20th Italian Symposium on Advanced Database Systems
SEBD 2012

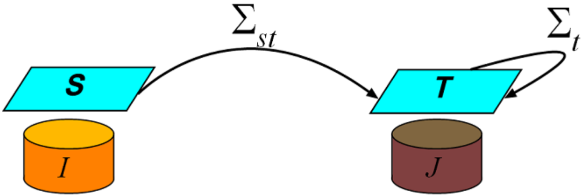
Stratification-based Criteria for Checking Chase Termination

Sergio Greco, Francesca Spezzano, Irina Trubitsyna

DEIS, Università della Calabria - Italy
{greco,fspezzano,irina}@deis.unical.it

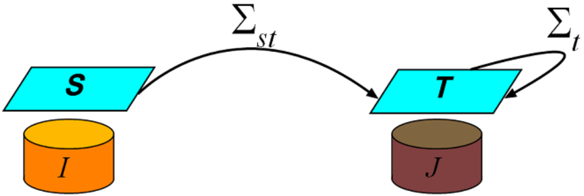
Data Exchange [FaginKMP05]

Schema Mapping $\mathcal{M} = (\mathbf{S}, \mathbf{T}, \Sigma_{st}, \Sigma_t)$



Data Exchange [FaginKMP05]

Schema Mapping $\mathcal{M} = (\mathbf{S}, \mathbf{T}, \Sigma_{st}, \Sigma_t)$



- Given a source instance I, a *solution for I under $\mathcal{M}$* is a target instance J s.t. $(I, J) \models \Sigma_{st}$ and $J \models \Sigma_t$.

Data Exchange - Universal solutions

- A **universal solution** for I under $\mathcal{M}$: is a solution J for I under $\mathcal{M}$ s.t. for every solution J' for I under $\mathcal{M}$, J can be "mapped" in J'.

Data Exchange - Universal solutions

- A **universal solution** for I under $\mathcal{M}$: is a solution J for I under $\mathcal{M}$ s.t. for every solution J' for I under $\mathcal{M}$, J can be "mapped" in J'.

Example

Let $\mathcal{M}$ be specified by the s-t TGD

$$\forall x \forall y (E(x, y) \rightarrow \exists z (F(x, z) \land F(z, y)).$$

If $I = E(1, 2)$, then:

Data Exchange - Universal solutions

- A **universal solution** for I under $\mathcal{M}$: is a solution J for I under $\mathcal{M}$ s.t. for every solution J' for I under $\mathcal{M}$, J can be "mapped" in J'.

Example

Let $\mathcal{M}$ be specified by the s-t TGD

$$\forall x \forall y (E(x, y) \rightarrow \exists z (F(x, z) \wedge F(z, y)).$$

If $I = E(1, 2)$, then:

- $J_1 = F(1, 1), F(1, 2)$ is a solution for I.

Data Exchange - Universal solutions

- A **universal solution** for I under $\mathcal{M}$: is a solution J for I under $\mathcal{M}$ s.t. for every solution J' for I under $\mathcal{M}$, J can be "mapped" in J'.

Example

Let $\mathcal{M}$ be specified by the s-t TGD

$$\forall x \forall y (E(x, y) \rightarrow \exists z (F(x, z) \wedge F(z, y)).$$

If $I = E(1, 2)$, then:

- $J_1 = F(1, 1), F(1, 2)$ is a solution for I.
- $J_2 = F(1, 2), F(2, 2)$ is a solution for I.

Data Exchange - Universal solutions

- A **universal solution** for I under $\mathcal{M}$: is a solution J for I under $\mathcal{M}$ s.t. for every solution J' for I under $\mathcal{M}$, J can be "mapped" in J'.

Example

Let $\mathcal{M}$ be specified by the s-t TGD

$$\forall x \forall y (E(x,y) \rightarrow \exists z (F(x,z) \wedge F(z,y)).$$

If $I = E(1,2)$, then:

- $J_1 = F(1,1), F(1,2)$ is a solution for I.
- $J_2 = F(1,2), F(2,2)$ is a solution for I.
- $J_3 = F(1,\eta), F(\eta,2)$ is a universal solution for I (labeled nulls can be mapped to constants).

Data Exchange - Universal solutions

- A **universal solution** for I under $\mathcal{M}$: is a solution J for I under $\mathcal{M}$ s.t. for every solution J' for I under $\mathcal{M}$, J can be "mapped" in J'.

Example

Let $\mathcal{M}$ be specified by the s-t TGD

$$\forall x \forall y (E(x,y) \rightarrow \exists z (F(x,z) \wedge F(z,y)).$$

If $I = E(1,2)$, then:

- $J_1 = F(1,1), F(1,2)$ is a solution for I.
- $J_2 = F(1,2), F(2,2)$ is a solution for I.
- $J_3 = F(1,\eta), F(\eta,2)$ is a universal solution for I (labeled nulls can be mapped to constants).

- Intuitively, universal solutions are the **most general solutions** in data exchange.

Data Exchange - Universal solutions

Theorem

A universal solution can be computed using the ***chase algorithm****.*

Chase Algorithm

- Add tuples to the database and set the value of nulls to satisfy integrity constraints.

Example

- $D = \{N(a), S(a)\}$

$$\Sigma = \begin{array}{l} \forall x\ [\ N(x) \rightarrow \exists y\ E(x,y)\ ] \\ \forall (x,y)\ [\ S(x) \wedge E(x,y) \rightarrow N(y)\ ] \end{array}$$

Chase Algorithm

- Add tuples to the database and set the value of nulls to satisfy integrity constraints.

Example

- $D = \{N(a), S(a)\}$

$$\Sigma = \begin{array}{l} \forall x \; [\; N(x) \rightarrow \exists y \; E(x,y) \;] \\ \forall (x,y) \; [\; S(x) \wedge E(x,y) \rightarrow N(y) \;] \end{array}$$

- $chase(D, \Sigma) = \{N(a), S(a)\}$

Chase Algorithm

- Add tuples to the database and set the value of nulls to satisfy integrity constraints.

