Chase Variants & k-Boundedness

Reunion GRAPHIK 2018

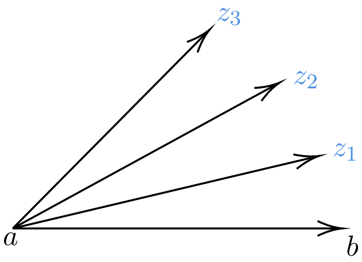
Stathis Delivorias, Marie-Laure Mugnier, Michel Leclère, Federico Ulliana

17 December, 2018

Contents:

Redundancy & Retraction

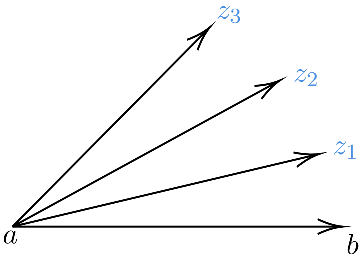
$$\exists z_1 \exists z_2 \exists z_3 \; p(a,b) \wedge p(a,z_1) \wedge p(a,z_2) \wedge p(a,z_3)$$



$$F = \{p(a,b), p(a,z_1), \\ p(a,z_2), p(a,z_3)\}$$

Redundancy & Retraction

$$\exists z_1 \exists z_2 \exists z_3 \; p(a,b) \wedge p(a,z_1) \wedge p(a,z_2) \wedge p(a,z_3)$$



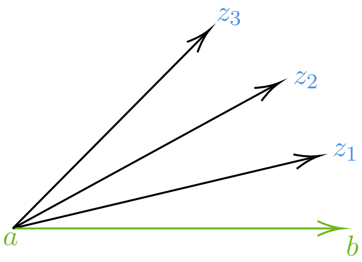
$$F = \{p(a,b), p(a,z_1), p(a,z_2), p(a,z_3)\}$$

$$\boxed{\begin{array}{l} z_1 \mapsto b \\ z_2 \mapsto b \\ z_3 \mapsto b \end{array}}$$

retraction σ

Redundancy & Retraction

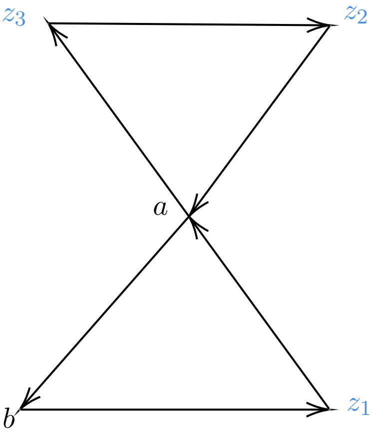
$$\exists z_1 \exists z_2 \exists z_3 \ p(a,b) \land p(a,z_1) \land p(a,z_2) \land p(a,z_3)$$



$$F = \{p(a,b), p(a,z_1), p(a,z_2), p(a,z_3)\}$$

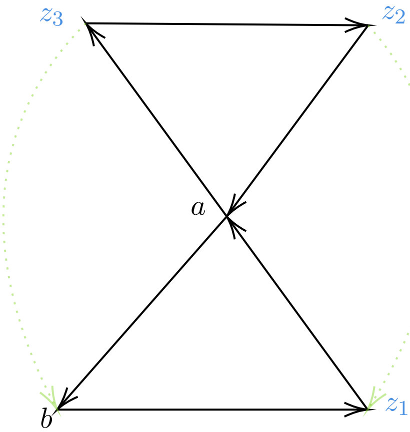
$$F \equiv \{p(a,b)\}$$

Redundancy & Retraction .2



$$F = \{p(a,b), p(b,z_1), p(z_1,a), p(a,z_3), p(z_3,z_2), p(z_2,a)\}$$

Redundancy & Retraction .2



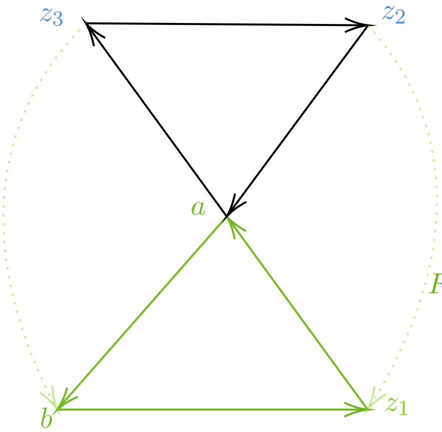
$$F = \{p(a,b), p(b,z_1), p(z_1,a), p(a,z_3), p(z_3,z_2), p(z_2,a)\}$$

