Implementing the Core Chase for **Horn-**$\mathcal{ALCH}$

Internship supervised by:
Jean-François Baget and David Carral
INRIA, Montpellier, GRAPHIK team

Maël Abily

2021

Entailment: a fundamental problem in knowledge representation

Conjunctive query entailment problem

Determine if the knowledge base K entails the query Q

Conjunctive query entailment problem

Determine if the knowledge base K entails the query Q

Knowledge base $K =$

Rules	**Facts (Initial Factbase)**
$\alpha = Mother(x) \rightarrow Parent(x)$	$Mother(\mathbf{Marie})$
$\beta = Parent(x) \rightarrow \exists y.IsTheParentOf(x, y)$	$IsTheSisterOf(\mathbf{Marie}, x)$
$\gamma = IsTheParentOf(x, y) \rightarrow Human(y)$	

Entailment: a fundamental problem in knowledge representation

Conjunctive query entailment problem

Determine if the knowledge base K entails the query Q

Knowledge base $K =$

Rules	**Facts (Initial Factbase)**
$\alpha = Mother(x) \rightarrow Parent(x)$ $\beta = Parent(x) \rightarrow \exists y. IsTheParentOf(x, y)$ $\gamma = IsTheParentOf(x, y) \rightarrow Human(y)$	$Mother(\textbf{Marie})$ $IsTheSisterOf(\textbf{Marie}, x)$

Query $Q = \{IsTheParentOf(\textbf{Marie}, x), Human(x)\}$

One of the solution to resolve the conjunctive query entailment problem: The Chase

Knowledge base $K =$

$\alpha = Mother(x) \rightarrow Parent(x)$ $\beta = Parent(x) \rightarrow \exists y. IsTheParentOf(x, y)$ $\gamma = IsTheParentOf(x, y) \rightarrow Human(y)$	$Mother(\textbf{Marie})$

Query $Q = \{IsTheParentOf(\textbf{Marie}, x), Human(x)\}$

One of the solution to resolve the conjunctive query entailment problem: The Chase

Knowledge base $K =$

$\alpha = Mother(x) \rightarrow Parent(x)$ $\beta = Parent(x) \rightarrow \exists y.IsTheParentOf(x, y)$ $\gamma = IsTheParentOf(x, y) \rightarrow Human(y)$	$Mother(\mathbf{Marie})$ $Parent(\mathbf{Marie})$

Query $Q = \{IsTheParentOf(\mathbf{Marie}, x), Human(x)\}$

One of the solution to resolve the conjunctive query entailment problem: The Chase

Knowledge base $K =$

$\alpha = Mother(x) \rightarrow Parent(x)$ $\beta = Parent(x) \rightarrow \exists y.IsTheParentOf(x, y)$ $\gamma = IsTheParentOf(x, y) \rightarrow Human(y)$	$Mother(\mathbf{Marie})$ $Parent(\mathbf{Marie})$ $IsTheParentOf(\mathbf{Marie}, y_0)$

$y_0 = f^y_\beta(\mathbf{Marie})$

Query $Q = \{IsTheParentOf(\mathbf{Marie}, x), Human(x)\}$

One of the solution to resolve the conjunctive query entailment problem: The Chase

Knowledge base $K =$

$\alpha = Mother(x) \rightarrow Parent(x)$ $\beta = Parent(x) \rightarrow \exists y. IsTheParentOf(x, y)$ $\gamma = IsTheParentOf(x, y) \rightarrow Human(y)$	$Mother(\textbf{Marie})$ $Parent(\textbf{Marie})$ $IsTheParentOf(\textbf{Marie}, y_0)$ $Human(y_0)$

$y_0 = f^y_\beta(\textbf{Marie})$

Query $Q = \{IsTheParentOf(\textbf{Marie}, x), Human(x)\}$

One of the solution to resolve the conjunctive query entailment problem: The Chase