Example

- $D = \{N(a), S(a)\}$

$$\Sigma = \begin{array}{l} \forall x\ [\ N(x) \rightarrow \exists y\ E(x,y)\ ] \\ \forall (x,y)\ [\ S(x) \wedge E(x,y) \rightarrow N(y)\ ] \end{array}$$

- $chase(D,\Sigma) = \{N(a), S(a)\}$

Chase Algorithm

- Add tuples to the database and set the value of nulls to satisfy integrity constraints.

Example

- $D = \{N(a), S(a)\}$

$$\Sigma = \begin{array}{l} \forall x\ [\ N(x) \rightarrow \exists y\ E(x,y)\ ] \\ \forall (x,y)\ [\ S(x) \wedge E(x,y) \rightarrow N(y)\ ] \end{array}$$

- $chase(D, \Sigma) = \{N(a), S(a), E(a, \eta_1)\}$

Chase Algorithm

- Add tuples to the database and set the value of nulls to satisfy integrity constraints.

Example

- $D = \{N(a), S(a)\}$

$$\Sigma = \begin{array}{l} \forall x\ [\ N(x) \rightarrow \exists y\ E(x,y)\ ] \\ \forall (x,y)\ [\ S(x) \wedge E(x,y) \rightarrow N(y)\ ] \end{array}$$

- $chase(D, \Sigma) = \{N(a), S(a), E(a, \eta_1)\}$

Chase Algorithm

- Add tuples to the database and set the value of nulls to satisfy integrity constraints.

Example

- $D = \{N(a), S(a)\}$

$$\Sigma = \begin{array}{l} \forall x \; [\; N(x) \rightarrow \exists y \; E(x, y) \;] \\ \forall (x, y) \; [\; S(x) \wedge E(x, y) \rightarrow N(y) \;] \end{array}$$

- $chase(D, \Sigma) = \{N(a), S(a), E(a, \eta_1), N(\eta_1)\}$

Chase Algorithm

- Add tuples to the database and set the value of nulls to satisfy integrity constraints.

Example

- $D = \{N(a), S(a)\}$

$$\Sigma = \begin{array}{l} \forall x\ [\ N(x) \rightarrow \exists y\ E(x,y)\ ] \\ \forall (x,y)\ [\ S(x) \wedge E(x,y) \rightarrow N(y)\ ] \end{array}$$

- $chase(D, \Sigma) = \{N(a), S(a), E(a, \eta_1), N(\eta_1)\}$

Chase Algorithm

- Add tuples to the database and set the value of nulls to satisfy integrity constraints.

Example

- $D = \{N(a), S(a)\}$

$$\Sigma = \begin{array}{l} \forall x \; [\; N(x) \rightarrow \exists y \; E(x,y) \;] \\ \forall (x,y) \; [\; S(x) \wedge E(x,y) \rightarrow N(y) \;] \end{array}$$

- $chase(D, \Sigma) = \{N(a), S(a), E(a, \eta_1), N(\eta_1), E(\eta_1, \eta_2)\}$

Chase Algorithm

- Add tuples to the database and set the value of nulls to satisfy integrity constraints.

Example

- $D = \{N(a), S(a)\}$

$$\Sigma = \begin{array}{l} \forall x\ [\ N(x) \rightarrow \exists y\ E(x,y)\ ] \\ \forall (x,y)\ [\ S(x) \wedge E(x,y) \rightarrow N(y)\ ] \end{array}$$

- $chase(D, \Sigma) = \{N(a), S(a), E(a, \eta_1), N(\eta_1), E(\eta_1, \eta_2)\}$

- Other applications: query containment & optimization, data integration, consistent query answering,. . .

A Non-terminating Chase Sequence

Example

$$\Sigma = \begin{array}{l} \forall x \; [\; N(x) \rightarrow \exists y \; E(x,y) \;] \\ \forall (x,y) \; [\; S(y), E(x,y) \rightarrow N(y) \;] \end{array}$$

$D = \{N(a)\}$

$chase(D, \Sigma) = \{N(a), E(a, \eta_1)\}$

A Non-terminating Chase Sequence

Example

$$\Sigma = \begin{array}{l} \forall x\ [\ N(x) \rightarrow \exists y\ E(x,y)\ ] \\ \forall (x,y)\ [\ S(y), E(x,y) \rightarrow N(y)\ ] \end{array}$$

$D = \{N(a)\}$

$chase(D, \Sigma) = \{N(a), E(a, \eta_1), N(\eta_1)\}$

A Non-terminating Chase Sequence

Example

$$\Sigma = \begin{array}{l} \forall x\ [\ N(x) \rightarrow \exists y\ E(x,y)\ ] \\ \forall (x,y)\ [\ S(y), E(x,y) \rightarrow N(y)\ ] \end{array}$$

$D = \{N(a)\}$

$chase(D,\Sigma) = \{N(a), E(a,\eta_1), N(\eta_1), E(\eta_1,\eta_2)\}$

A Non-terminating Chase Sequence

Example

$$\Sigma = \begin{array}{l} \forall x\ [\ N(x) \rightarrow \exists y\ E(x,y)\ ] \\ \forall (x,y)\ [\ S(y), E(x,y) \rightarrow N(y)\ ] \end{array}$$

$D = \{N(a)\}$

$chase(D,\Sigma) = \{N(a), E(a,\eta_1), N(\eta_1), E(\eta_1,\eta_2), N(\eta_2), \ldots\}$

A Non-terminating Chase Sequence

Example

$$\Sigma = \begin{array}{l} \forall x\ [\ N(x) \rightarrow \exists y\ E(x,y)\ ] \\ \forall (x,y)\ [\ S(y), E(x,y) \rightarrow N(y)\ ] \end{array}$$

$D = \{N(a)\}$

$chase(D,\Sigma) = \{N(a), E(a,\eta_1), N(\eta_1), E(\eta_1,\eta_2), N(\eta_2), \ldots\}$

- Chase termination: undecidable problem (Deutsch et al. [2008])

A Non-terminating Chase Sequence

Example

$$\Sigma = \begin{array}{l} \forall x\ [\ N(x) \rightarrow \exists y\ E(x,y)\ ] \\ \forall (x,y)\ [\ S(y), E(x,y) \rightarrow N(y)\ ] \end{array}$$

$D = \{N(a)\}$

$chase(D,\Sigma) = \{N(a), E(a,\eta_1), N(\eta_1), E(\eta_1,\eta_2), N(\eta_2), \ldots\}$

