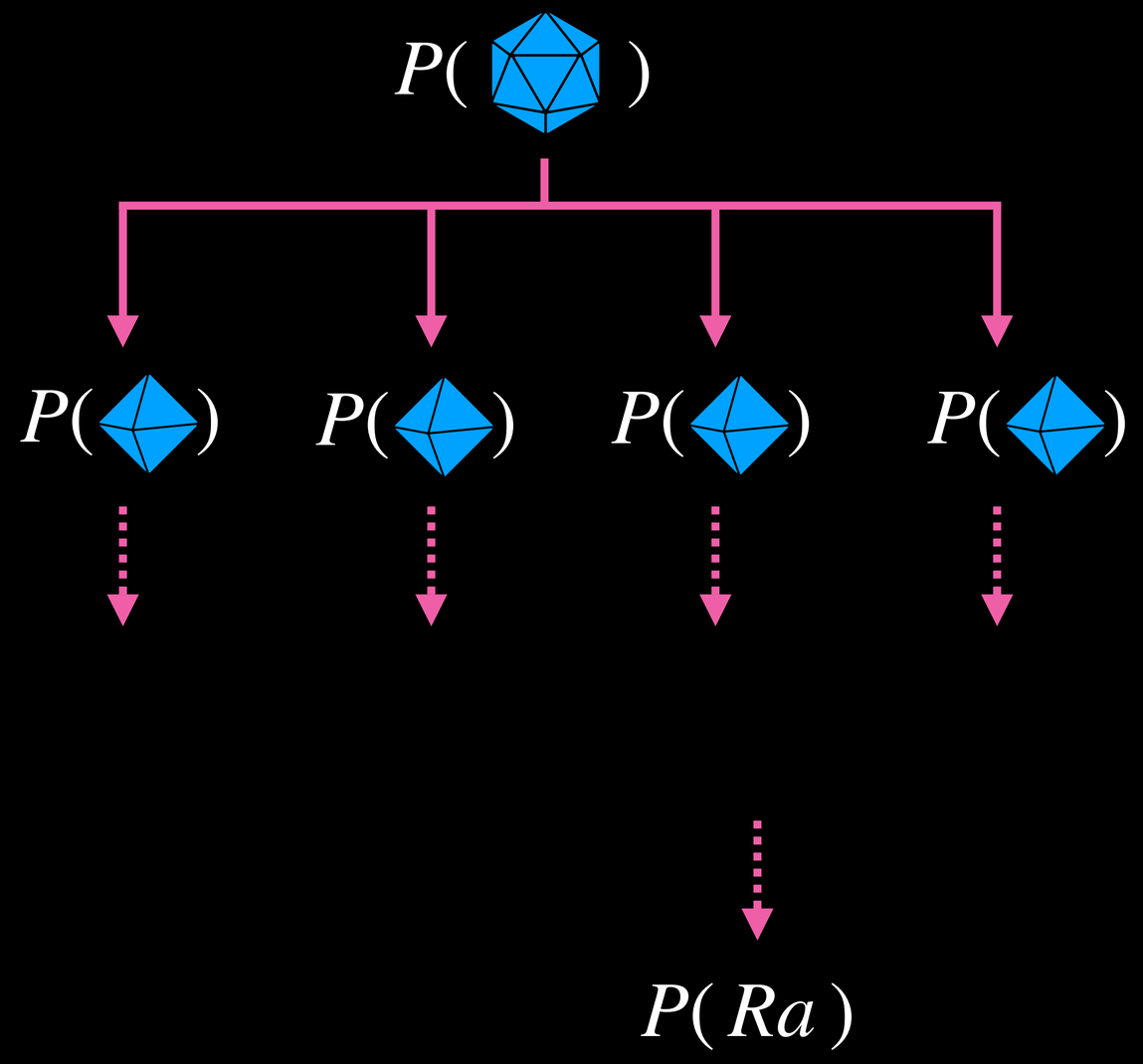
Dalvi & Suciu Algorithm




= a large query expression


= smaller queries

Dalvi & Suciu Algorithm




= a large query expression


= smaller queries

← atomic queries

Dalvi & Suciu Algorithm

Dalvi & Suciu Algorithm

Syntactical Independence 1

$$P(\ \exists x\, Rx\ \wedge\ \exists y\, Sy\ )$$

Dalvi & Suciu Algorithm

Syntactical Independence 1

$$P(\exists x\, Rx \wedge \exists y\, Sy)$$

Dalvi & Suciu Algorithm

Syntactical Independence 1

$$P(\exists x\, Rx \wedge \exists y\, Sy)$$

$$= P(\exists x\, Rx) \cdot P(\exists y\, Sy)$$

Dalvi & Suciu Algorithm

Syntactical Independence 1

$$P(\exists x\, Rx \wedge \exists y\, Sy)$$

$$= P(\exists x\, Rx) \cdot P(\exists y\, Sy)$$

Syntactical Independence 2

Dalvi & Suciu Algorithm

Syntactical Independence 1

$$P(\exists x\, Rx \wedge \exists y\, Sy)$$

$$= P(\exists x\, Rx) \cdot P(\exists y\, Sy)$$

Syntactical Independence 2

$$P(\exists x\, Rx)$$

Dalvi & Suciu Algorithm

Syntactical Independence 1

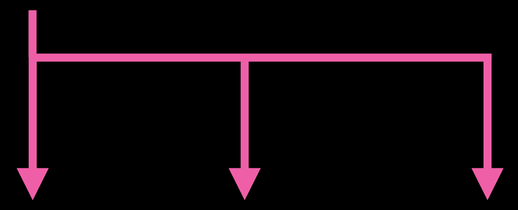
$$P(\exists x\, Rx \wedge \exists y\, Sy)$$

$$= P(\exists x\, Rx) \cdot P(\exists y\, Sy)$$

Syntactical Independence 2

$$P(\exists x\, Rx)$$

$x = ?$



Dalvi & Suciu Algorithm

Syntactical Independence 1

$$P(\exists x\, Rx \wedge \exists y\, Sy)$$

$$= P(\exists x\, Rx) \cdot P(\exists y\, Sy)$$

Syntactical Independence 2

$$P(\exists x\, Rx)$$

$x = ?$

a

$$P(Ra)$$

Dalvi & Suciu Algorithm

Syntactical Independence 1

$$P(\exists x\, Rx \wedge \exists y\, Sy)$$

$$= P(\exists x\, Rx) \cdot P(\exists y\, Sy)$$

Syntactical Independence 2

$$P(\exists x\, Rx)$$

$x = ?$

a b c

$$P(Ra) \quad P(Rb) \quad P(Rc)$$

Dalvi & Suciu Algorithm

Syntactical Independence 1

$$P(\exists x\, Rx \wedge \exists y\, Sy)$$

$$= P(\exists x\, Rx) \cdot P(\exists y\, Sy)$$

Syntactical Independence 2

$$P(\exists x\, Rx)$$

...if x is a separator variable!

$x = ?$

a b c

$$P(Ra) \quad P(Rb) \quad P(Rc)$$

Dalvi & Suciu Algorithm

Syntactical Independence 1

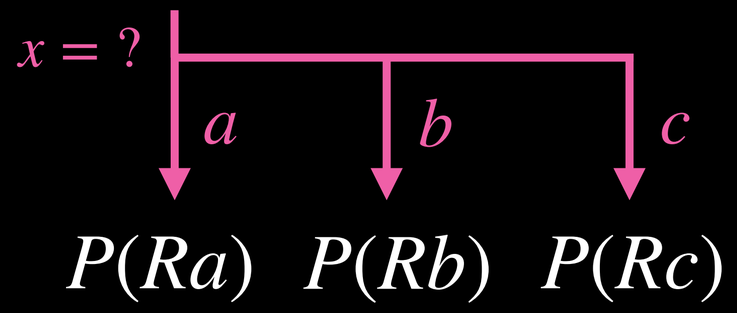
$$P(\exists x\, Rx \wedge \exists y\, Sy)$$

$$= P(\exists x\, Rx) \cdot P(\exists y\, Sy)$$

Syntactical Independence 2

$$P(\exists x\, Rx)$$

...if x is a separator variable!



Inclusion/Exclusion

Dalvi & Suciu Algorithm

Syntactical Independence 1

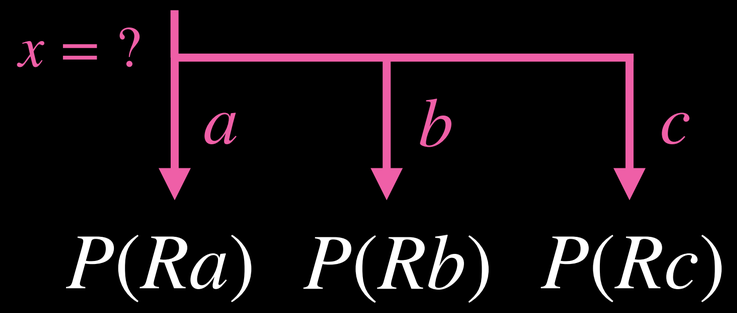
$$P(\exists x\, Rx \wedge \exists y\, Sy)$$

$$= P(\exists x\, Rx) \cdot P(\exists y\, Sy)$$

Syntactical Independence 2

$$P(\exists x\, Rx)$$

...if x is a separator variable!



Inclusion/Exclusion

$$P(\varphi_1 \wedge \varphi_2)$$

Dalvi & Suciu Algorithm

Syntactical Independence 1

$$P(\exists x\, Rx \wedge \exists y\, Sy)$$

$$= P(\exists x\, Rx) \cdot P(\exists y\, Sy)$$

Syntactical Independence 2

$$P(\exists x\, Rx)$$

...if x is a separator variable!

$x = ?$

a b c

$$P(Ra) \quad P(Rb) \quad P(Rc)$$

Inclusion/Exclusion

$$P(\varphi_1 \wedge \varphi_2)$$

$$P(\varphi_1) \qquad P(\varphi_2) \qquad P(\varphi_1 \vee \varphi_2)$$

Dalvi & Suciu Algorithm

Syntactical Independence 1

$$P(\exists x\, Rx \wedge \exists y\, Sy)$$

$$= P(\exists x\, Rx) \cdot P(\exists y\, Sy)$$

Syntactical Independence 2

$$P(\exists x\, Rx)$$

...if x is a separator variable!