Knowledge base $K =$

$\alpha = Mother(x) \rightarrow Parent(x)$	$Mother(\textbf{Marie})$
$\beta = Parent(x) \rightarrow \exists y. IsTheParentOf(x, y)$	$Parent(\textbf{Marie})$
$\gamma = IsTheParentOf(x, y) \rightarrow Human(y)$	$IsTheParentOf(\textbf{Marie}, y_0)$
	$Human(y_0)$

$y_0 = f^y_\beta(\textbf{Marie})$

Query $Q = \{IsTheParentOf(\textbf{Marie}, x), Human(x)\}$

$$\bullet \xrightarrow{\text{IsTheParentOf}} y_0\text{: Human}$$

Marie: Mother, Parent

The chase computes a universal model

Definition

A factbase U is a universal model of a knowledge base K if for every query Q, K entails Q if and only if U entails Q.

Theorem

For a knowledge base K, the result of the oblivious chase on K is an universal model of K.

The oblivious chase

one rule: $\alpha = R(x, y) \rightarrow \exists z.R(y, z) \wedge R(z, y)$ where $R = \mathit{IsTheFriendOf}$
two facts: $R(\textbf{Francis}, \textbf{Bob}), R(\textbf{Bob}, \textbf{Francis})$

$$\begin{tikzcd} \textbf{Francis} \arrow[r, bend left, "R"] & \textbf{Bob} \arrow[l, bend left, "R"] \end{tikzcd}$$

The oblivious chase

one rule: $\alpha = R(x,y) \rightarrow \exists z. R(y,z) \land R(z,y)$ where $R = \mathit{IsTheFriendOf}$
two facts: $R(\textbf{Francis}, \textbf{Bob}), R(\textbf{Bob}, \textbf{Francis})$

$$\begin{tikzcd}
\textbf{Francis} \arrow[r, bend left, "R"] & \textbf{Bob} \arrow[l, bend left, "R"] \arrow[r, bend left, "R"] & z_1 \arrow[l, bend left, "R"]
\end{tikzcd}$$

$z_1 = f^z_\alpha(\textbf{Francis}, \textbf{Bob})$

The oblivious chase

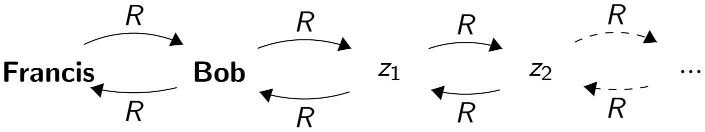
one rule: $\alpha = R(x, y) \rightarrow \exists z.R(y, z) \wedge R(z, y)$ where $R = \mathit{IsTheFriendOf}$
two facts: $R(\textbf{Francis}, \textbf{Bob}), R(\textbf{Bob}, \textbf{Francis})$

$$\begin{tikzcd}
\textbf{Francis} \arrow[r, bend left, "R"] & \textbf{Bob} \arrow[l, bend left, "R"] \arrow[r, bend left, "R"] & z_1 \arrow[l, bend left, "R"] \arrow[r, bend left, "R"] & z_2 \arrow[l, bend left, "R"]
\end{tikzcd}$$

$z_1 = f_\alpha^z(\textbf{Francis}, \textbf{Bob})$
$z_2 = f_\alpha^z(\textbf{Bob}, z_1)$

The oblivious chase

one rule: $\alpha = R(x, y) \rightarrow \exists z.R(y, z) \wedge R(z, y)$ where $R = IsTheFriendOf$
two facts: $R(\textbf{Francis}, \textbf{Bob}), R(\textbf{Bob}, \textbf{Francis})$



$z_1 = f_\alpha^z(\textbf{Francis}, \textbf{Bob})$
$z_2 = f_\alpha^z(\textbf{Bob}, z_1)$

The core chase

Definition

A core of a factbase F is one of the minimal subset of F that is logically equivalent to F.

For example, if $F = \{Human(x), Human(\mathbf{Marie})\}$, then $Core(F) = \{Human(\mathbf{Marie})\}$.