- Chase termination: undecidable problem (Deutsch et al. [2008])
- Individuate sufficient chase termination conditions

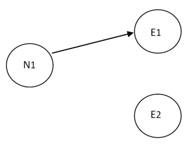
Weak Acyclicity [FaginKMP05]

(Dependency Graph)

- represent the flow of values
- nodes are positions
- special edges $\stackrel{*}{\rightarrow}$ indicates the generation of new null values

Example

$$\Sigma = \begin{array}{l} \forall x\ [\ N(x) \rightarrow \exists y\ E(x,y)\ ] \\ \forall (x,y)\ [\ E(x,y) \rightarrow N(y)\ ] \end{array}$$



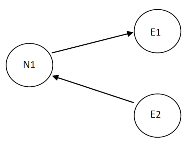
Weak Acyclicity [FaginKMP05]

(Dependency Graph)

- represent the flow of values
- nodes are positions
- special edges $\xrightarrow{*}$ indicates the generation of new null values

Example

$$\Sigma = \begin{array}{l} \forall x\ [\ N(x) \rightarrow \exists y\ E(x,y)\ ] \\ \forall (x,y)\ [\ E(x,y) \rightarrow N(y)\ ] \end{array}$$



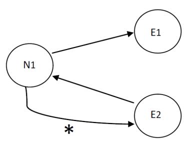
Weak Acyclicity [FaginKMP05]

(Dependency Graph)

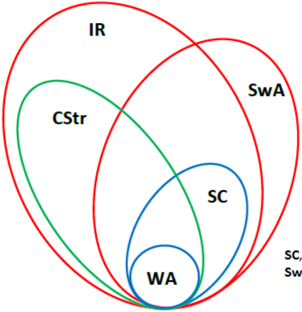
- represent the flow of values
- nodes are positions
- special edges $\xrightarrow{*}$ indicates the generation of new null values

Example

$$\Sigma = \begin{array}{l} \forall x\ [\ N(x) \rightarrow \exists y\ E(x,y)\ ] \\ \forall (x,y)\ [\ E(x,y) \rightarrow N(y)\ ] \end{array}$$



$\mathcal{WA}$, $\mathcal{SC}$, $\mathcal{CS}tr$, $\mathcal{IR}$ and $\mathcal{S}w\mathcal{A}$ Relationships



SC,CStr,IR [MeierSL09]
SwA [Marnette09]

Contributions

- Introduction of new chase termination criteria
  - *WA*-stratification (*SC*-stratification and *SwA*-stratification)
  - Local Stratification

Chase Termination Criteria

Stratification [DeutschNR08]

1. Partition $\Sigma = \Sigma_1 \cup \cdots \cup \Sigma_n$ according to how constraints fire each other;

Stratification [DeutschNR08]

1. Partition $\Sigma = \Sigma_1 \cup \cdots \cup \Sigma_n$ according to how constraints fire each other;
2. verify if each subset $\Sigma_1 \cup \cdots \cup \Sigma_n$ is weakly-acyclic.

Stratification [DeutschNR08]

1. Partition $\Sigma = \Sigma_1 \cup \cdots \cup \Sigma_n$ according to how constraints fire each other;
2. verify if each subset $\Sigma_1 \cup \cdots \cup \Sigma_n$ is weakly-acyclic.

- Build the *chase graph*

Stratification [DeutschNR08]

1. Partition $\Sigma = \Sigma_1 \cup \cdots \cup \Sigma_n$ according to how constraints fire each other;
2. verify if each subset $\Sigma_1 \cup \cdots \cup \Sigma_n$ is weakly-acyclic.

- *Precedence relation*

  $r_1 \prec r_2$ if $\exists K_1$ s.t.

  1) $K_1 \models r_2 \wedge K_1 \xrightarrow{r_1} K_2$,

  2) $K_2 \not\models r_2$.

- Build the *chase graph*

Stratification [DeutschNR08]

1. Partition $\Sigma = \Sigma_1 \cup \cdots \cup \Sigma_n$ according to how constraints fire each other;
2. verify if each subset $\Sigma_1 \cup \cdots \cup \Sigma_n$ is weakly-acyclic.

- *Precedence relation*

  $r_1 \prec r_2$ if $\exists K_1$ s.t.

  1) $K_1 \models r_2 \wedge K_1 \xrightarrow{r_1} K_2$,

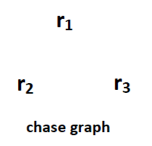
  2) $K_2 \not\models r_2$.

- Build the *chase graph*
- Partition in strongly connected components

Example: Stratification

Example

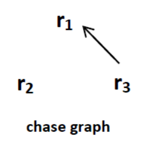
$$\Sigma = \begin{array}{ll} r_1: & R(x) \rightarrow \exists y\, T(x, y) \\ r_2: & R(x) \rightarrow T(x, x) \\ r_3: & T(x, y) \wedge T(x, x) \rightarrow R(y) \end{array}$$



Example: Stratification

Example

$$\Sigma = \begin{array}{l} r_1 : \; R(x) \rightarrow \exists y \, T(x,y) \\ r_2 : \; R(x) \rightarrow T(x,x) \\ r_3 : \; T(x,y) \wedge T(x,x) \rightarrow R(y) \end{array}$$

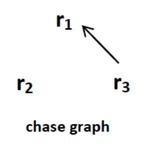


- $r_3 \prec r_1$ because if $K_1 = \{T(a,b), T(a,a)\}$ then

Example: Stratification

Example

$$\Sigma = \begin{array}{ll} r_1: & R(x) \rightarrow \exists y\, T(x,y) \\ r_2: & R(x) \rightarrow T(x,x) \\ r_3: & T(x,y) \wedge T(x,x) \rightarrow R(y) \end{array}$$

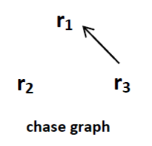


- $r_3 \prec r_1$ because if $K_1 = \{T(a,b), T(a,a)\}$ then
  - $K_1 \models r_1$ and $K_1 \xrightarrow{r_3} K_2 = \{T(a,b), T(a,a), R(b)\}$,

Example: Stratification

Example

$$\Sigma = \begin{array}{l} r_1 : \ R(x) \rightarrow \exists y\, T(x,y) \\ r_2 : \ R(x) \rightarrow T(x,x) \\ r_3 : \ T(x,y) \wedge T(x,x) \rightarrow R(y) \end{array}$$

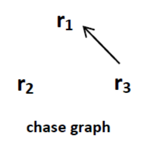


- $r_3 \prec r_1$ because if $K_1 = \{T(a,b), T(a,a)\}$ then
  - $K_1 \models r_1$ and $K_1 \xrightarrow{r_3} K_2 = \{T(a,b), T(a,a), R(b)\}$,
  - $K_2 \not\models r_1$.