$z_2 \mapsto z_1$

$z_3 \mapsto b$

retraction σ

Redundancy & Retraction .2



$$F = \{p(a,b), p(b,z_1), p(z_1,a), p(a,z_3), p(z_3,z_2), p(z_2,a)\}$$

$$F \equiv \{p(a,b), p(b,z_1), p(z_1,a)\}$$

Existential Rules:

$$\forall \bar{x} \forall \bar{y} \big(\mathrm{Body}(\bar{x}, \bar{y}) \to \exists \bar{z}\ \mathrm{Head}(\bar{x}, \bar{z})\big)$$

$\mathrm{Body}(\bar{x}, \bar{y}), \mathrm{Head}(\bar{x}, \bar{z})$: conjunctions of atoms with variables from the sets $\bar{x}, \bar{y}$ and $\bar{x}, \bar{z}$ respectively.

Without loss of clarity we can omit the quantifiers:

$$\mathrm{Body}(\bar{x}, \bar{y}) \to \mathrm{Head}(\bar{x}, \bar{z})$$

for example

$$p(x, y) \wedge q(x, z) \to p(x, w) \wedge s(z, y)$$

Rule Application

for example the rule

$$R = \texttt{mammal}(x) \rightarrow \texttt{motherof}(w, x) \wedge \texttt{mammal}(w)$$

applied to the *factbase*

$$F = \{\texttt{mammal}(Ringo)\}$$

gives

$$F' = \{\texttt{mammal}(Ringo), \texttt{motherof}(w_0, Ringo), \texttt{mammal}(w_0)\}$$

Rule Application

the *trigger*

$$\big(R, \{x \mapsto \texttt{Ringo}\}\big)$$

applied to

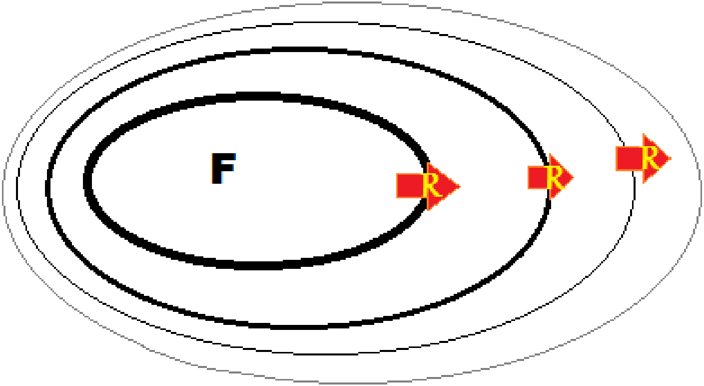
$$F = \ \big\{\texttt{mammal}(Ringo)\big\}$$

gives

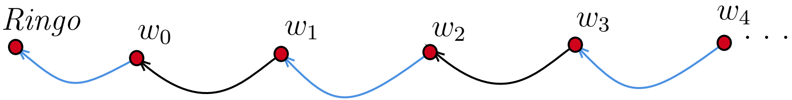
$$F' = \ \big\{\texttt{mammal}(Ringo), \texttt{motherof}(w_0, Ringo), \texttt{mammal}(w_0)\big\}$$

Forward Chaining

is the process of using a set of rules $\mathcal{R}$
to infer information from a factbase F.



The sequential application of rules on an evolving factbase is called a *derivation*. A derivation is not necessarily terminating...