$x = ?$

a b c

$$P(Ra) \quad P(Rb) \quad P(Rc)$$

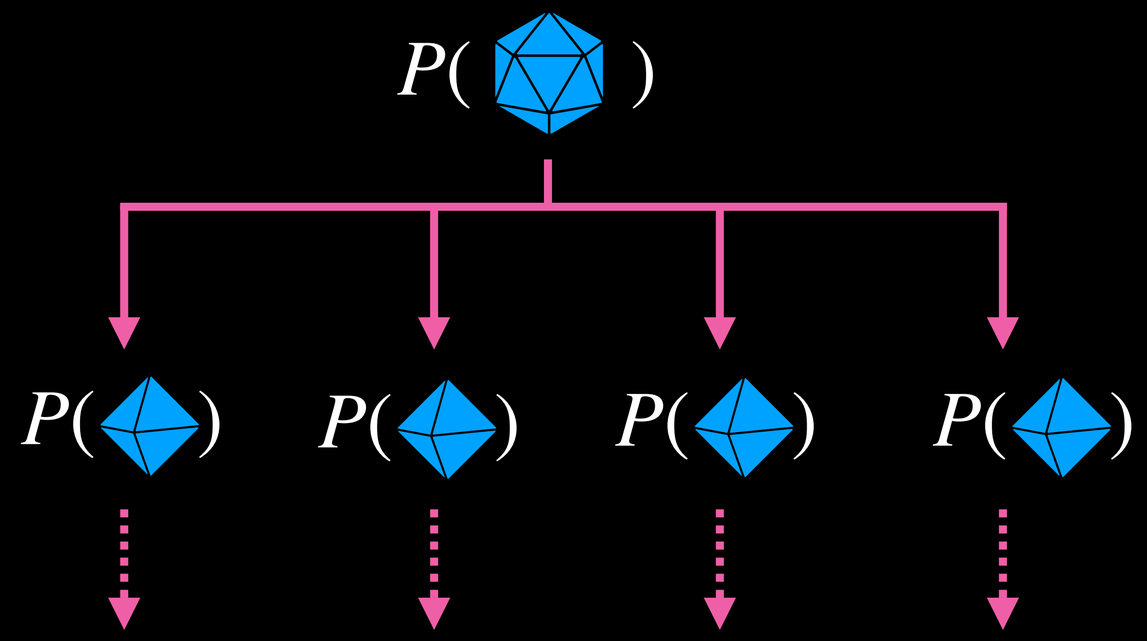
Inclusion/Exclusion

$$P(\varphi_1 \wedge \varphi_2)$$

really: Möbius transformation

$$P(\varphi_1) \qquad P(\varphi_2) \qquad P(\varphi_1 \vee \varphi_2)$$

Dalvi & Suciu Algorithm

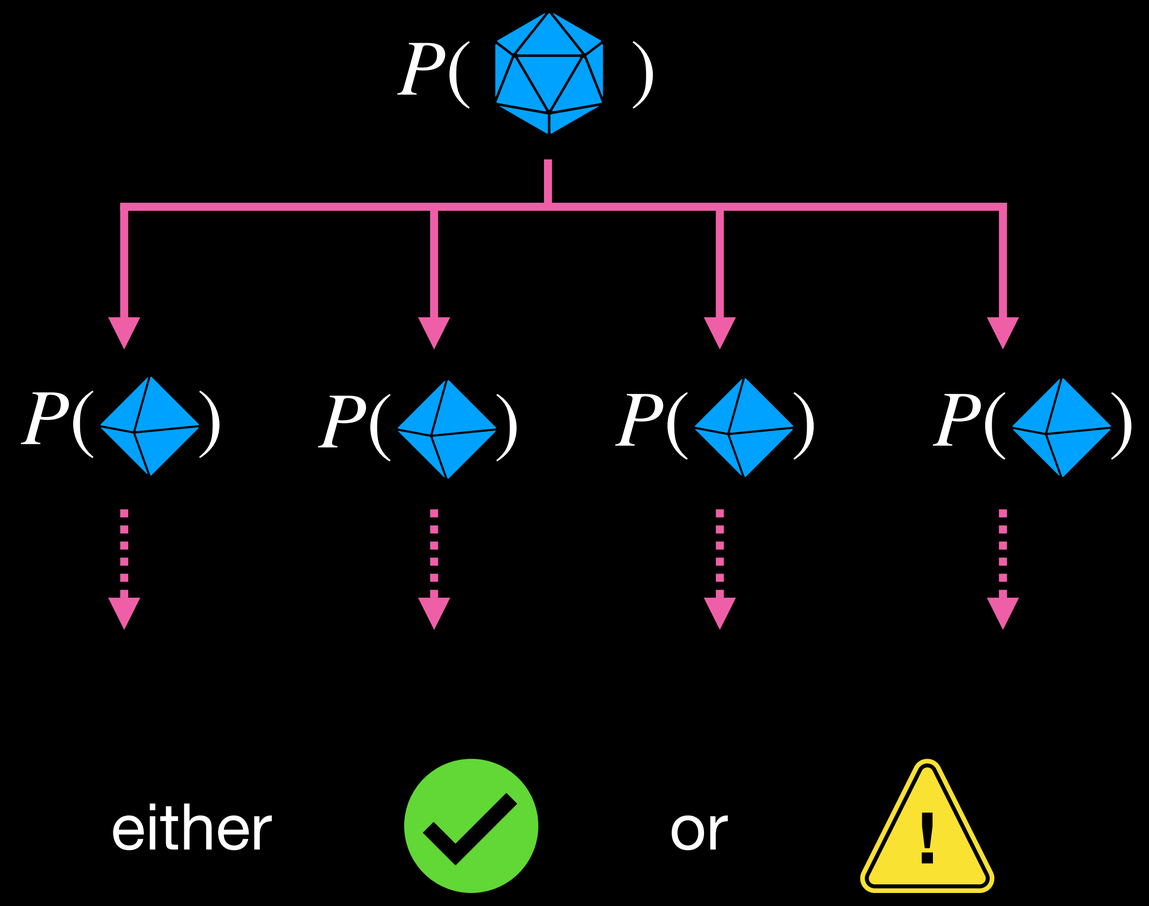


= a large query expression

= smaller queries

either  or 

Dalvi & Suciu Algorithm

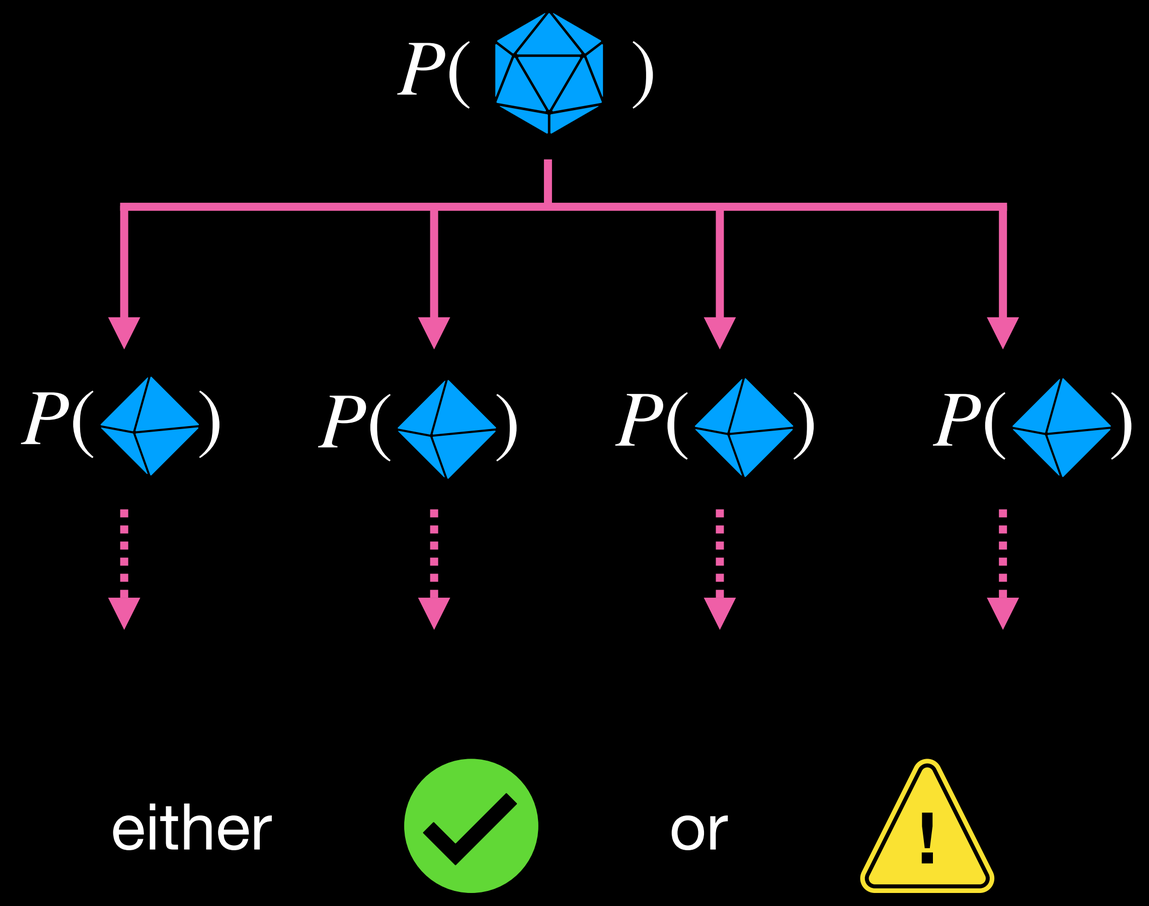



= a large query expression


= smaller queries

Dichotomy [Dalvi & Suciu 2012]:

Dalvi & Suciu Algorithm



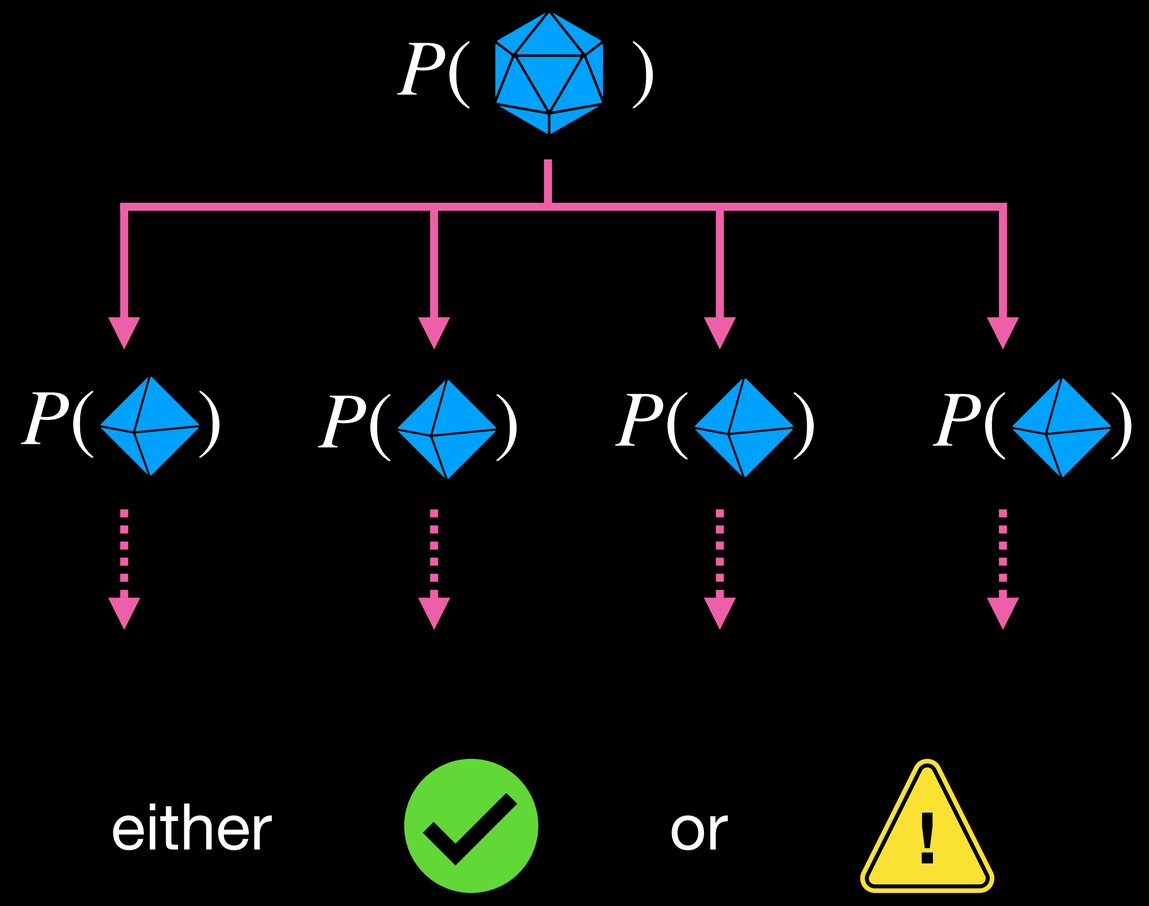
 = a large query expression

 = smaller queries

Dichotomy [Dalvi & Suciu 2012]:

 then $P($  $)$ is in PTIME

Dalvi & Suciu Algorithm



 = a large query expression

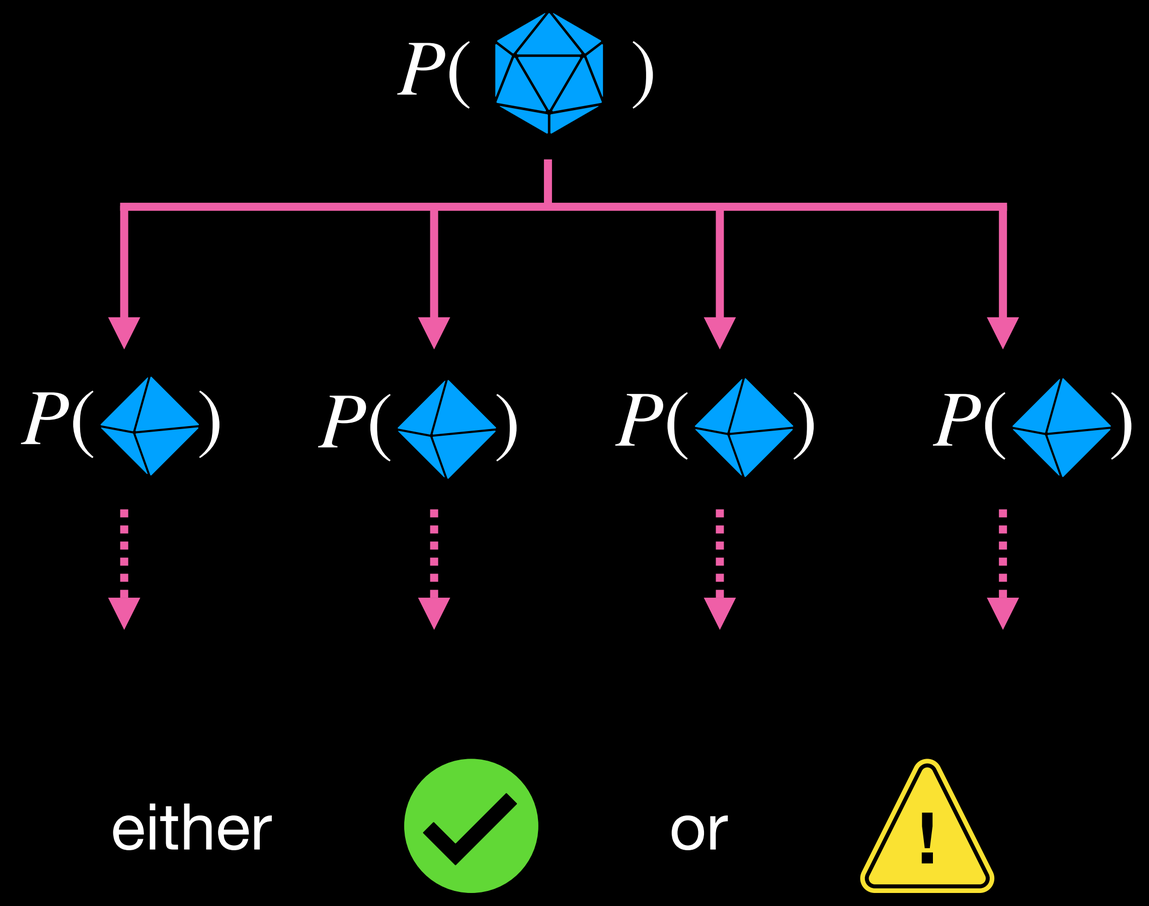
 = smaller queries

Dichotomy [Dalvi & Suciu 2012]:

 then $P($  $)$ is in PTIME

 then $P($  $)$ is #P-hard

Dalvi & Suciu Algorithm

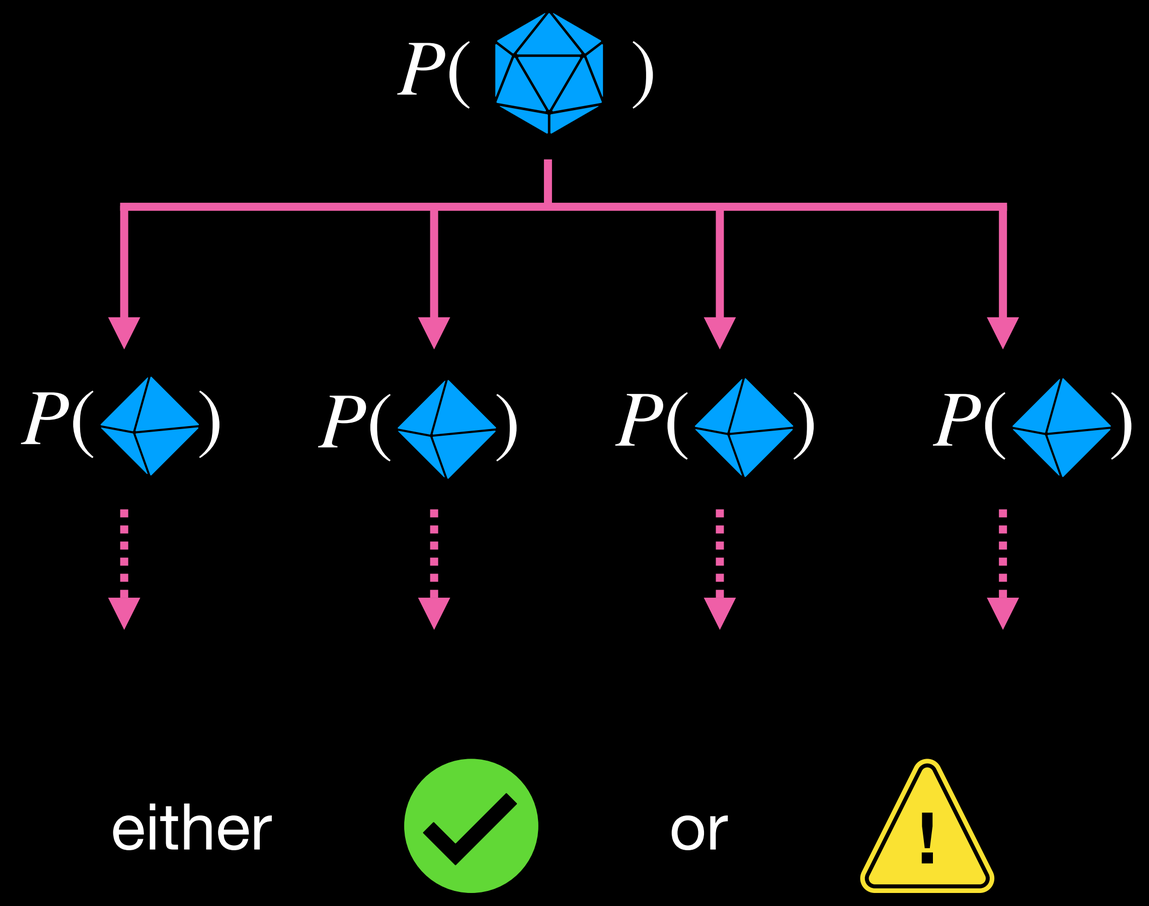


Dichotomy [Dalvi & Suciu 2012]:

then $P($ $)$ is in PTIME

then $P($ $)$ is #P-hard

Dalvi & Suciu Algorithm



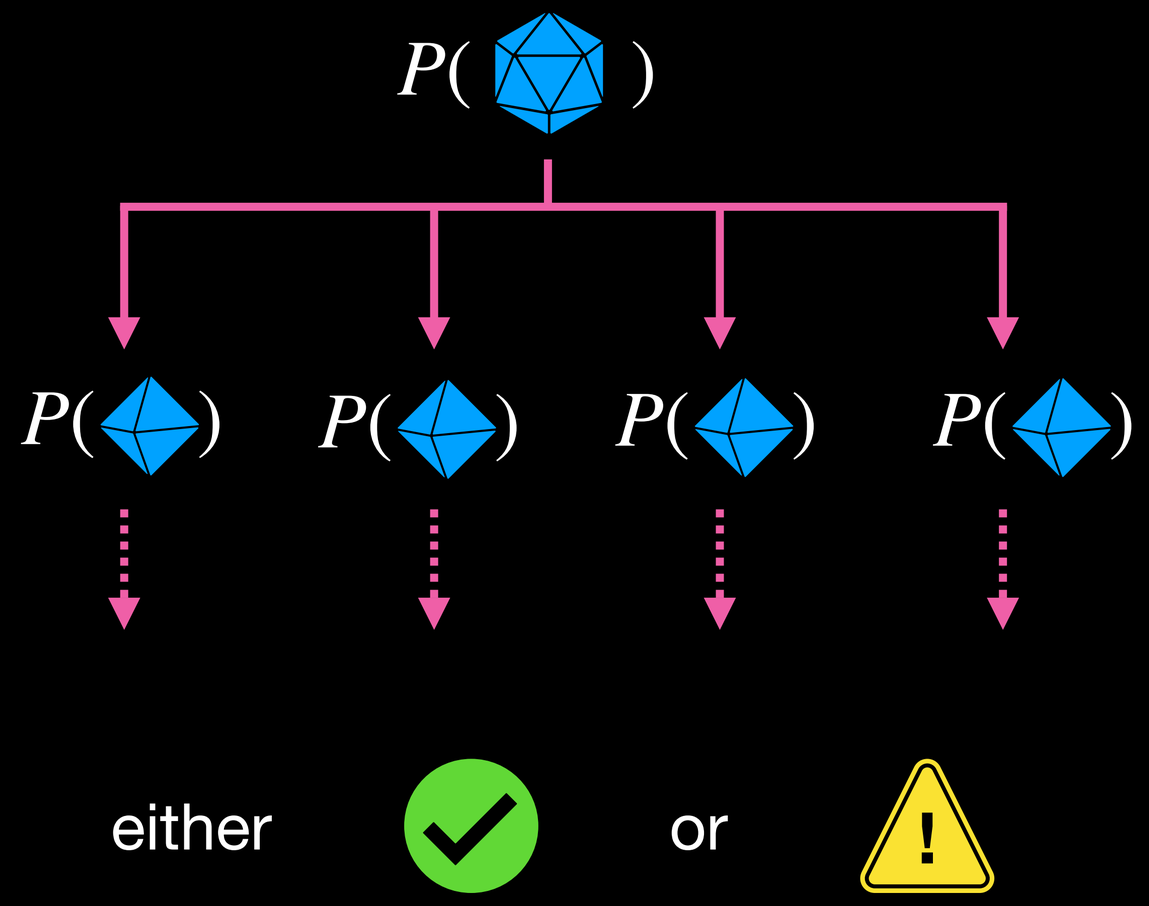
Dichotomy [Dalvi & Suciu 2012]:

then $P($ $)$ is in PTIME

then $P($ $)$ is #P-hard

Boolean Query Evaluation

Dalvi & Suciu Algorithm



Dichotomy [Dalvi & Suciu 2012]**:**

then $P($ $)$ is in PTIME

then $P($ $)$ is #P-hard

Our Result:

we can maintain
the result of this computation
under single-tuple updates
with constant update time

Probabilistic Databases under Updates

Boolean Query Evaluation
and Ranked Enumeration

Christoph Berkholz, Maximilian Merz

Talk at AlMoTh 2022, based on the PODS 2021 paper

Probabilistic Databases under Updates

Boolean Query Evaluation
4 and Ranked Enumeration

Christoph Berkholz, Maximilian Merz

Talk at AlMoTh 2022, based on the PODS 2021 paper

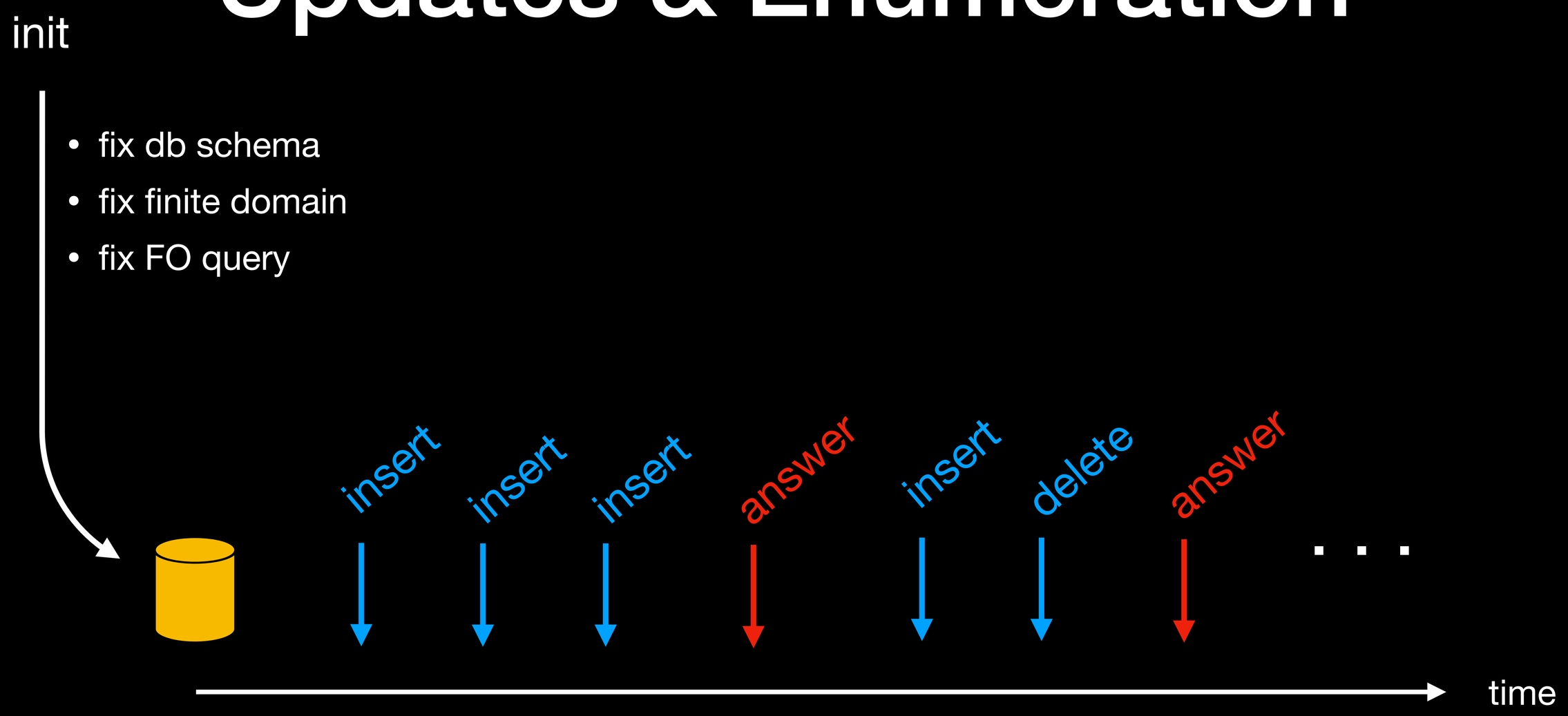
✔ Probabilistic Databases ✔ under Updates