Example: Stratification

Example

$$\Sigma = \begin{array}{l} r_1 : \; R(x) \rightarrow \exists y\, T(x,y) \\ r_2 : \; R(x) \rightarrow T(x,x) \\ r_3 : \; T(x,y) \wedge T(x,x) \rightarrow R(y) \end{array}$$

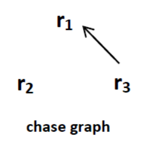


- $r_1 \not\prec r_3$ because

Example: Stratification

Example

$$\Sigma = \begin{array}{l} r_1 : \; R(x) \rightarrow \exists y \, T(x,y) \\ r_2 : \; R(x) \rightarrow T(x,x) \\ r_3 : \; T(x,y) \wedge T(x,x) \rightarrow R(y) \end{array}$$

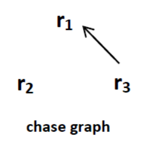


- $r_1 \not\prec r_3$ because
  - if $K_1 = \{R(a)\}$ then $K_2 = \{R(a), T(a, \eta_1)\} \models r_3$ $(T(a,a) \notin K_2)$

Example: Stratification

Example

$$\Sigma = \begin{array}{ll} r_1: & R(x) \rightarrow \exists y\, T(x,y) \\ r_2: & R(x) \rightarrow T(x,x) \\ r_3: & T(x,y) \wedge T(x,x) \rightarrow R(y) \end{array}$$

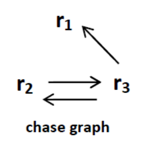


- $r_1 \not\prec r_3$ because
  - if $K_1 = \{R(a)\}$ then $K_2 = \{R(a), T(a,\eta_1)\} \models r_3$ $(T(a,a) \notin K_2)$
  - if $K_1 = \{R(a), T(a,a)\}$ then $K_1 \models \Sigma$ (K_1 universal solution)

Example: Stratification

Example

$$\Sigma = \begin{array}{ll} r_1: & R(x) \rightarrow \exists y\, T(x,y) \\ r_2: & R(x) \rightarrow T(x,x) \\ r_3: & T(x,y) \wedge T(x,x) \rightarrow R(y) \end{array}$$

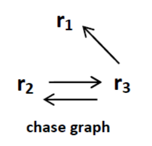


Σ is stratified

Example: Stratification

Example

$$\Sigma = \begin{array}{l} r_1 : \; R(x) \rightarrow \exists y \, T(x,y) \\ r_2 : \; R(x) \rightarrow T(x,x) \\ r_3 : \; T(x,y) \wedge T(x,x) \rightarrow R(y) \end{array}$$



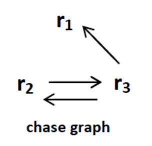
Σ is stratified

- chase sequence $r_1 \rightarrow r_2 \rightarrow r_3 \rightarrow \ldots$ is not terminating

Example: Stratification

Example

$$\Sigma = \begin{array}{l} r_1 : \ R(x) \rightarrow \exists y \ T(x,y) \\ r_2 : \ R(x) \rightarrow T(x,x) \\ r_3 : \ T(x,y) \wedge T(x,x) \rightarrow R(y) \end{array}$$



Σ is stratified

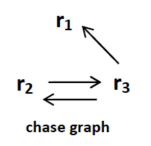
- chase sequence $r_1 \rightarrow r_2 \rightarrow r_3 \rightarrow \ldots$ is not terminating

  $D = \{R(a)\}$

Example: Stratification

Example

$$\Sigma = \begin{array}{l} r_1 : \ R(x) \rightarrow \exists y\, T(x,y) \\ r_2 : \ R(x) \rightarrow T(x,x) \\ r_3 : \ T(x,y) \wedge T(x,x) \rightarrow R(y) \end{array}$$



Σ is stratified

- chase sequence $r_1 \rightarrow r_2 \rightarrow r_3 \rightarrow \ldots$ is not terminating

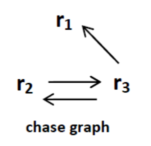
  $D = \{R(a)\}$

  $D \xrightarrow{r_1,h_1} D_1 = \{R(a), T(a,\eta_1)\}$

Example: Stratification

Example

$$\Sigma = \begin{array}{l} r_1 : \; R(x) \rightarrow \exists y\, T(x,y) \\ r_2 : \; R(x) \rightarrow T(x,x) \\ r_3 : \; T(x,y) \wedge T(x,x) \rightarrow R(y) \end{array}$$



Σ is stratified

- chase sequence $r_1 \rightarrow r_2 \rightarrow r_3 \rightarrow \ldots$ is not terminating

  $D = \{R(a)\}$

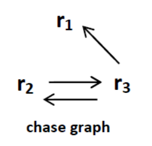
  $D \xrightarrow{r_1, h_1} D_1 = \{R(a), T(a, \eta_1)\}$

  $D_1 \xrightarrow{r_2, h_2} D_2 = \{R(a), T(a, \eta_1), T(a, a)\}$

Example: Stratification

Example

$$\Sigma = \begin{array}{l} r_1 : \; R(x) \rightarrow \exists y \, T(x,y) \\ r_2 : \; R(x) \rightarrow T(x,x) \\ r_3 : \; T(x,y) \wedge T(x,x) \rightarrow R(y) \end{array}$$



Σ is stratified

- chase sequence $r_1 \rightarrow r_2 \rightarrow r_3 \rightarrow \ldots$ is not terminating