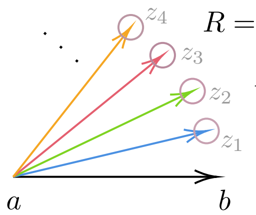


$$\{\texttt{mammal}(Ringo), \texttt{motherof}(w_0, Ringo), \texttt{mammal}(w_0), \texttt{motherof}(w_1, w_0), \texttt{mammal}(w_1), \texttt{motherof}(w_2, w_1), \texttt{mammal}(w_2), \texttt{motherof}(w_3, w_2), \texttt{mammal}(w_3), ..$$

Chase Variants: Classes of Derivations

In existential rules, several *chase variants* handle differently the possible redundancies caused by the introduction of *nulls* to the factbase:

The **oblivious chase** performs all possible rule applications without repeating triggers:



$$R = p(x, y) \to p(x, z) \wedge r(z)$$

$$F = \{p(a, b)\}$$

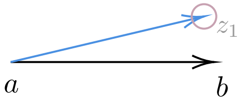
Chase Variants: Classes of Derivations

In existential rules, several *chase variants* handle differently the possible redundancies caused by the introduction of *nulls* to the factbase:

The **semi-oblivious chase** does not repeat triggers which map the *frontier* variables similarly:

$$R = p(x, y) \to p(x, z) \wedge r(z)$$

$$F = \{p(a, b)\}$$



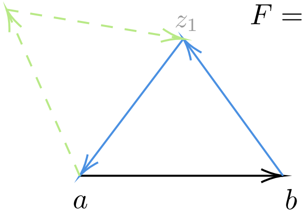
Chase Variants: Classes of Derivations

In existential rules, several *chase variants* handle differently the possible redundancies caused by the introduction of *nulls* to the factbase:

The **restricted chase** does not apply triggers when their output can be retracted back to the existing facbase:

$$R = p(x, y) \rightarrow p(y, z) \wedge p(z, x)$$

$$F = \{p(a, b)\}$$



Chase Variants: Classes of Derivations

In existential rules, several *chase variants* handle differently the possible redundancies caused by the introduction of *nulls* to the factbase:

The **equivalent chase**[1] applies a trigger only if the resulting factbase is not logically equivalent to the existing facbase:

$$R = p(x, y) \rightarrow p(x, x) \wedge p(y, z)$$

$$F = \{p(a, b)\}$$

[1] the equivalent chase is equivalent with the *core chase* in terms of termination

Boundedness

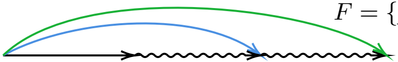
Even if a rule(set) is such that forward chaining always terminates, like

$$p(x, y) \wedge q(y, z) \rightarrow p(x, z)$$

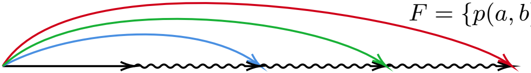
it might not be bounded:

$F = \{p(a, b), q(b, c)\}$

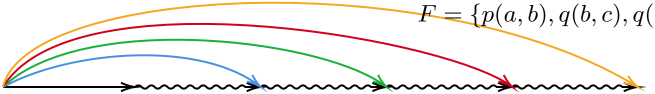
$F = \{p(a, b), q(b, c), q(c, d)\}$



$F = \{p(a, b), q(b, c), q(c, d), q(d, e)\}$

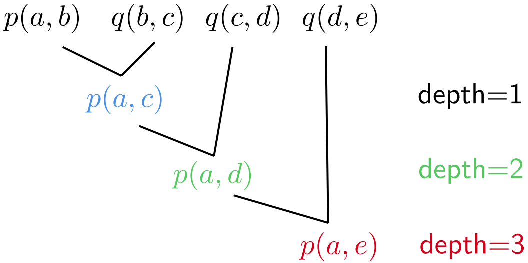


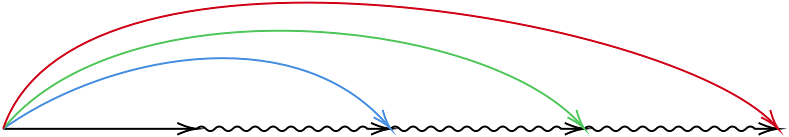
$F = \{p(a, b), q(b, c), q(c, d), q(d, e), q(e, f)\}$



Boundedness

In particular a ruleset is *bounded* if there is a bound to the depth of derivations for every factbase.





Boundedness

- Several semantic properties are ensured when a ruleset is bounded (chase termination, query rewriting termination, non-recursivity).
- There has been extensive research around boundedness in *datalog*, where it has been shown to be undecidable.
- We parametrize boundedness by chase variants, and introduce k-*boundedness*.