✔ Boolean Query Evaluation
➡ and Ranked Enumeration

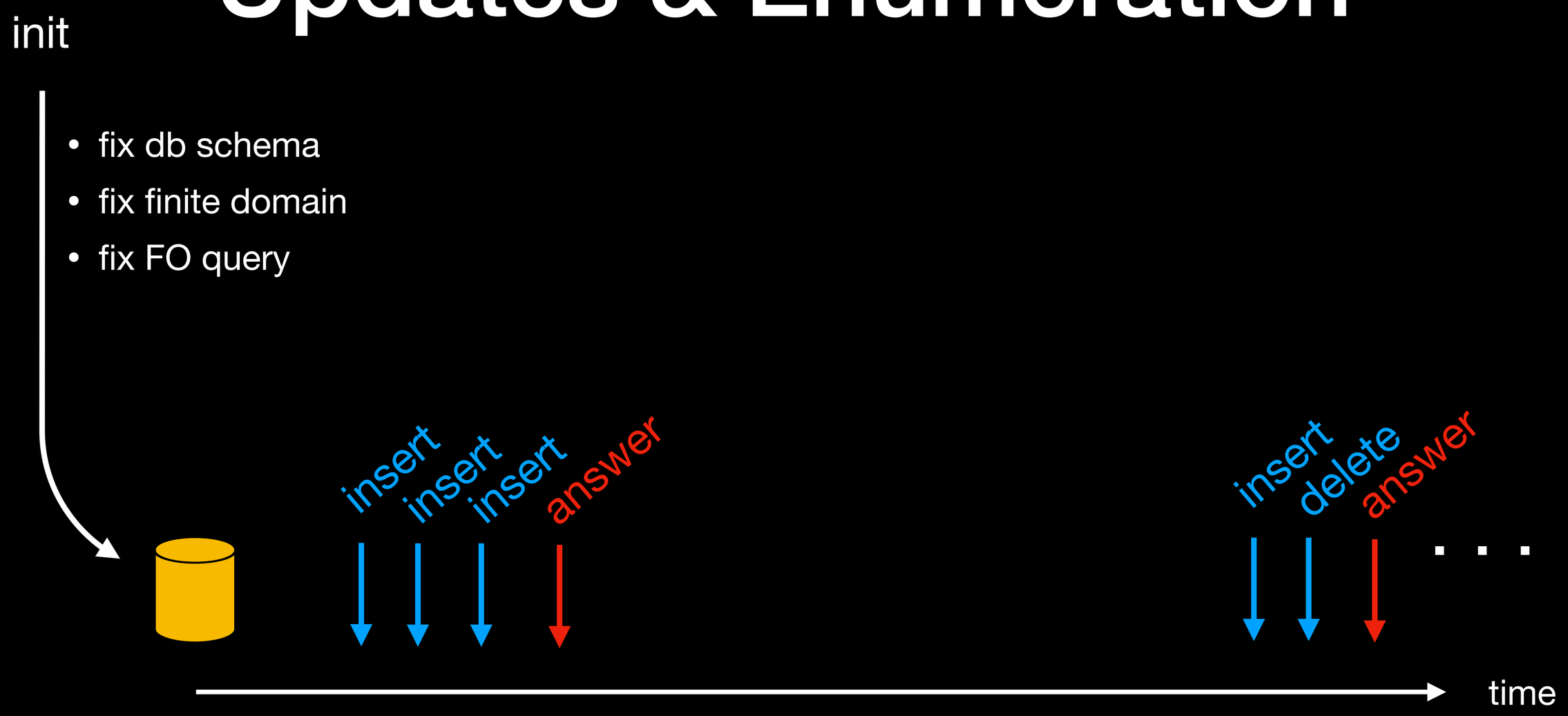
Christoph Berkholz, Maximilian Merz

Talk at AlMoTh 2022, based on the PODS 2021 paper

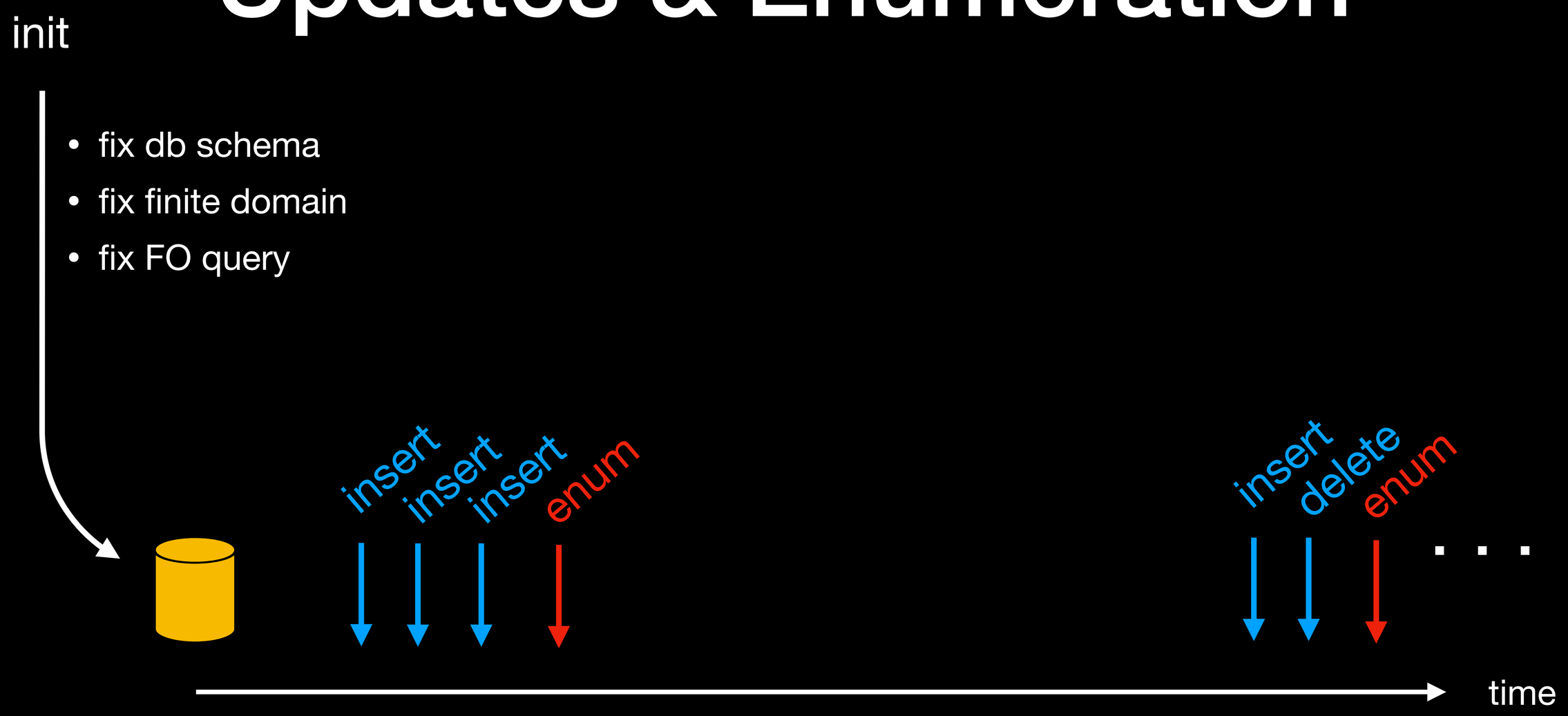
Updates & Enumeration



Updates & Enumeration



Updates & Enumeration

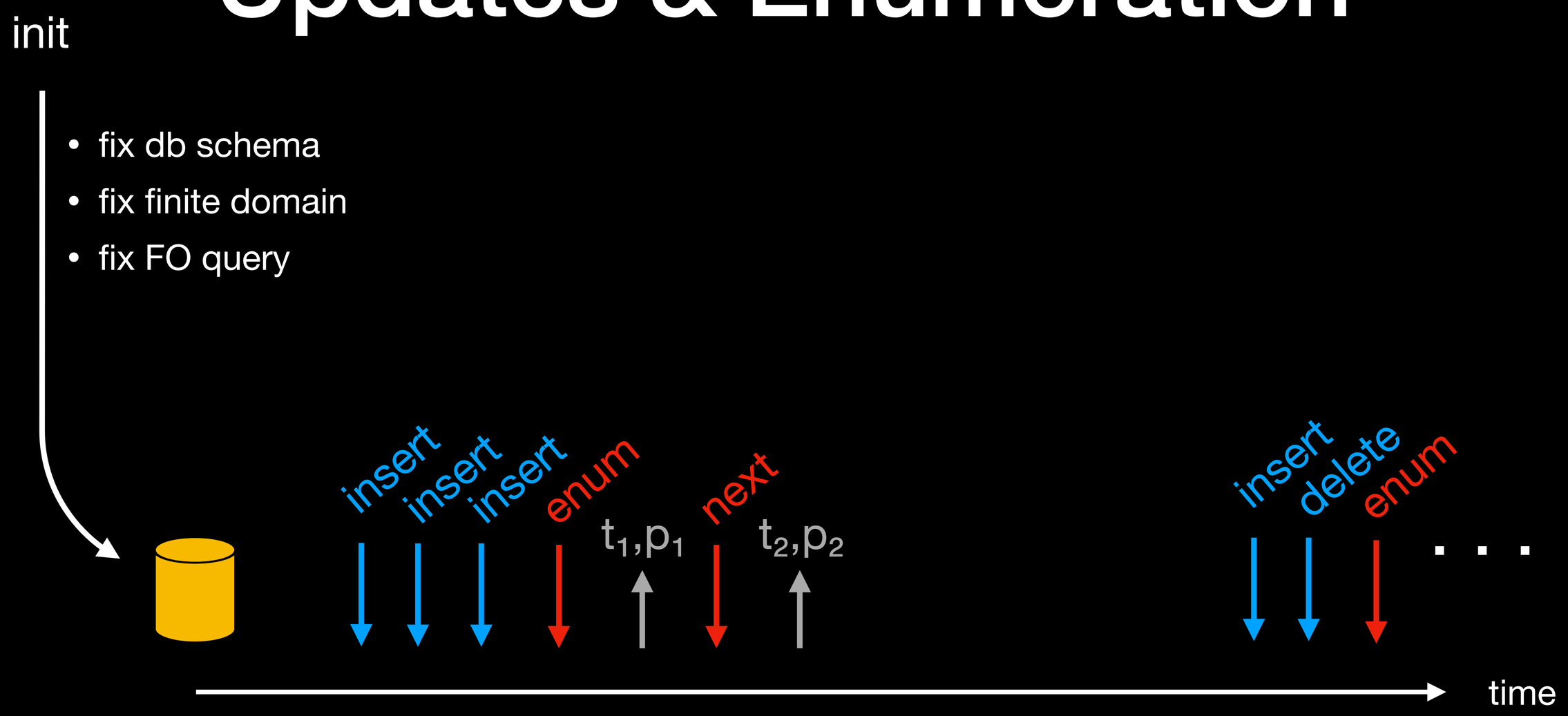


Updates & Enumeration

init

- fix db schema
- fix finite domain
- fix FO query

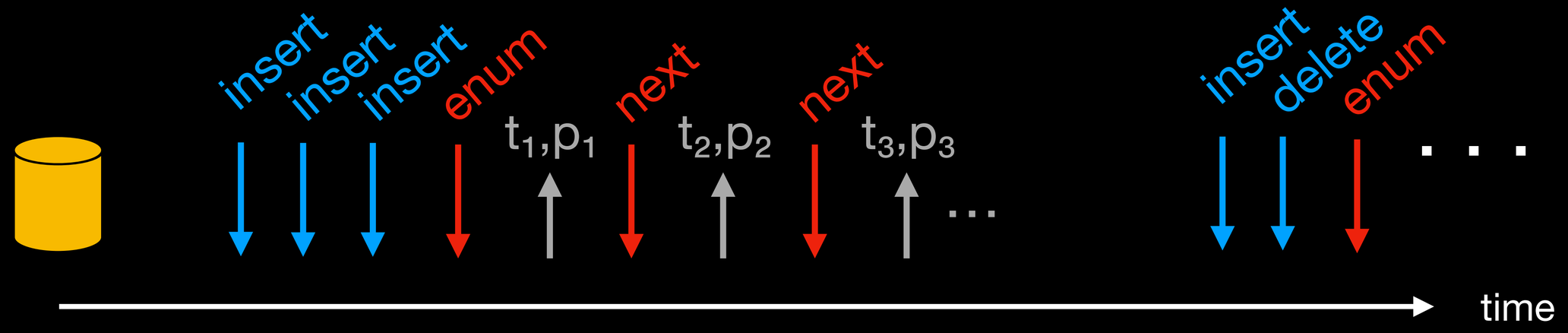
Updates & Enumeration



Updates & Enumeration

init

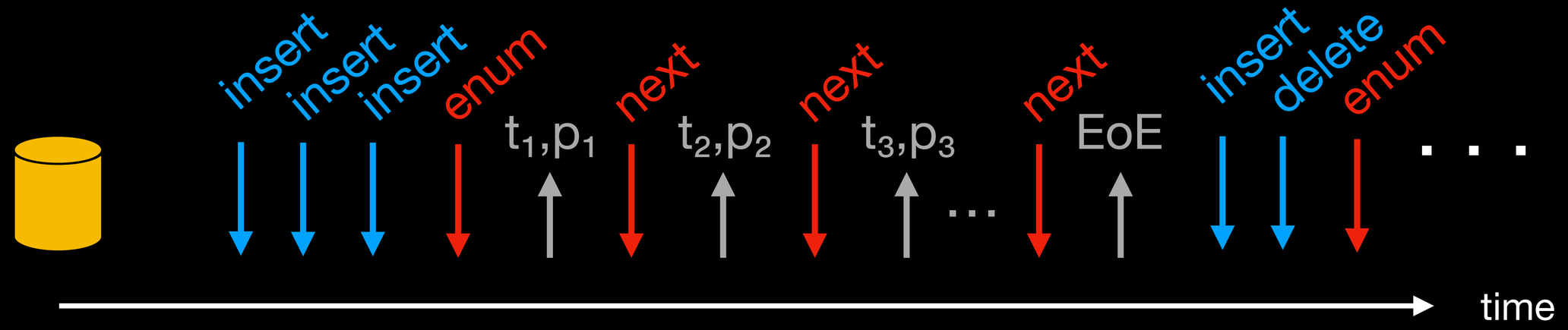
- fix db schema
- fix finite domain
- fix FO query