  $D = \{R(a)\}$

  $D \xrightarrow{r_1,h_1} D_1 = \{R(a), T(a,\eta_1)\}$

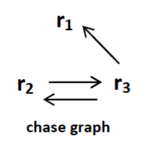
  $D_1 \xrightarrow{r_2,h_2} D_2 = \{R(a), T(a,\eta_1), T(a,a)\}$

  $D_2 \xrightarrow{r_3,h_3} D_3 = \{R(a), T(a,\eta_1), T(a,a), R(\eta_1)\}$

Example: Stratification

Example

$$\Sigma = \begin{array}{l} r_1 : \; R(x) \rightarrow \exists y \, T(x,y) \\ r_2 : \; R(x) \rightarrow T(x,x) \\ r_3 : \; T(x,y) \wedge T(x,x) \rightarrow R(y) \end{array}$$



Σ is stratified

- chase sequence $r_1 \rightarrow r_2 \rightarrow r_3 \rightarrow \ldots$ is not terminating

  $D = \{R(a)\}$

  $D \xrightarrow{r_1, h_1} D_1 = \{R(a), T(a, \eta_1)\}$

  $D_1 \xrightarrow{r_2, h_2} D_2 = \{R(a), T(a, \eta_1), T(a, a)\}$

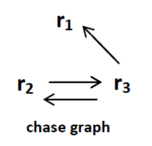
  $D_2 \xrightarrow{r_3, h_3} D_3 = \{R(a), T(a, \eta_1), T(a, a), R(\eta_1)\}$

  $D_3 \xrightarrow{r_1, h_4} D_4 = \{R(a), T(a, \eta_1), T(a, a), R(\eta_1), T(\eta_1, \eta_2)\} \ldots$

Example: Stratification

Example

$$\Sigma = \begin{array}{l} r_1 : \; R(x) \rightarrow \exists y \, T(x,y) \\ r_2 : \; R(x) \rightarrow T(x,x) \\ r_3 : \; T(x,y) \wedge T(x,x) \rightarrow R(y) \end{array}$$



Σ is stratified

- chase sequence $r_1 \rightarrow r_2 \rightarrow r_3 \rightarrow \ldots$ is not terminating

  $D = \{R(a)\}$

  $D \xrightarrow{r_1, h_1} D_1 = \{R(a), T(a, \eta_1)\}$

  $D_1 \xrightarrow{r_2, h_2} D_2 = \{R(a), T(a, \eta_1), T(a, a)\}$

  $D_2 \xrightarrow{r_3, h_3} D_3 = \{R(a), T(a, \eta_1), T(a, a), R(\eta_1)\}$

  $D_3 \xrightarrow{r_1, h_4} D_4 = \{R(a), T(a, \eta_1), T(a, a), R(\eta_1), T(\eta_1, \eta_2)\} \ldots$

- chase sequence $r_2 \rightarrow r_3 \rightarrow r_1$ terminates

Stratification

- Variant of stratification: C-Stratification

Stratification

- Variant of stratification: C-Stratification
- Limited because uses a weaker form of chase

WA-stratification [Greco,Spezzano,Trubitsyna VLDB2011]

- Builds a *firing graph* $\Gamma(\Sigma) = (\Sigma, E)$ representing how constraints fire each other.
- $(r_1, r_2) \in E$ if $r_1 < r_2$ (firing r_1 can cause r_2 to fire)
- $r_1 < r_2$ if:

  1) $K_1 \xrightarrow{r_1} K_2$,

  2) $K_1 \cup S \models r_2$,

  3) $K_2 \cup S \not\models r_2$ and

  4) $Null(S) \cap (Null(K_2) - Null(K_1)) = \emptyset$.

- $r_1 \prec r_2$ if:

  1) $K_1 \xrightarrow{r_1} K_2$,

  2) $K_1 \models r_2$,

  3) $K_2 \not\models r_2$.

WA-stratification [Greco,Spezzano,Trubitsyna VLDB2011]

- Builds a *firing graph* $\Gamma(\Sigma) = (\Sigma, E)$ representing how constraints fire each other.
- $(r_1, r_2) \in E$ if $r_1 < r_2$ (firing r_1 can cause r_2 to fire)
- $r_1 < r_2$ if:

  1) $K_1 \xrightarrow{r_1} K_2$,

  2) $K_1 \cup S \models r_2$,

  3) $K_2 \cup S \not\models r_2$ and

  4) $Null(S) \cap (Null(K_2) - Null(K_1)) = \emptyset$.

- $r_1 \prec r_2$ if:

  1) $K_1 \overset{r_1}{\nrightarrow} K_2$,

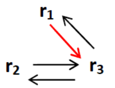
  2) $K_1 \models r_2$,

  3) $K_2 \not\models r_2$.

Example

$$\Sigma = \begin{array}{ll} r_1: & R(x) \rightarrow \exists y\, T(x,y) \\ r_2: & R(x) \rightarrow T(x,x) \\ r_3: & T(x,y) \wedge T(x,x) \rightarrow R(y) \end{array}$$

- $K_1 = \{R(a)\}$ and $K_2 = \{R(a), T(a, \eta_1)\}$
- $S = \{T(a,a)\}$
- $r_3: \; T(a, \eta_1) \wedge T(a,a) \rightarrow R(\eta_1)$
- r_3 is fired by r_1, then we have $r_1 < r_3$



firing graph

WA-stratification [Greco,Spezzano,Trubitsyna VLDB2011]

(WA-Stratification)

We say that Σ is *WA-stratified* (*WAStr*) iff the constraints in every nontrivial strongly connected component of the *firing graph* $\Gamma(\Sigma) = (\Sigma, \{(r_1, r_2) | r_1 < r_2\})$ are weakly acyclic.

WA-stratification [Greco,Spezzano,Trubitsyna VLDB2011]

(WA-Stratification)

We say that Σ is *WA-stratified* (*WAStr*) iff the constraints in every nontrivial strongly connected component of the *firing graph* $\Gamma(\Sigma) = (\Sigma, \{(r_1, r_2) | r_1 < r_2\})$ are weakly acyclic.

Proposition

$\mathcal{C}\mathcal{S}tr \subsetneq \mathcal{W}\mathcal{A}\mathcal{S}tr$ and $\mathcal{S}\mathcal{C} \nparallel \mathcal{W}\mathcal{A}\mathcal{S}tr$.