The core chase

Definition

A core of a factbase F is one of the minimal subset of F that is logically equivalent to F.

one rule: $\alpha = R(x, y) \rightarrow \exists z.R(y, z) \wedge R(z, y)$
two facts: $R(\textbf{Francis}, \textbf{Bob}), R(\textbf{Bob}, \textbf{Francis})$
$z_1 = f^z_\alpha(\textbf{Francis}, \textbf{Bob}), z_0 = f^z_\alpha(\textbf{Bob}, \textbf{Francis})$

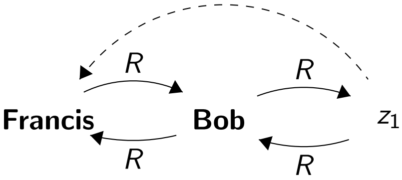
$$\begin{tikzcd} \textbf{Francis} \arrow[r, bend left, "R"] & \textbf{Bob} \arrow[l, bend left, "R"] \arrow[r, bend left, "R"] & z_1 \arrow[l, bend left, "R"] \end{tikzcd}$$

The core chase

Definition

A core of a factbase F is one of the minimal subset of F that is logically equivalent to F.

one rule: $\alpha = R(x, y) \rightarrow \exists z.R(y, z) \wedge R(z, y)$
two facts: $R(\textbf{Francis}, \textbf{Bob}), R(\textbf{Bob}, \textbf{Francis})$
$z_1 = f_\alpha^z(\textbf{Francis}, \textbf{Bob}), z_0 = f_\alpha^z(\textbf{Bob}, \textbf{Francis})$



Definition

A core of a factbase F is one of the minimal subset of F that is logically equivalent to F.

one rule: $\alpha = R(x, y) \rightarrow \exists z.R(y, z) \wedge R(z, y)$
two facts: $R(\textbf{Francis}, \textbf{Bob}), R(\textbf{Bob}, \textbf{Francis})$
$z_1 = f^z_\alpha(\textbf{Francis}, \textbf{Bob}), z_0 = f^z_\alpha(\textbf{Bob}, \textbf{Francis})$

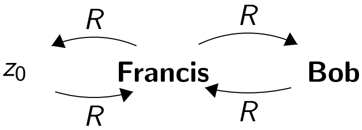
$$\begin{tikzcd} \textbf{Francis} \arrow[r, bend left, "R"] & \textbf{Bob} \arrow[l, bend left, "R"] \end{tikzcd}$$

The core chase

Definition

A core of a factbase F is one of the minimal subset of F that is logically equivalent to F.

one rule: $\alpha = R(x,y) \rightarrow \exists z.R(y,z) \wedge R(z,y)$
two facts: $R(\mathbf{Francis}, \mathbf{Bob}), R(\mathbf{Bob}, \mathbf{Francis})$
$z_1 = f_\alpha^z(\mathbf{Francis}, \mathbf{Bob}), z_0 = f_\alpha^z(\mathbf{Bob}, \mathbf{Francis})$

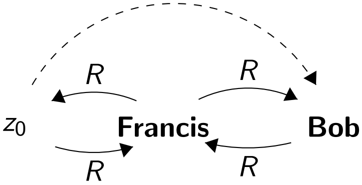


The core chase

Definition

A core of a factbase F is one of the minimal subset of F that is logically equivalent to F.

one rule: $\alpha = R(x, y) \rightarrow \exists z.R(y, z) \wedge R(z, y)$
two facts: $R(\mathbf{Francis}, \mathbf{Bob}), R(\mathbf{Bob}, \mathbf{Francis})$
$z_1 = f^z_\alpha(\mathbf{Francis}, \mathbf{Bob}), z_0 = f^z_\alpha(\mathbf{Bob}, \mathbf{Francis})$



The core chase

Definition

A core of a factbase F is one of the minimal subset of F that is logically equivalent to F.