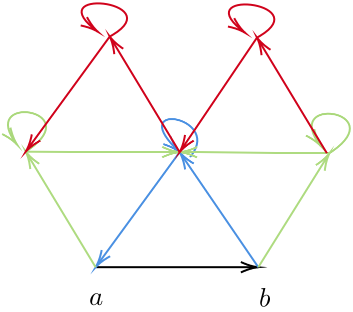
X-k-boundedness

X $\in \{\mathbf{O}, \mathbf{SO}, \mathbf{R}, \mathbf{E}\}$ is used as a parameter for each (X-)chase variant.

A ruleset $\mathcal{R}$ is *X-k-bounded* if for every factbase F, every X-derivation from $(F, \mathcal{R})$ is of depth at most k.

For example

$$\mathcal{R} = \{p(x, y) \to p(y, z) \wedge p(z, z) \wedge p(z, x)\}$$



$\mathcal{R}$ is $\mathbf{R}$-3-bounded

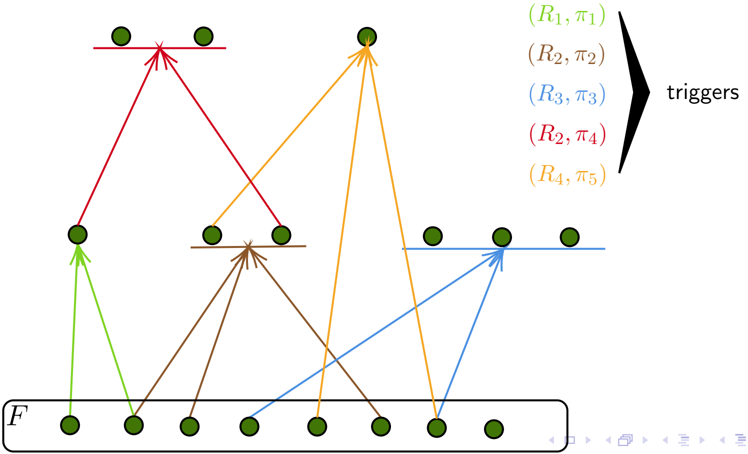
but

$\mathcal{R}$ is $\mathbf{SO}$-unbounded

Chase Graph

Chase derivations can be represented using the concept of a *chase graph*:

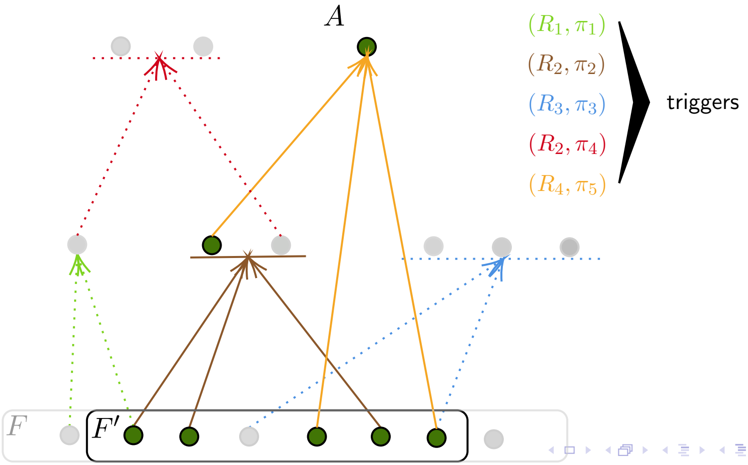
$$\mathcal{R} = \{R_1, R_2, R_3, R_4\}$$



Ancestors

Chase derivations can be represented using the concept of a *chase graph*:

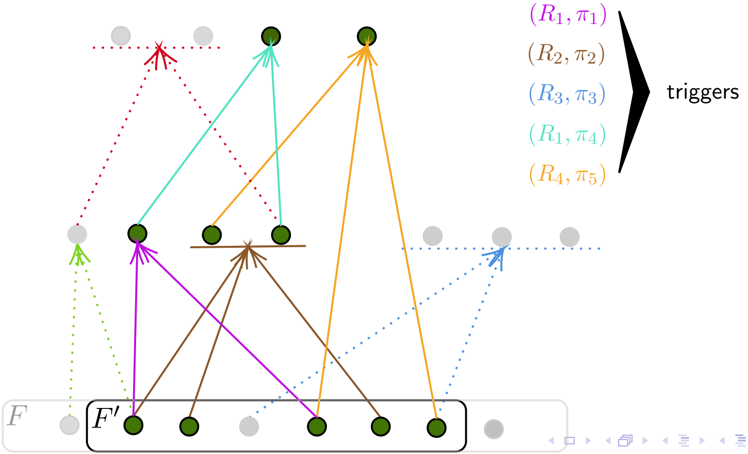
$$\mathcal{R} = \{R_1, R_2, R_3, R_4\}$$