WA-stratification [Greco,Spezzano,Trubitsyna VLDB2011]

(WA-Stratification)

We say that Σ is *WA-stratified* (*WAStr*) iff the constraints in every nontrivial strongly connected component of the *firing graph* $\Gamma(\Sigma) = (\Sigma, \{(r_1, r_2) | r_1 < r_2\})$ are weakly acyclic.

Proposition

$\mathcal{CS}tr \subsetneq \mathcal{WAS}tr$ and $\mathcal{SC} \nparallel \mathcal{WAS}tr$.

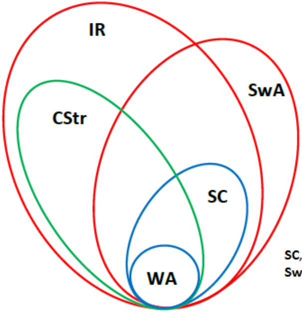
Example

$$\Sigma = \{r : T(x, y) \wedge T(y, x) \rightarrow \exists z\ T(y, z) \wedge T(z, x)\}$$

is *WA*-stratified, but not c-stratified.

SC-stratification and *SwA*-stratification

[Greco,Spezzano,Trubitsyna VLDB2011]



SC,CStr,IR [MeierSL09]
SwA [Marnette09]

SC-stratification and *SwA*-stratification

[Greco,Spezzano,Trubitsyna VLDB2011]

Definition

Given a set of TDGs Σ, we say that Σ is

- *SC-stratified* (*SCStr*) if the constraints in every nontrivial strongly connected component of the firing graph $\Gamma(\Sigma)$ are safe.
- *SwA-stratified* (*SwAStr*) if the constraints in every nontrivial strongly connected component of the firing graph $\Gamma(\Sigma)$ are super-weak acyclic.

SC-stratification and *SwA*-stratification

[Greco,Spezzano,Trubitsyna VLDB2011]

Definition

Given a set of TDGs Σ, we say that Σ is

- *SC-stratified* (*SCStr*) if the constraints in every nontrivial strongly connected component of the firing graph $\Gamma(\Sigma)$ are safe.
- *SwA-stratified* (*SwAStr*) if the constraints in every nontrivial strongly connected component of the firing graph $\Gamma(\Sigma)$ are super-weak acyclic.

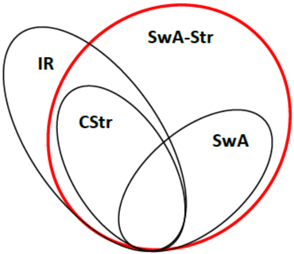
Proposition

Let Σ be a set of WA-stratified (resp. SC-stratified, SwA-stratified) TGDs and let D be a database. Then, the length of every chase sequence of Σ over D is polynomial in the size of D.

SwA-stratification [Greco,Spezzano,Trubitsyna VLDB2011]

Theorem

1. $\mathcal{WAS}tr \subsetneq \mathcal{SCS}tr \subsetneq \mathcal{S}w\mathcal{AS}tr$,
2. for $C \in \{WA, SC, SwA\}$, $\mathcal{C} \subsetneq \mathcal{C}\text{-}\mathcal{S}tr$ and
3. $\mathcal{IR} \nparallel \mathcal{S}w\mathcal{AS}tr$.



Local Stratification [Greco,Spezzano,Trubitsyna VLDB2011]

New chase termination condition: **Local Stratification (LS)**

Local Stratification [Greco,Spezzano,Trubitsyna VLDB2011]

New chase termination condition: **Local Stratification (LS)**

- Combine $<$-relation of $\mathcal{WA}$-$\mathcal{S}tr$, with *SwA* criterion,

Local Stratification [Greco,Spezzano,Trubitsyna VLDB2011]

New chase termination condition: **Local Stratification (LS)**

- Combine $<$-relation of $\mathcal{WA}$-$\mathcal{S}tr$, with *SwA* criterion,
- ensures the termination of all chase sequences in polynomial data complexity,

Local Stratification [Greco,Spezzano,Trubitsyna VLDB2011]

New chase termination condition: **Local Stratification (LS)**

- Combine $<$-relation of $\mathcal{WA}$-$\mathcal{S}tr$, with *SwA* criterion,
- ensures the termination of all chase sequences in polynomial data complexity,
- $\mathcal{S}w\mathcal{A}$-$\mathcal{S}tr \subsetneq \mathcal{LS}$ and $\mathcal{IR} \subsetneq \mathcal{LS}$.

Example: Super-weak Acyclicity [Marnette09]

Analyzes nulls propagation through constraints

Example

$$\Sigma = \begin{array}{ll} r_1: & N(x) \rightarrow \exists y\, \exists z\, E(x,y) \wedge S(z,y) \\ r_2: & E(x_1,y_1) \wedge S(x_1,y_1) \rightarrow N(y_1) \\ r_3: & E(x_2,y_2) \rightarrow E(y_2,x_2) \end{array}$$

Example: Super-weak Acyclicity [Marnette09]

Analyzes nulls propagation through constraints

Example

$$\Sigma = \begin{array}{ll} r_1: & N(x) \rightarrow \exists y\, \exists z\, E(x,y) \wedge S(z,y) \\ r_2: & E(x_1,y_1) \wedge S(x_1,y_1) \rightarrow N(y_1) \\ r_3: & E(x_2,y_2) \rightarrow E(y_2,x_2) \end{array}$$

$$\Sigma = \begin{array}{ll} r_1: & N(\overset{p_1}{x}) \rightarrow \exists y\, \exists z\, E(\overset{p_2}{x}, \overset{p_3}{y}) \wedge S(\overset{p_4}{z}, \overset{p_5}{y}) \\ r_2: & E(\overset{p_6}{x_1}, \overset{p_7}{y_1}) \wedge S(\overset{p_8}{x_1}, \overset{p_9}{y_1}) \rightarrow N(\overset{p_{10}}{y_1}) \\ r_3: & E(\overset{p_{11}}{x_2}, \overset{p_{12}}{y_2}) \rightarrow E(\overset{p_{13}}{y_2}, \overset{p_{14}}{x_2}) \end{array}$$