one rule: $\alpha = R(x, y) \rightarrow \exists z.R(y, z) \wedge R(z, y)$
two facts: $R(\mathbf{Francis}, \mathbf{Bob}), R(\mathbf{Bob}, \mathbf{Francis})$
$z_1 = f_\alpha^z(\mathbf{Francis}, \mathbf{Bob}), z_0 = f_\alpha^z(\mathbf{Bob}, \mathbf{Francis})$

$$\begin{tikzcd} \mathbf{Francis} \arrow[r, bend left, "R"] & \mathbf{Bob} \arrow[l, bend left, "R"] \end{tikzcd}$$

The core chase is the best in term of terminaison

Theorem

The knowledge base K admits a finite universal model if and only if the core chase terminates on K. In this case, the core chase computes a universal model.

The goals of my internship

- The core chase is hard to compute but it terminates more often.
- Computing the core of a factbase is NP-complete.

The goals of my internship

- The core chase is hard to compute but it terminates more often.
- Computing the core of a factbase is NP-complete.

- Identify a fragment of rules such that we can quickly compute the core.
- Give an algorithm.
- Prove the correctness and the completeness of the algorithm.

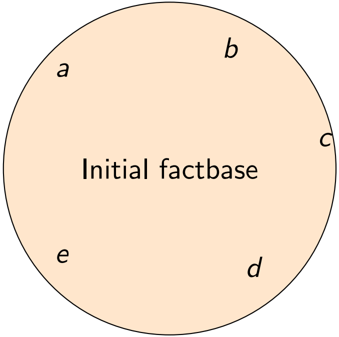
Our restriction: Horn-$\mathcal{ALCH}$

The rules are of the form:

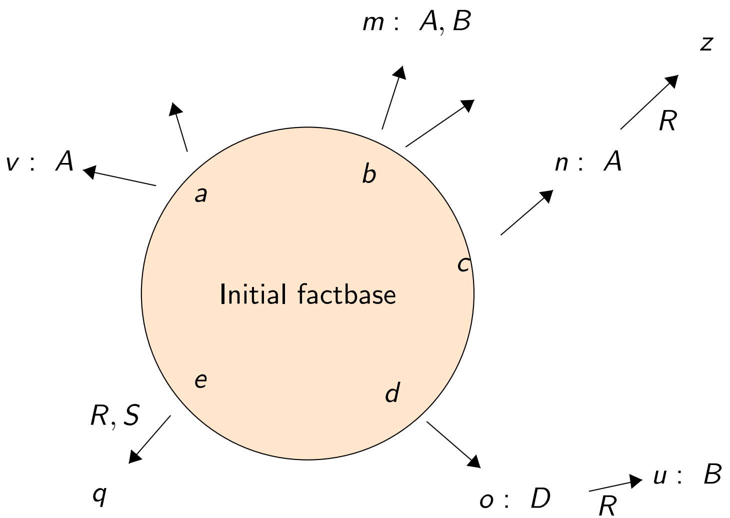
$$
\begin{aligned}
A(x) &\rightarrow \exists y.R(x,y) \wedge B(y) && (1)\\
A(x) \wedge B(x) &\rightarrow C(x) && (2)\\
A(x) \wedge R(x,y) &\rightarrow B(y) && (3)\\
R(x,y) \wedge B(y) &\rightarrow A(x) && (4)\\
R(x,y) \wedge S(x,y) &\rightarrow V(x,y) && (5)
\end{aligned}
$$

The saturation of the initial factbase will create facts that have a tree shape.