Preservation of Ancestry

The X-chase *preserves ancestry* if for every atom A in an X-derivation, there is an X-derivation from its ancestors (& same ruleset) that produces A at the same rank.

$$\mathcal{R} = \{R_1, R_2, R_3, R_4\}$$



Some Results

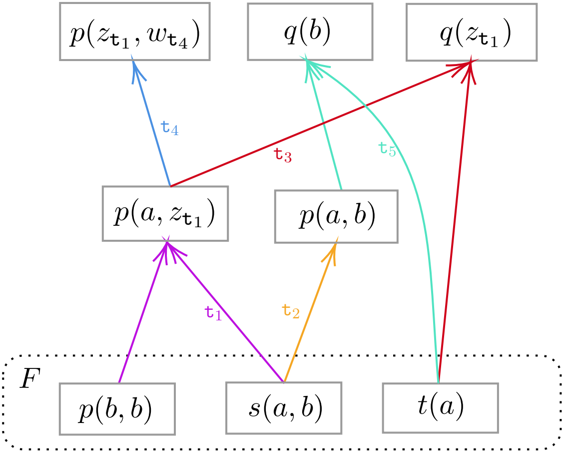
I. Determining if a set of rules is X-k-bounded is decidable if the X-chase preserves ancestry.

II. The X-chase preserves ancestry for X $\in \{\mathbf{O}, \mathbf{SO}, \mathbf{R}\}$.

III. The $\mathbf{E}$-chase does not preserve ancestry.

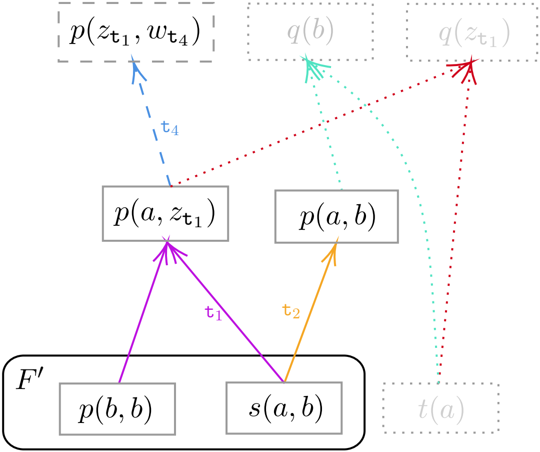
$\mathbb{E}$-chase Does Not Preserve Ancestry

$$R_1 = s(x,y) \wedge p(y,y) \rightarrow \exists z\; p(x,z)$$
$$R_2 = s(x,y) \rightarrow p(x,y)$$
$$R_3 = t(x) \wedge p(x,y) \rightarrow q(y)$$
$$R_4 = p(x,y) \rightarrow \exists w\; p(y,w)$$



$\mathbb{E}$-chase Does Not Preserve Ancestry

$$R_1 = s(x,y) \wedge p(y,y) \to \exists z\ p(x,z)$$
$$R_2 = s(x,y) \to p(x,y)$$
$$R_3 = t(x) \wedge p(x,y) \to q(y)$$
$$R_4 = p(x,y) \to \exists w\ p(y,w)$$



Summary of thesis work

- Extensive study of the notion of chase variant and its potential properties, new framework that encorporates many chase variants.
- Defined two new chase variants, the vacuum chase and the local core chase, that detect more redundancies than the restricted chase.
- Decidability of k-boundedness shown for many chase variants (oblivious, semi-oblivious, restricted, breadth-first oblivious, breadth-first semi-oblivious, breadth-first restricted, parallel, local core chase).
- Open Question: (Un)Decidability of k-boundedness for the equivalent chase, the core chase, the frugal chase and the vacuum chase to be researched.
- Open Question: (Un)Decidability of $\exists$-*boundedness* of restricted chase.

THANK YOU !!