Updates & Enumeration

init

- fix db schema
- fix finite domain
- fix FO query

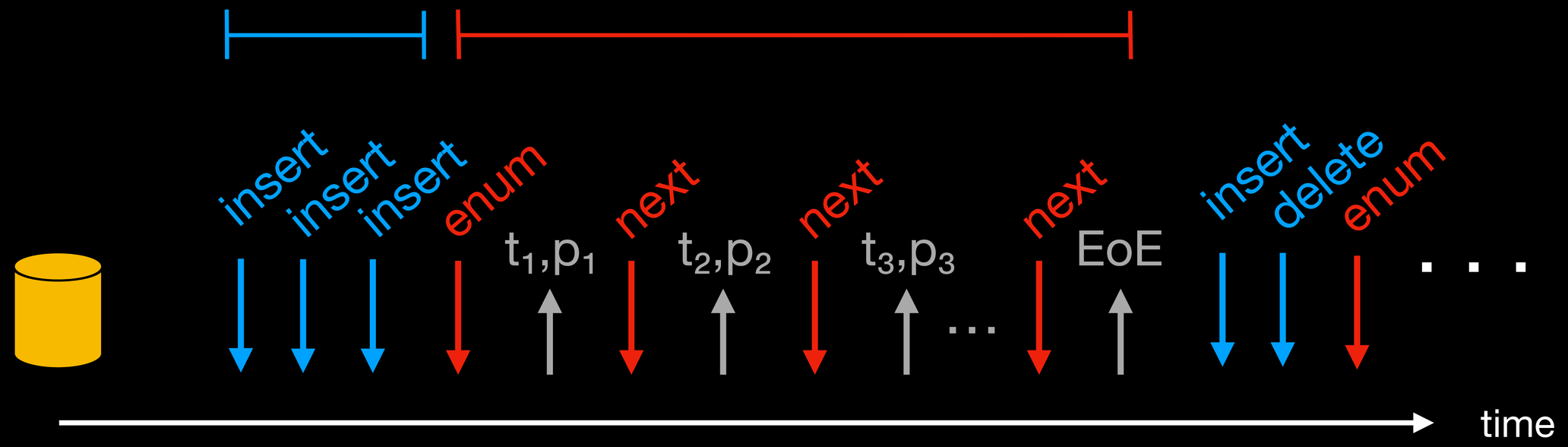


Updates & Enumeration

init

- fix db schema
- fix finite domain
- fix FO query

- update phase: insert / delete
- enumeration phase

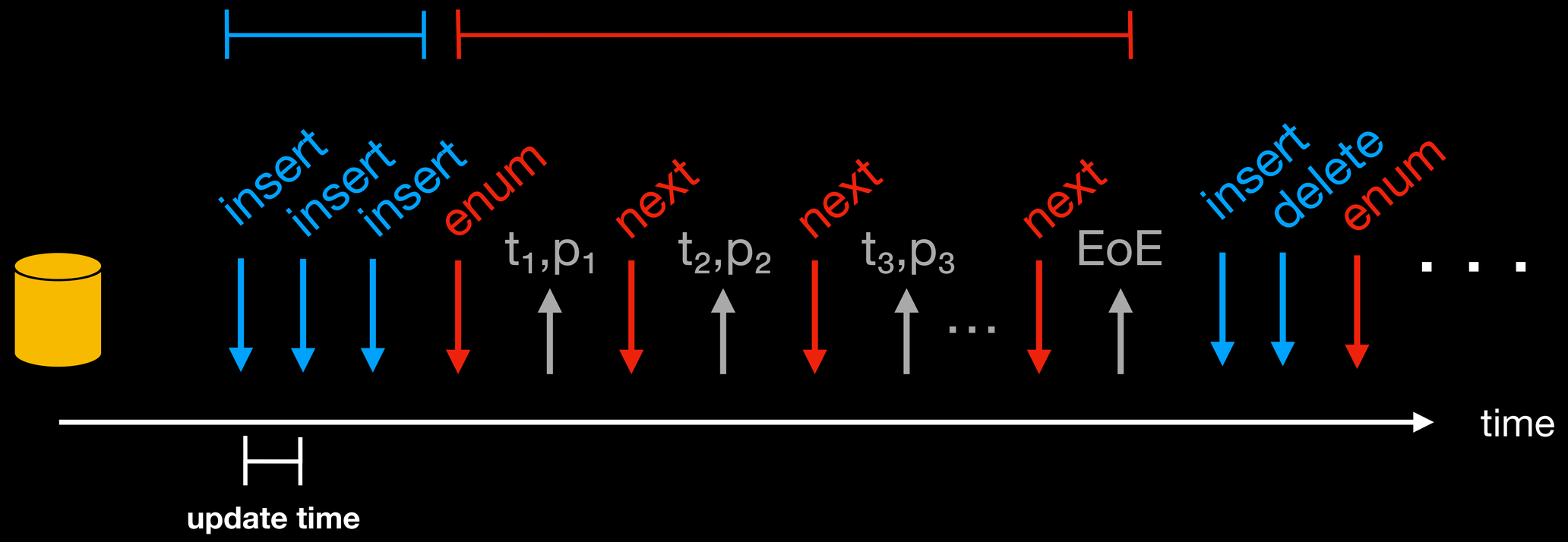


Updates & Enumeration

init

- fix db schema
- fix finite domain
- fix FO query

- update phase: insert / delete
- enumeration phase

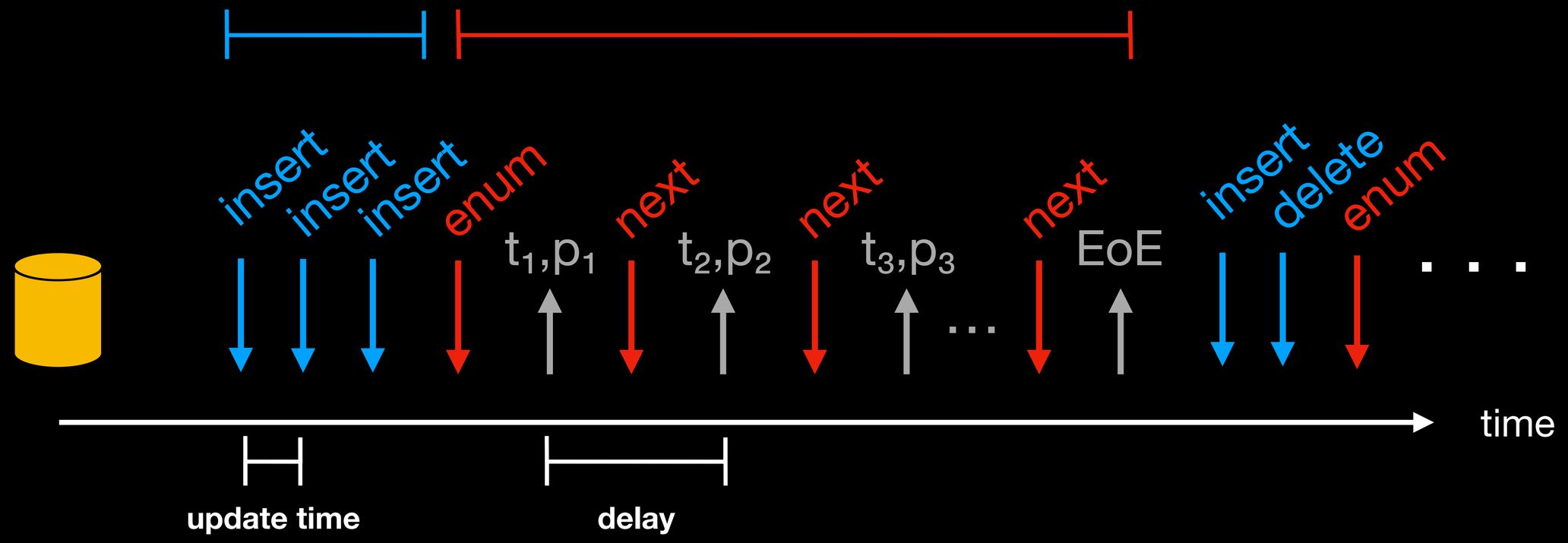


Updates & Enumeration

init

- fix db schema
- fix finite domain
- fix FO query

- update phase: insert / delete
- enumeration phase

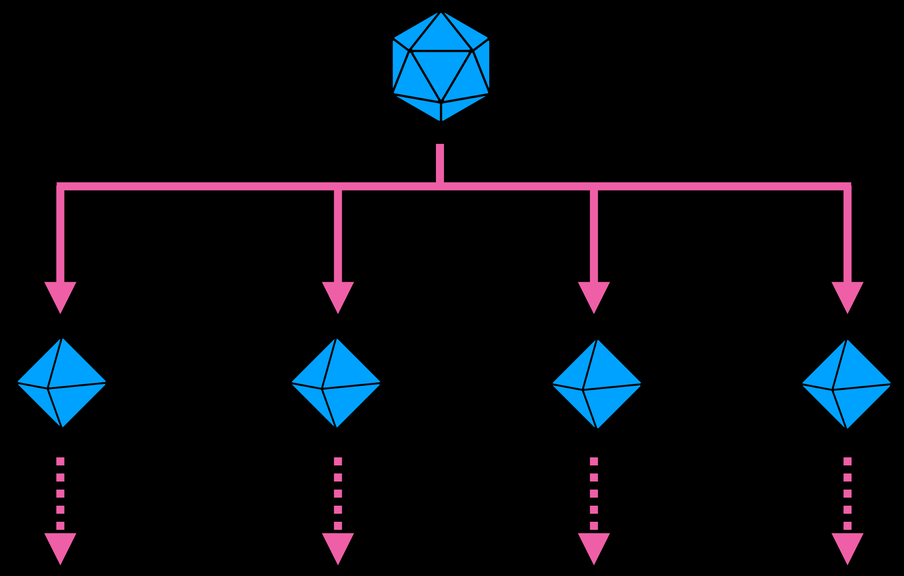


Ranked Enumeration: Our Algorithm

Ranked Enumeration: Our Algorithm

Berkholz, Keppeler, Schweikardt: (2017)

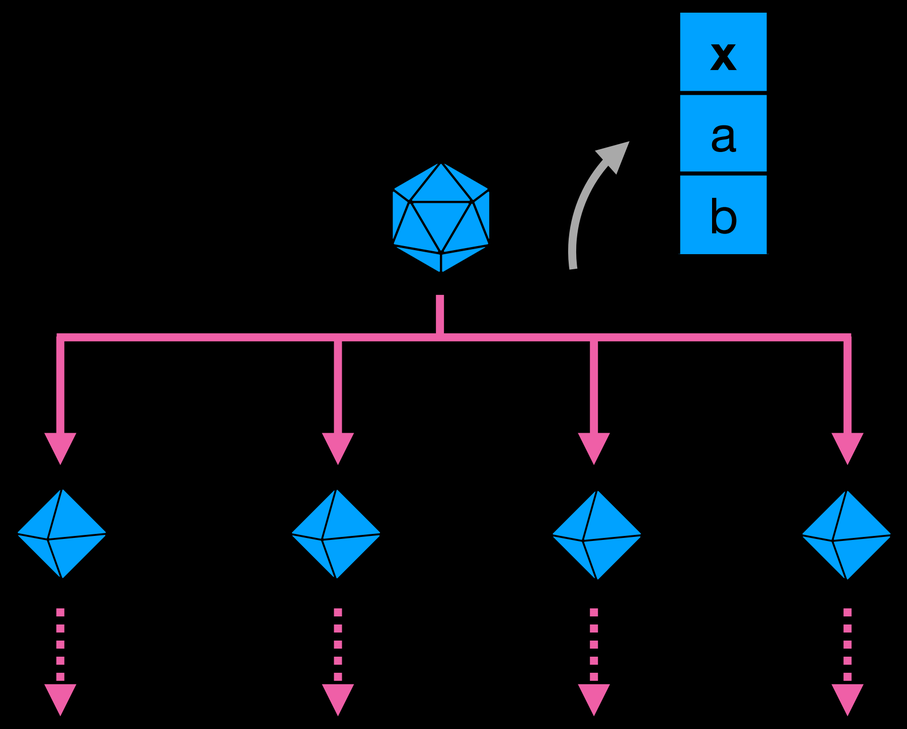
…on deterministic DB's



Ranked Enumeration: Our Algorithm

Berkholz, Keppeler, Schweikardt: (2017)

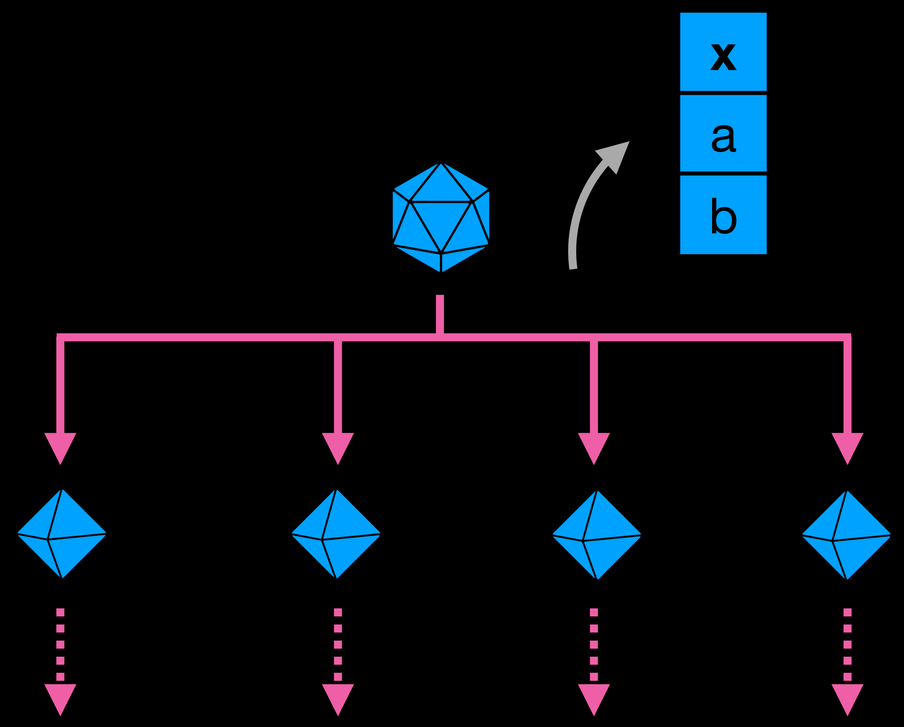
…on deterministic DB's



Ranked Enumeration: Our Algorithm

Berkholz, Keppeler, Schweikardt: (2017)

…on deterministic DB's



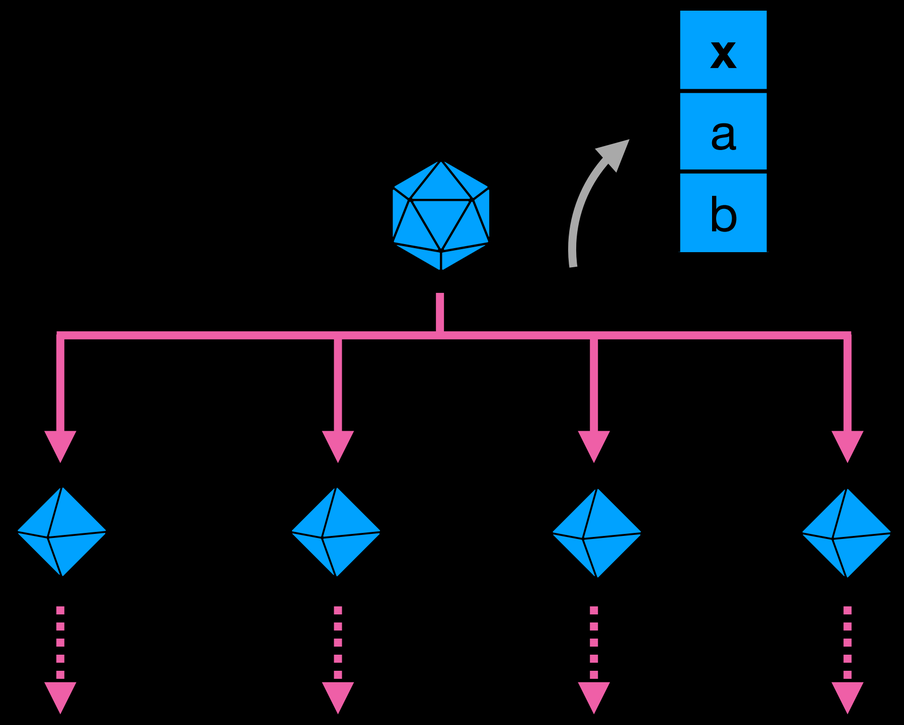
Our Result:

…on tuple-independent probabilistic DB's

Ranked Enumeration: Our Algorithm

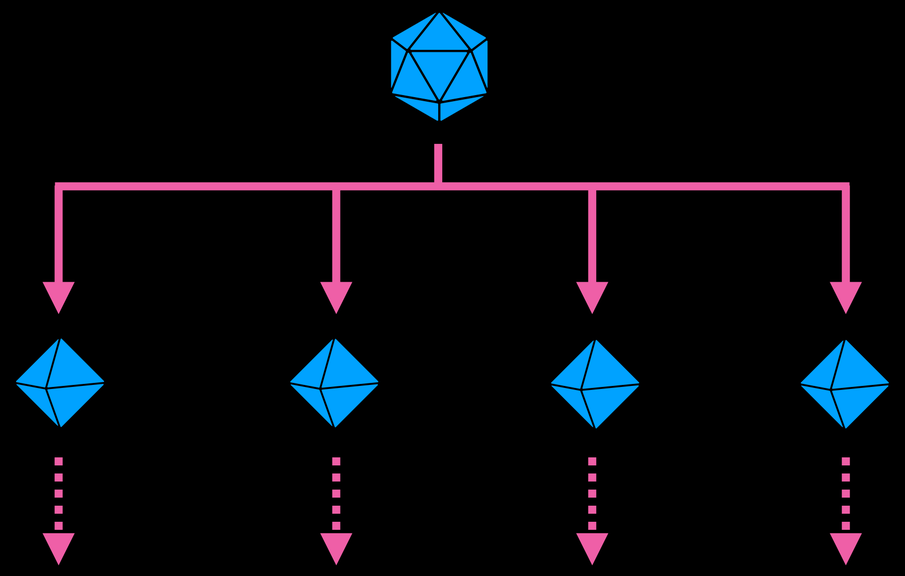
Berkholz, Keppeler, Schweikardt: (2017)

…on deterministic DB's



Our Result:

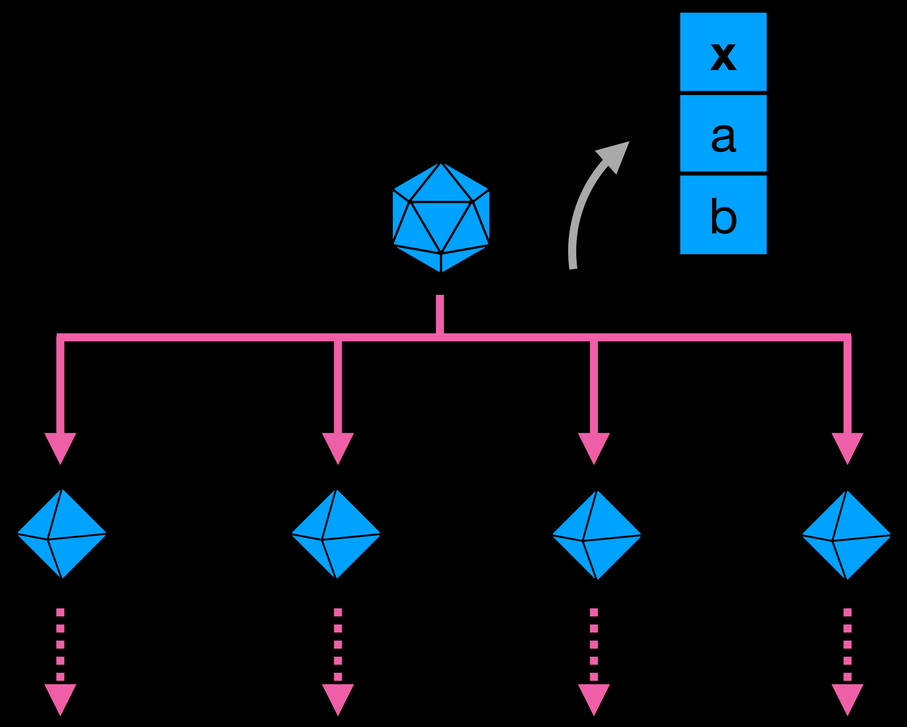
…on tuple-independent probabilistic DB's



Ranked Enumeration: Our Algorithm

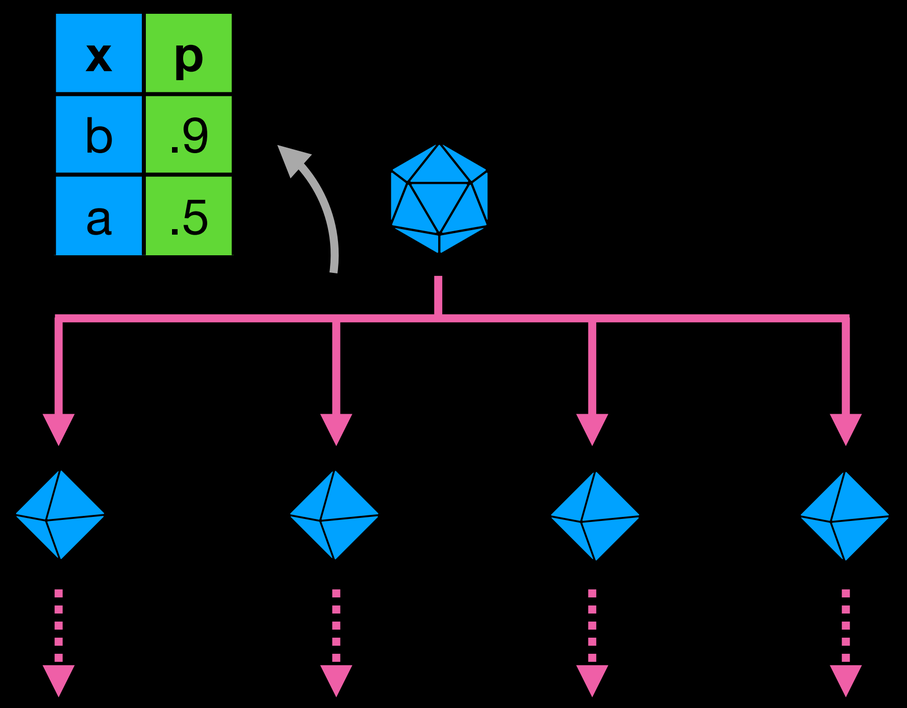
Berkholz, Keppeler, Schweikardt: (2017)

…on deterministic DB's



Our Result:

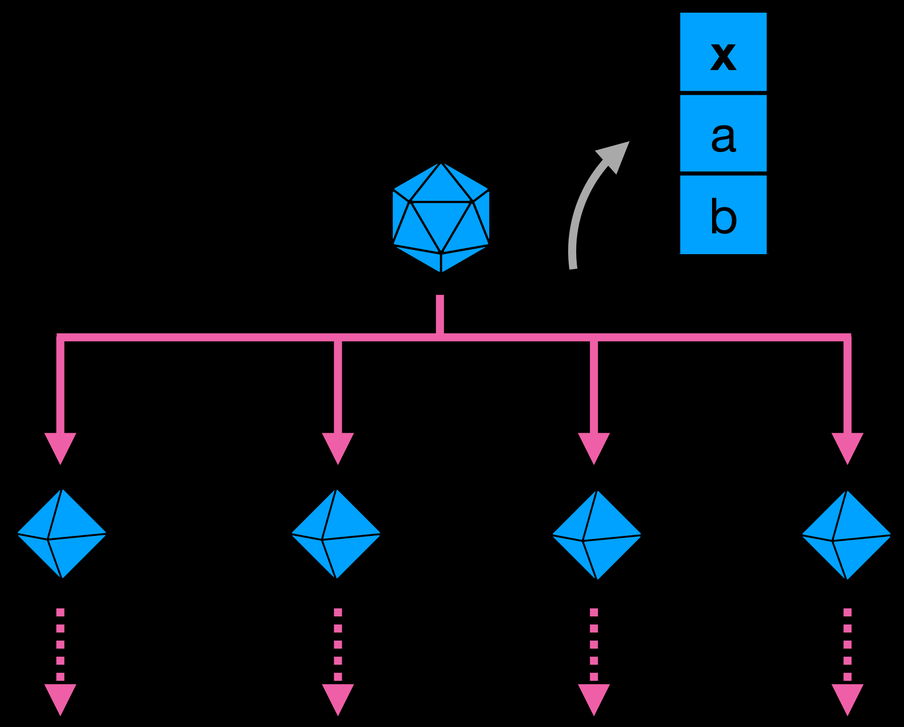
…on tuple-independent probabilistic DB's



Ranked Enumeration: Our Algorithm

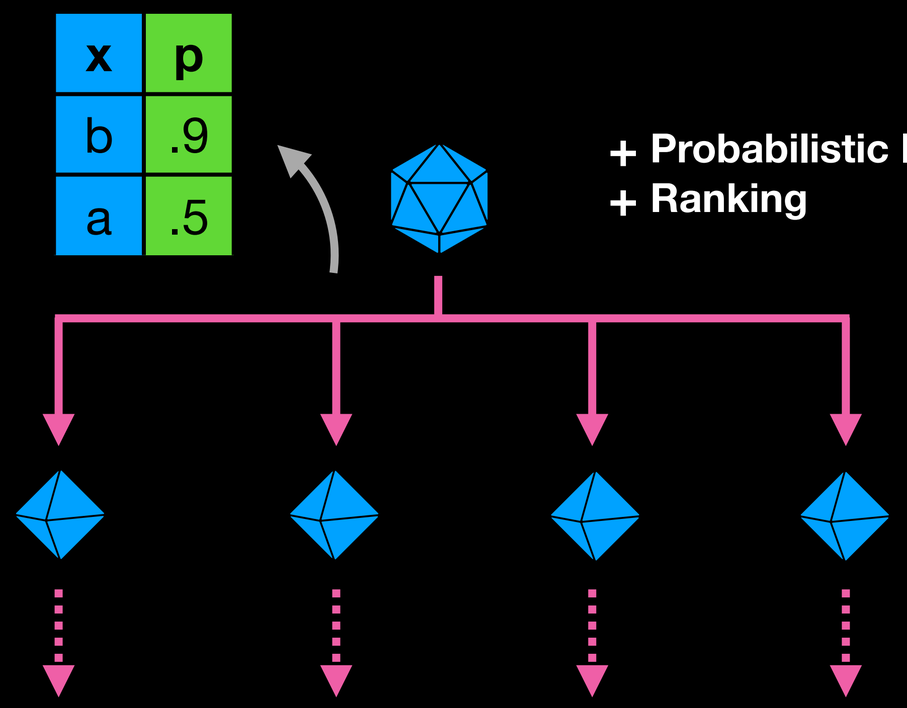
Berkholz, Keppeler, Schweikardt: (2017)

…on deterministic DB's



Our Result:

…on tuple-independent probabilistic DB's



+ Probabilistic Inference
+ Ranking

Ranked Enumeration: Our Algorithm

Berkholz, Keppeler, Schweikardt: (2017)

…on deterministic DB's

Our Result:

…on tuple-independent probabilistic DB's

Ranked Enumeration: Our Algorithm

Berkholz, Keppeler, Schweikardt: (2017)

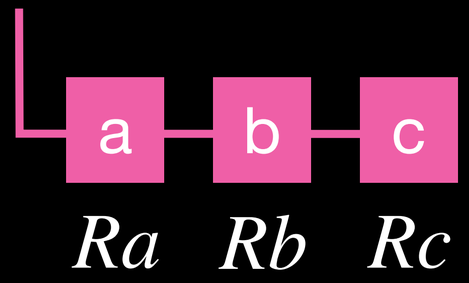
…on deterministic DB's

Our Result:

…on tuple-independent probabilistic DB's

$$\varphi(x) = Rx$$

$x = ?$



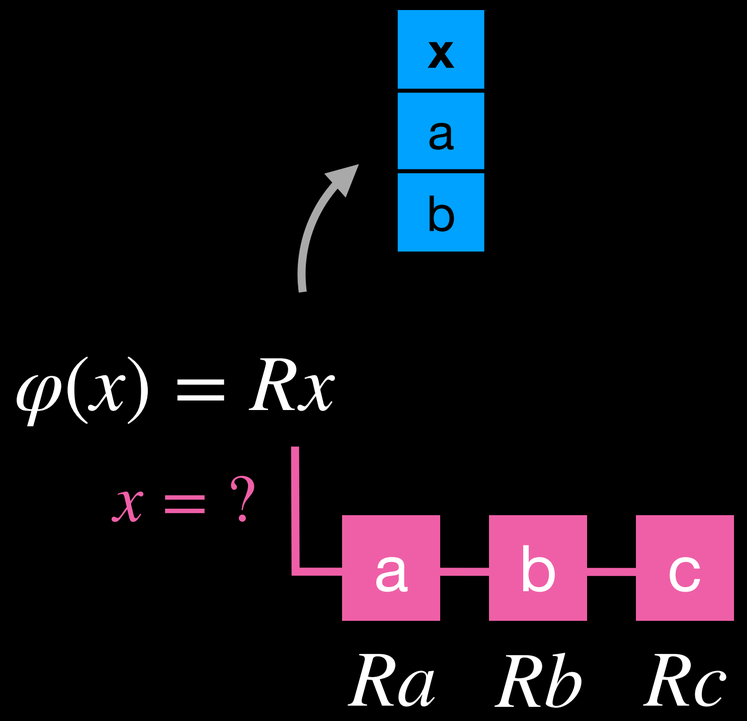
Ranked Enumeration: Our Algorithm

Berkholz, Keppeler, Schweikardt: (2017)

…on deterministic DB's

Our Result:

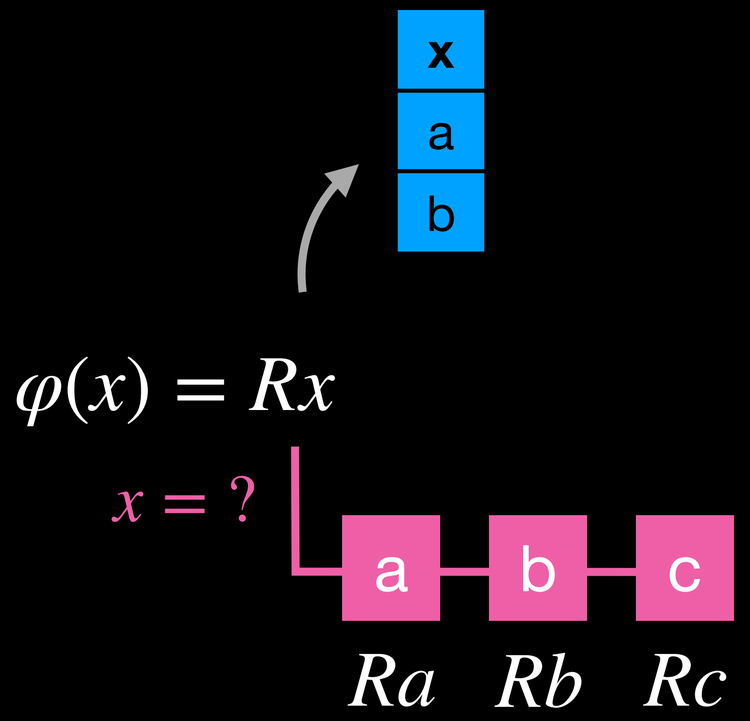
…on tuple-independent probabilistic DB's



Ranked Enumeration: Our Algorithm

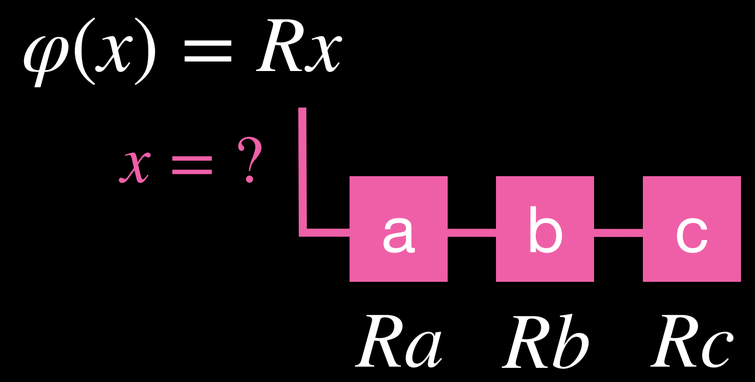
Berkholz, Keppeler, Schweikardt: (2017)

…on deterministic DB's



Our Result:

…on tuple-independent probabilistic DB's



Ranked Enumeration: Our Algorithm

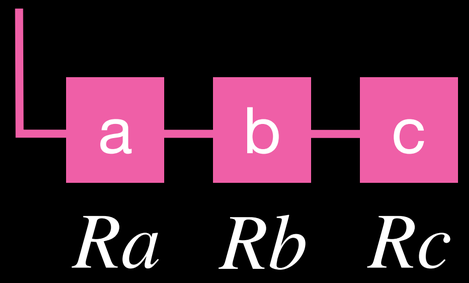
Berkholz, Keppeler, Schweikardt: (2017)

…on deterministic DB's

x
a
b

$\varphi(x) = Rx$

$x = ?$



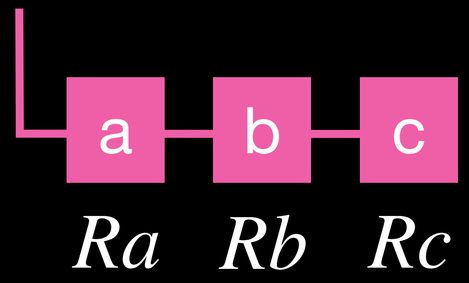
Our Result:

…on tuple-independent probabilistic DB's

x	p
b	.9
a	.5

$\varphi(x) = Rx$

$x = ?$



Ranked Enumeration: Our Algorithm

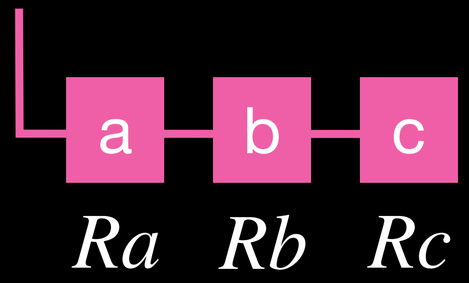
Berkholz, Keppeler, Schweikardt: (2017)

…on deterministic DB's

x
a
b

$\varphi(x) = Rx$

$x = ?$



Our Result:

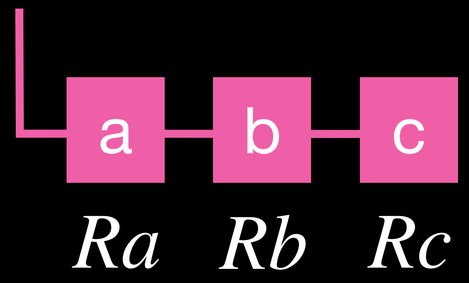
…on tuple-independent probabilistic DB's

x	p
b	.9
a	.5

$\varphi(x) = Rx$

+ priority queue

$x = ?$



Ranked Enumeration: Our Algorithm

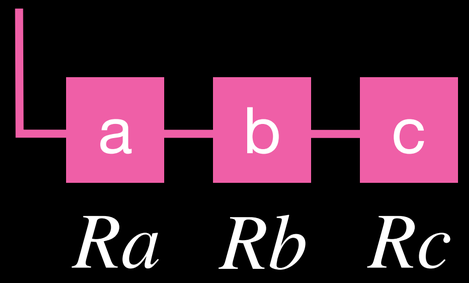
Berkholz, Keppeler, Schweikardt: (2017)

…on deterministic DB's

x
a
b

$\varphi(x) = Rx$

$x = ?$



Our Result:

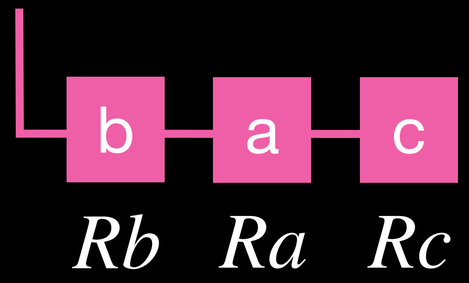
…on tuple-independent probabilistic DB's

x	p
b	.9
a	.5

$\varphi(x) = Rx$

+ priority queue

$x = ?$



Ranked Enumeration: Our Algorithm

Berkholz, Keppeler, Schweikardt: (2017)

…on deterministic DB's

Our Result:

…on tuple-independent probabilistic DB's

Ranked Enumeration: Our Algorithm

Berkholz, Keppeler, Schweikardt: (2017)

…on deterministic DB's

Our Result:

…on tuple-independent probabilistic DB's

$$\varphi(x, y) = Rx \wedge Sy$$

Ranked Enumeration: Our Algorithm

Berkholz, Keppeler, Schweikardt: (2017)

…on deterministic DB's

Our Result:

…on tuple-independent probabilistic DB's

$$\varphi(x, y) = Rx \wedge Sy$$

$$Rx \qquad Sy$$

Ranked Enumeration: Our Algorithm

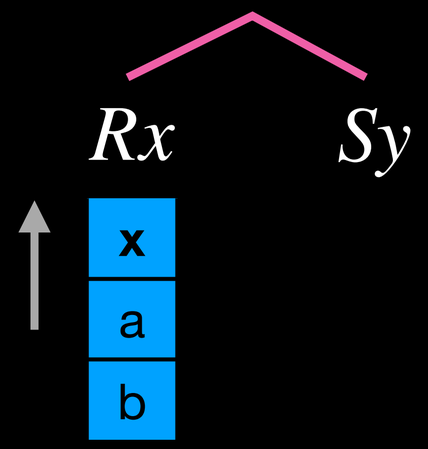
Berkholz, Keppeler, Schweikardt: (2017)

…on deterministic DB's

Our Result:

…on tuple-independent probabilistic DB's

$$\varphi(x, y) = Rx \wedge Sy$$



Ranked Enumeration: Our Algorithm

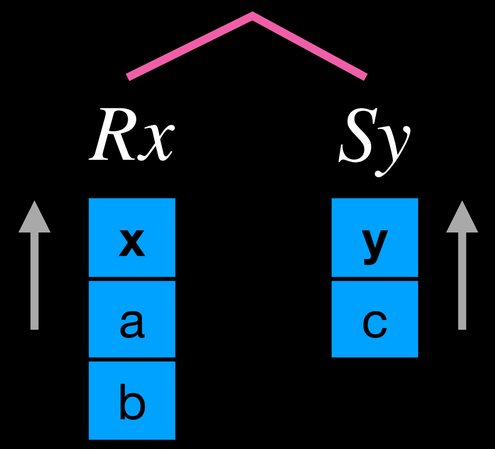
Berkholz, Keppeler, Schweikardt: (2017)

…on deterministic DB's

Our Result:

…on tuple-independent probabilistic DB's

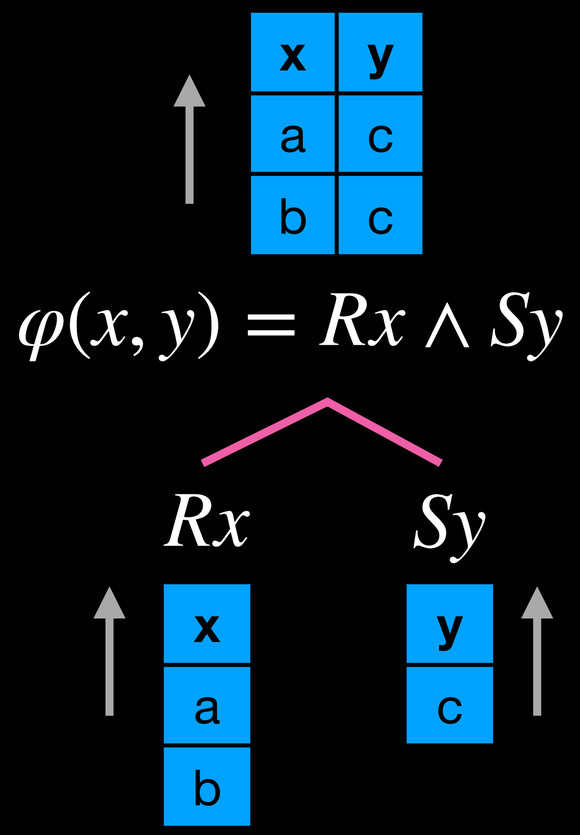
$$\varphi(x, y) = Rx \wedge Sy$$



Ranked Enumeration: Our Algorithm

Berkholz, Keppeler, Schweikardt: (2017)

…on deterministic DB's



Our Result:

…on tuple-independent probabilistic DB's

Ranked Enumeration: Our Algorithm

Berkholz, Keppeler, Schweikardt: (2017)

…on deterministic DB's

x	y
a	c
b	c

$$\varphi(x, y) = Rx \wedge Sy$$

Rx Sy

x
a
b

y
c

Our Result:

…on tuple-independent probabilistic DB's

$$\varphi(x, y) = Rx \wedge Sy$$

Rx Sy

Ranked Enumeration: Our Algorithm

Berkholz, Keppeler, Schweikardt: (2017)

…on deterministic DB's

x	y
a	c
b	c

$$\varphi(x, y) = Rx \wedge Sy$$

Rx Sy

x
a
b

y
c

Our Result:

…on tuple-independent probabilistic DB's

$$\varphi(x, y) = Rx \wedge Sy$$

Rx Sy

x	p
b	.9
a	.5

Ranked Enumeration: Our Algorithm

Berkholz, Keppeler, Schweikardt: (2017)

…on deterministic DB's

x	y
a	c
b	c

$$\varphi(x, y) = Rx \wedge Sy$$

Rx

x
a
b

Sy

y
c

Our Result:

…on tuple-independent probabilistic DB's

$$\varphi(x, y) = Rx \wedge Sy$$

Rx

x	p
b	.9
a	.5

Sy

y	p
c	.2

Ranked Enumeration: Our Algorithm

Berkholz, Keppeler, Schweikardt: (2017)

…on deterministic DB's

x	y
a	c
b	c

$$\varphi(x, y) = Rx \wedge Sy$$

Rx

x
a
b

Sy

y
c

Our Result:

…on tuple-independent probabilistic DB's

must be independent!

$$\varphi(x, y) = Rx \wedge Sy$$

Rx

x	p
b	.9
a	.5

Sy

y	p
c	.2

Ranked Enumeration: Our Algorithm

Berkholz, Keppeler, Schweikardt: (2017)

…on deterministic DB's

x	y
a	c
b	c

$$\varphi(x, y) = Rx \wedge Sy$$

Rx

x
a
b

Sy

y
c

Our Result:

…on tuple-independent probabilistic DB's

x	y	p
a	c	.9 • .2
b	c	.5 • .2

$$\varphi(x, y) = Rx \wedge Sy$$

Rx

x	p
b	.9
a	.5

Sy

y	p
c	.2

must be independent!