Example

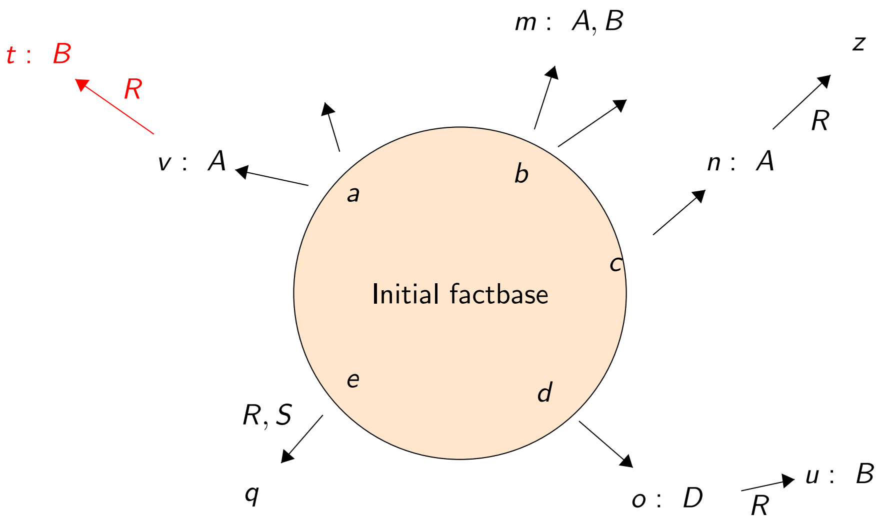


Example

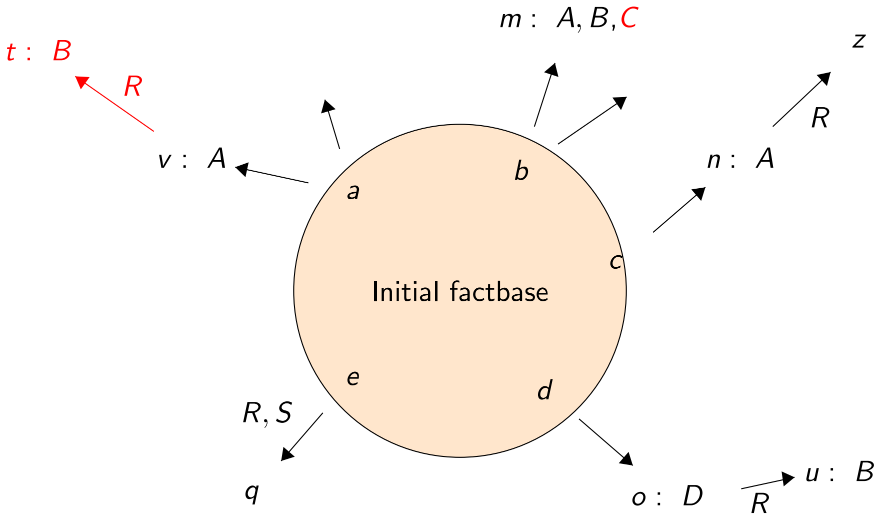


$$A(x) \rightarrow \exists y. R(x, y) \wedge B(y)$$

Example

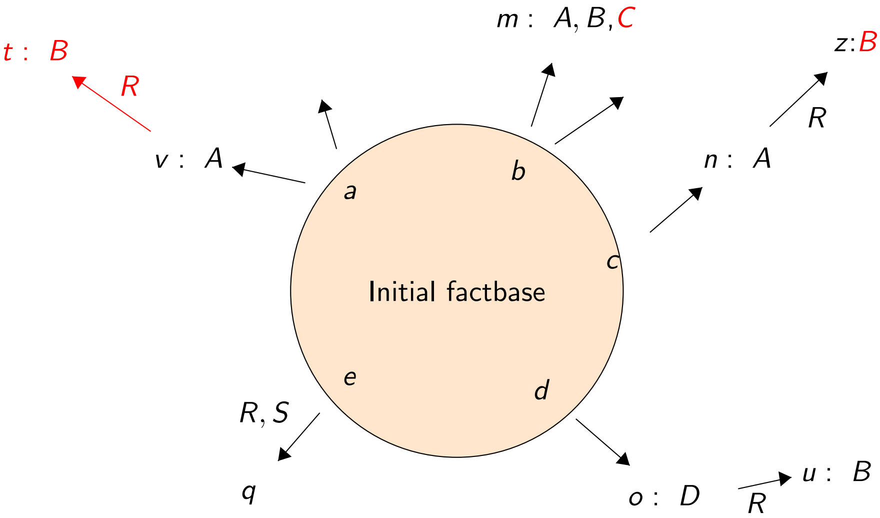


$$A(x) \rightarrow \exists y. R(x,y) \wedge B(y)$$



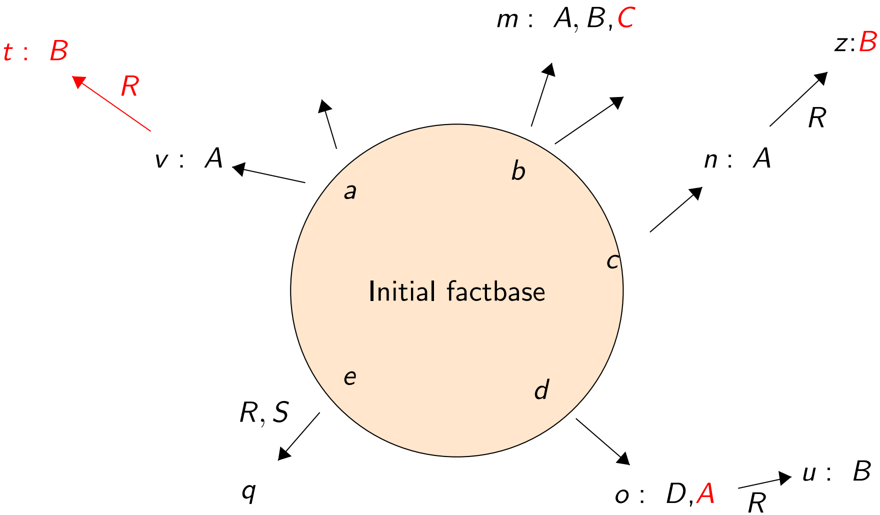
$$A(x) \wedge B(x) \rightarrow C(x)$$

Example