Example: Super-weak Acyclicity [Marnette09]

Analyzes nulls propagation through constraints

Example

$$\Sigma = \begin{array}{l} r_1 : N(\overset{p_1}{x}) \rightarrow \exists y\, \exists z\, E(\overset{p_2}{x}, \overset{p_3}{y}) \wedge S(\overset{p_4}{z}, \overset{p_5}{y}) \\ r_2 : E(\overset{p_6}{x_1}, \overset{p_7}{y_1}) \wedge S(\overset{p_8}{x_1}, \overset{p_9}{y_1}) \rightarrow N(\overset{p_{10}}{y_1}) \\ r_3 : E(\overset{p_{11}}{x_2}, \overset{p_{12}}{y_2}) \rightarrow E(\overset{p_{13}}{y_2}, \overset{p_{14}}{x_2}) \end{array}$$

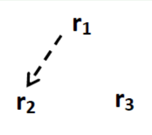
$$Move(r_1, y) = \{p_3, p_5\}$$

Example: Super-weak Acyclicity [Marnette09]

Analyzes nulls propagation through constraints

Example

$$\Sigma = \begin{array}{l} r_1 : \ N(\overset{p_1}{x}) \rightarrow \exists y\, \exists z\, E(\overset{p_2}{x}, \overset{p_3}{y}) \wedge S(\overset{p_4}{z}, \overset{p_5}{y}) \\ r_2 : \ E(\overset{p_6}{x_1}, \overset{p_7}{y_1}) \wedge S(\overset{p_8}{x_1}, \overset{p_9}{y_1}) \rightarrow N(\overset{p_{10}}{y_1}) \\ r_3 : \ E(\overset{p_{11}}{x_2}, \overset{p_{12}}{y_2}) \rightarrow E(\overset{p_{13}}{y_2}, \overset{p_{14}}{x_2}) \end{array}$$



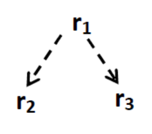
$$Move(r_1, y) = \{p_3, p_5, p_{10}\}$$

Example: Super-weak Acyclicity [Marnette09]

Analyzes nulls propagation through constraints

Example

$$\Sigma = \begin{array}{ll} r_1: & N(\overset{p_1}{x}) \rightarrow \exists y \, \exists z \; E(\overset{p_2}{x}, \overset{p_3}{y}) \wedge S(\overset{p_4}{z}, \overset{p_5}{y}) \\ r_2: & E(\overset{p_6}{x_1}, \overset{p_7}{y_1}) \wedge S(\overset{p_8}{x_1}, \overset{p_9}{y_1}) \rightarrow N(\overset{p_{10}}{y_1}) \\ r_3: & E(\overset{p_{11}}{x_2}, \overset{p_{12}}{y_2}) \rightarrow E(\overset{p_{13}}{y_2}, \overset{p_{14}}{x_2}) \end{array}$$



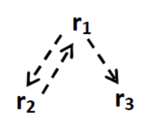
$$Move(r_1, y) = \{p_3, p_5, p_{10}, p_{13}\}$$

Example: Super-weak Acyclicity [Marnette09]

Analyzes nulls propagation through constraints

Example

$$\Sigma = \begin{array}{l} r_1 : \; N(\overset{p_1}{\color{red}{x}}) \rightarrow \exists y \, \exists z \; E(\overset{p_2}{\color{red}{x}}, \overset{p_3}{y}) \wedge S(\overset{p_4}{z}, \overset{p_5}{y}) \\ r_2 : \; E(\overset{p_6}{x_1}, \overset{p_7}{y_1}) \wedge S(\overset{p_8}{x_1}, \overset{p_9}{y_1}) \rightarrow N(\overset{p_{10}}{y_1}) \\ r_3 : \; E(\overset{p_{11}}{x_2}, \overset{p_{12}}{y_2}) \rightarrow E(\overset{p_{13}}{y_2}, \overset{p_{14}}{x_2}) \end{array}$$



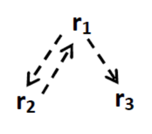
$$Move(r_1, y) = \{p_3, p_5, p_{10}, p_{13}, {\color{red}{p_2}}\}$$

Example: Super-weak Acyclicity [Marnette09]

Analyzes nulls propagation through constraints

Example

$$\Sigma = \begin{array}{l} r_1 : N(\overset{p_1}{x}) \rightarrow \exists y\, \exists z\, E(\overset{p_2}{x}, \overset{p_3}{y}) \wedge S(\overset{p_4}{z}, \overset{p_5}{y}) \\ r_2 : E(\overset{p_6}{x_1}, \overset{p_7}{y_1}) \wedge S(\overset{p_8}{x_1}, \overset{p_9}{y_1}) \rightarrow N(\overset{p_{10}}{y_1}) \\ r_3 : E(\overset{p_{11}}{x_2}, \overset{p_{12}}{y_2}) \rightarrow E(\overset{p_{13}}{y_2}, \overset{p_{14}}{x_2}) \end{array}$$



$$Move(r_1, y)) = \{p_3, p_5, p_{10}, p_{13}, p_2, p_{14}\}$$

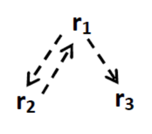
Example: Super-weak Acyclicity [Marnette09]

Analyzes nulls propagation through constraints

Example

$$\Sigma = \begin{array}{l} r_1 : N(\overset{p_1}{x}) \rightarrow \exists y \, \exists z \, E(\overset{p_2}{x}, \overset{p_3}{y}) \wedge S(\overset{p_4}{z}, \overset{p_5}{y}) \\ r_2 : E(\overset{p_6}{x_1}, \overset{p_7}{y_1}) \wedge S(\overset{p_8}{x_1}, \overset{p_9}{y_1}) \rightarrow N(\overset{p_{10}}{y_1}) \\ r_3 : E(\overset{p_{11}}{x_2}, \overset{p_{12}}{y_2}) \rightarrow E(\overset{p_{13}}{y_2}, \overset{p_{14}}{x_2}) \end{array}$$

$$Move(r_1, y) = \{p_3, p_5, p_{10}, p_{13}, p_2, p_{14}\}$$

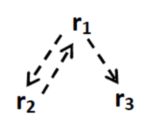


Example: Super-weak Acyclicity [Marnette09]