Ranked Enumeration: Our Algorithm

Berkholz, Keppeler, Schweikardt: (2017)

…on deterministic DB's

Our Result:

…on tuple-independent probabilistic DB's

Ranked Enumeration: Our Algorithm

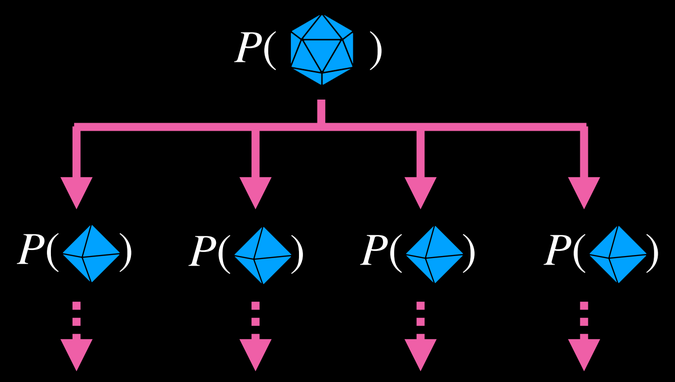
Berkholz, Keppeler, Schweikardt: (2017)

…on deterministic DB's

Our Result:

…on tuple-independent probabilistic DB's

Ranked Enumeration: Our Algorithm



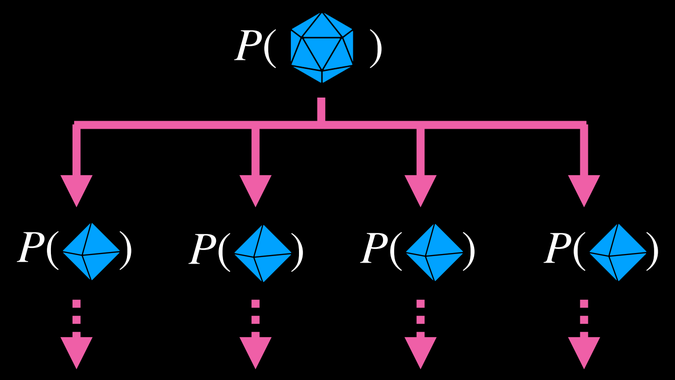
Berkholz, Keppeler, Schweikardt: (2017)

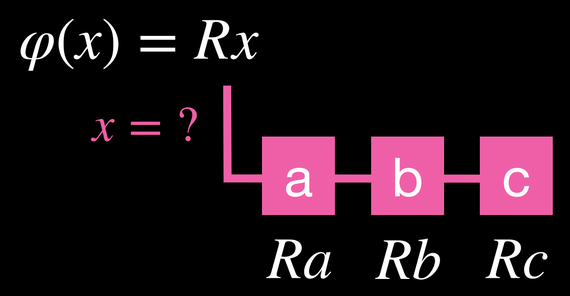
…on deterministic DB's

Our Result:

…on tuple-independent probabilistic DB's

Ranked Enumeration: Our Algorithm





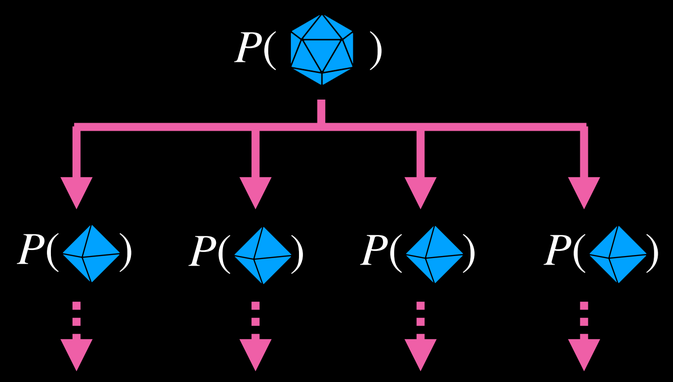
Berkholz, Keppeler, Schweikardt: (2017)

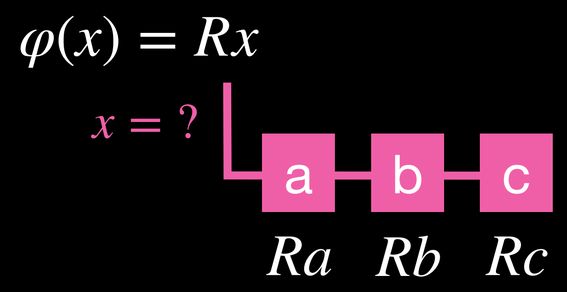
…on deterministic DB's

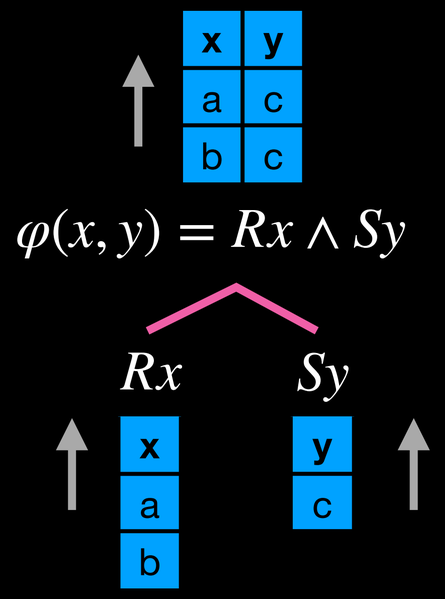
Our Result:

…on tuple-independent probabilistic DB's

Ranked Enumeration: Our Algorithm







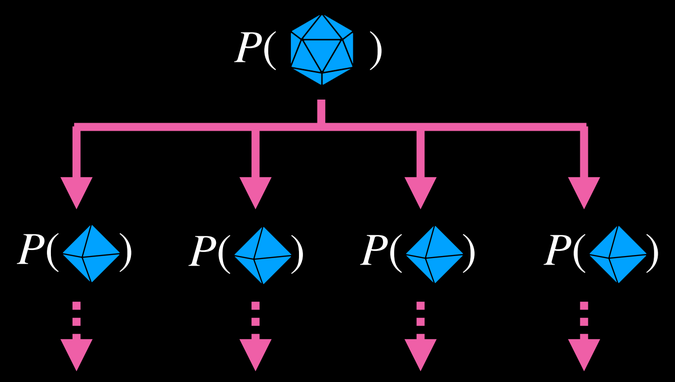
Berkholz, Keppeler, Schweikardt: (2017)

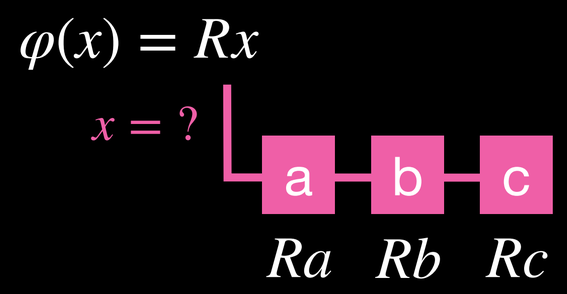
…on deterministic DB's

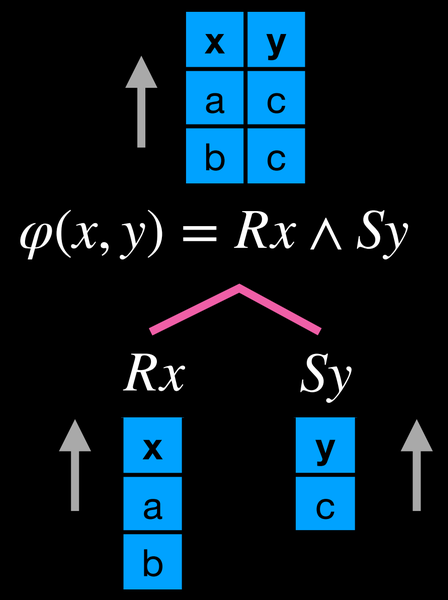
Our Result:

…on tuple-independent probabilistic DB's

Ranked Enumeration: Our Algorithm







Berkholz, Keppeler, Schweikardt: (2017)

…on deterministic DB's

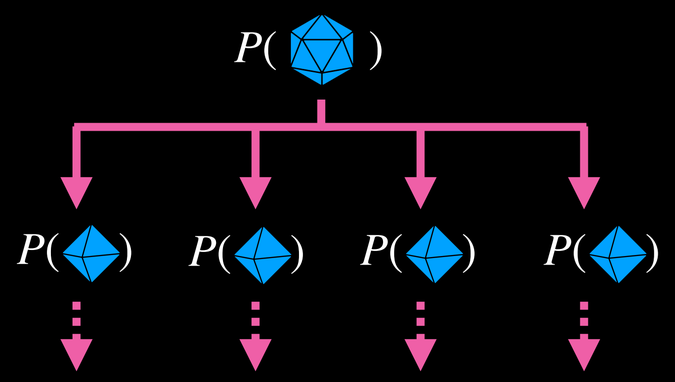
Query is q-hierarchical ⇒

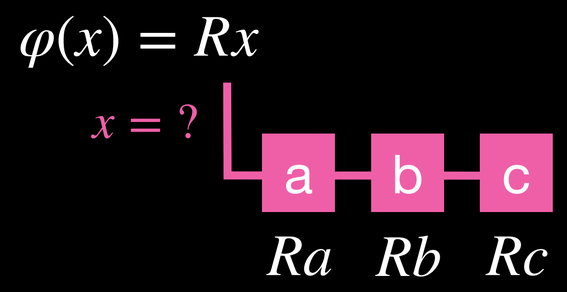
Dynamic Enumeration with O(1) delay and O(1) updates

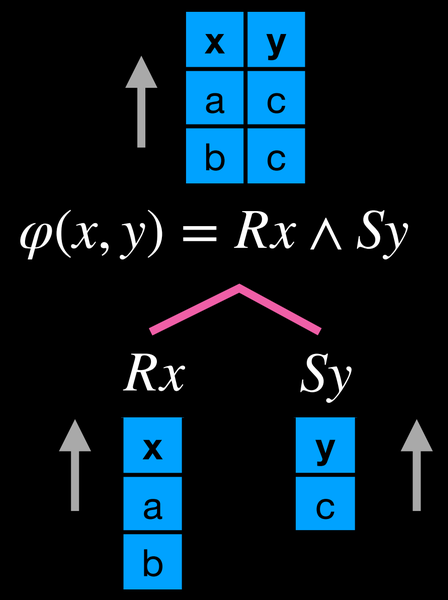
Our Result:

…on tuple-independent probabilistic DB's

Ranked Enumeration: Our Algorithm







Berkholz, Keppeler, Schweikardt: (2017)

…on deterministic DB's

Query is q-hierarchical ⇒

Dynamic Enumeration with O(1) delay and O(1) updates

Our Result:

…on tuple-independent probabilistic DB's

Query is q-hierarchical ⇒

Dynamic Ranked Enumeration with O(log n) delay and O(log n) updates

Ranked Enumeration for non-repeating CQs

Ranked Enumeration for non-repeating CQs

Dichotomy:
[Berkholz, Keppeler, Schweikardt 2017]

…on deterministic DB's

Non-repeating CQs

Ranked Enumeration for non-repeating CQs

Dichotomy:
[Berkholz, Keppeler, Schweikardt 2017]

…on deterministic DB's

Non-repeating CQs

If the OMv-conjecture holds:

Query is not q-hierarchical ⇒

Dynamic Enumeration with
less than $O(n^{1/2})$ delay and $O(n^{1/2})$ updates
is **impossible**

Ranked Enumeration for non-repeating CQs

Dichotomy:
[Berkholz, Keppeler, Schweikardt 2017]

…on deterministic DB's

Non-repeating CQs

If the OMv-conjecture holds:

Query is not q-hierarchical ⇒

Dynamic Enumeration with
less than $O(n^{1/2})$ delay and $O(n^{1/2})$ updates
is **impossible**

Query is q-hierarchical ⇒

Dynamic Enumeration with
$O(1)$ delay and $O(1)$ updates

Ranked Enumeration for non-repeating CQs

Dichotomy:
[Berkholz, Keppeler, Schweikardt 2017]

…on deterministic DB's

Our Result:

…on TI probabilistic DB's

Non-repeating CQs

If the OMv-conjecture holds:

Query is not q-hierarchical ⇒

Dynamic Enumeration with
less than $O(n^{1/2})$ delay and $O(n^{1/2})$ updates
is **impossible**

Query is q-hierarchical ⇒

Dynamic Enumeration with
$O(1)$ delay and $O(1)$ updates

Ranked Enumeration for non-repeating CQs

Non-repeating CQs

Dichotomy:
[Berkholz, Keppeler, Schweikardt 2017]

…on deterministic DB's

If the OMv-conjecture holds:

Query is not q-hierarchical ⇒

Dynamic Enumeration with
less than $O(n^{1/2})$ delay and $O(n^{1/2})$ updates
is **impossible**

Query is q-hierarchical ⇒

Dynamic Enumeration with
$O(1)$ delay and $O(1)$ updates

Our Result:

…on TI probabilistic DB's

If the OMv-conjecture holds:

Query is not q-hierarchical ⇒

Dynamic Enumeration with
less than $O(n^{1/2})$ delay and $O(n^{1/2})$ updates
is **impossible**

Ranked Enumeration for non-repeating CQs

Non-repeating CQs

Dichotomy:
[Berkholz, Keppeler, Schweikardt 2017]

…on deterministic DB's

If the OMv-conjecture holds:

Query is not q-hierarchical ⇒

Dynamic Enumeration with
less than $O(n^{1/2})$ delay and $O(n^{1/2})$ updates
is **impossible**

Query is q-hierarchical ⇒

Dynamic Enumeration with
$O(1)$ delay and $O(1)$ updates

Our Result:

…on TI probabilistic DB's

If the OMv-conjecture holds:

Query is not q-hierarchical ⇒

Dynamic Enumeration with
less than $O(n^{1/2})$ delay and $O(n^{1/2})$ updates
is **impossible**

Query is q-hierarchical ⇒

Dynamic Ranked Enumeration
with $O(\log n)$ delay and $O(\log n)$ updates

✔ Probabilistic Databases ✔ under Updates

✔ Boolean Query Evaluation
➡ and Ranked Enumeration

Christoph Berkholz, Maximilian Merz

Talk at AlMoTh 2022, based on the PODS 2021 paper

Probabilistic Databases under Updates

Boolean Query Evaluation
and Ranked Enumeration

Christoph Berkholz, Maximilian Merz

Talk at AlMoTh 2022, based on the PODS 2021 paper

Probabilistic Databases under Updates

Boolean Query Evaluation
and Ranked Enumeration

Christoph Berkholz, Maximilian Merz

Talk at AlMoTh 2022, based on the PODS 2021 paper