Analyzes nulls propagation through constraints

Example

$$\Sigma = \begin{array}{l} r_1 : N(\overset{p_1}{x}) \rightarrow \exists y\, \exists z\, E(\overset{p_2}{x}, \overset{p_3}{y}) \wedge S(\overset{p_4}{z}, \overset{p_5}{y}) \\ r_2 : E(\overset{p_6}{x_1}, \overset{p_7}{y_1}) \wedge S(\overset{p_8}{x_1}, \overset{p_9}{y_1}) \rightarrow N(\overset{p_{10}}{y_1}) \\ r_3 : E(\overset{p_{11}}{x_2}, \overset{p_{12}}{y_2}) \rightarrow E(\overset{p_{13}}{y_2}, \overset{p_{14}}{x_2}) \end{array}$$



$Move(r_1, y) = \{p_3, p_5, p_{10}, p_{13}, p_2, p_{14}\}$

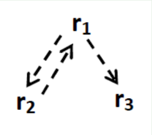
Then, we say that $r_1 \rightsquigarrow r_1$
because the null introduced by r_1 (p_3 and p_5) is copied again in the head of r_1 (p_2)

Example: Super-weak Acyclicity [Marnette09]

Analyzes nulls propagation through constraints

Example

$$\Sigma = \begin{array}{l} r_1 : \; N(\overset{p_1}{x}) \rightarrow \exists y \, \exists z \; E(\overset{p_2}{x}, \overset{p_3}{y}) \wedge S(\overset{p_4}{z}, \overset{p_5}{y}) \\ r_2 : \; E(\overset{p_6}{x_1}, \overset{p_7}{y_1}) \wedge S(\overset{p_8}{x_1}, \overset{p_9}{y_1}) \rightarrow N(\overset{p_{10}}{y_1}) \\ r_3 : \; E(\overset{p_{11}}{x_2}, \overset{p_{12}}{y_2}) \rightarrow E(\overset{p_{13}}{y_2}, \overset{p_{14}}{x_2}) \end{array}$$



$Move(r_1, y) = \{p_3, p_5, p_{10}, p_{13}, p_2, p_{14}\}$

Then, we say that $r_1 \rightsquigarrow r_1$
because the null introduced by r_1 (p_3 and p_5) is copied again in the head of r_1 (p_2)

The relation $\rightsquigarrow$ is cyclic and Σ is not in SwA

Local Stratification [Greco,Spezzano,Trubitsyna VLDB2011]

IDEA: testing the firing relation $<$ from *WA*-stratification in the *MOVE* construction

Example

$$\Sigma = \begin{array}{l} r_1 : N(\overset{p_1}{x}) \rightarrow \exists y\, \exists z\, E(\overset{p_2}{x}, \overset{p_3}{y}) \wedge S(\overset{p_4}{z}, \overset{p_5}{y}) \\ r_2 : E(\overset{p_6}{x_1}, \overset{p_7}{y_1}) \wedge S(\overset{p_8}{x_1}, \overset{p_9}{y_1}) \rightarrow N(\overset{p_{10}}{y_1}) \\ r_3 : E(\overset{p_{11}}{x_2}, \overset{p_{12}}{y_2}) \rightarrow E(\overset{p_{13}}{y_2}, \overset{p_{14}}{x_2}) \end{array}$$

Local Stratification [Greco,Spezzano,Trubitsyna VLDB2011]

IDEA: testing the firing relation $<$ from *WA*-stratification in the *MOVE* construction

Example

$$\Sigma = \begin{array}{ll} r_1: & N(\overset{p_1}{x}) \rightarrow \exists y\, \exists z\, E(\overset{p_2}{x}, \overset{p_3}{y}) \wedge S(\overset{p_4}{z}, \overset{p_5}{y}) \\ r_2: & E(\overset{p_6}{x_1}, \overset{p_7}{y_1}) \wedge S(\overset{p_8}{x_1}, \overset{p_9}{y_1}) \rightarrow N(\overset{p_{10}}{y_1}) \\ r_3: & E(\overset{p_{11}}{x_2}, \overset{p_{12}}{y_2}) \rightarrow E(\overset{p_{13}}{y_2}, \overset{p_{14}}{x_2}) \end{array}$$

$r_1 \not< r_2$

$MOVE(r_1, y) = \{p_3, p_5\}$

$p_{10} \notin MOVE(r_1, y)$ as $r_1 \not< r_2$

Local Stratification [Greco,Spezzano,Trubitsyna VLDB2011]

IDEA: testing the firing relation $<$ from *WA*-stratification in the *MOVE* construction

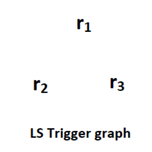
Example

$$\Sigma = \begin{array}{l} r_1 : \; N(\overset{p_1}{x}) \rightarrow \exists y \, \exists z \; E(\overset{p_2}{x}, \overset{p_3}{y}) \wedge S(\overset{p_4}{z}, \overset{p_5}{y}) \\ r_2 : \; E(\overset{p_6}{x_1}, \overset{p_7}{y_1}) \wedge S(\overset{p_8}{x_1}, \overset{p_9}{y_1}) \rightarrow N(\overset{p_{10}}{y_1}) \\ r_3 : \; E(\overset{p_{11}}{x_2}, \overset{p_{12}}{y_2}) \rightarrow E(\overset{p_{13}}{y_2}, \overset{p_{14}}{x_2}) \end{array}$$

$r_1 \not< r_2$

$MOVE(r_1, y) = \{p_3, p_5\}$

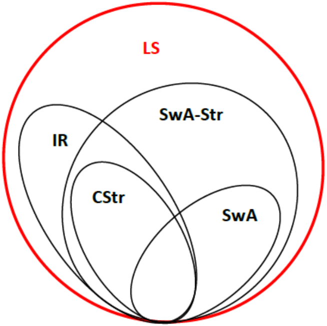
$p_{10} \notin MOVE(r_1, y)$ as $r_1 \not< r_2$



Conclusions

Conclusions

- We proposed new sufficient chase termination criteria, such as *WA*(*SC*,*SwA*)-Stratification and Local Stratification



Thank you!

Any question?