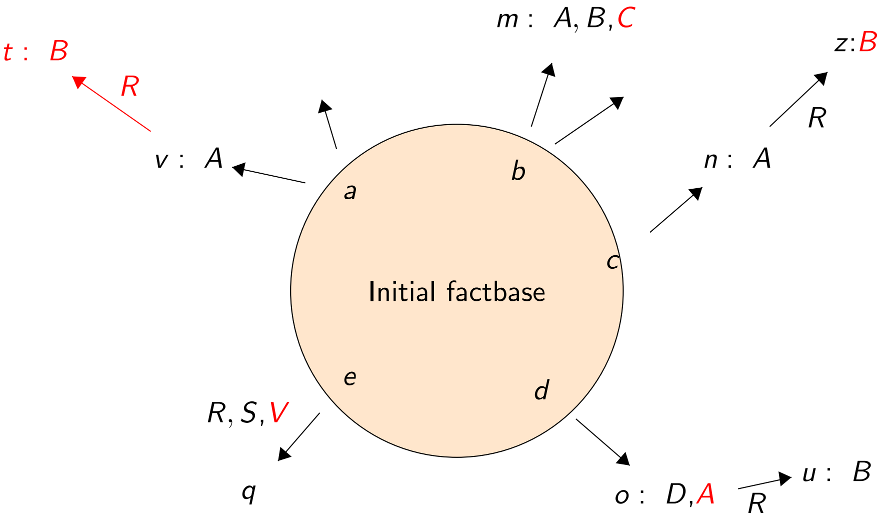
$$A(x) \wedge R(x, y) \rightarrow B(y)$$

Example



$$R(x, y) \wedge B(y) \rightarrow A(x)$$

Example



$$R(x, y) \land S(x, y) \rightarrow V(x, y)$$

A better algorithm in this restriction: The merge chase

The variable z is mergeable on the term t:

$$\begin{tikzcd}
t : A, B & & z : A \\
& u : A \arrow[ul, "{R, S}"'] \arrow[ur, "R"'] &
\end{tikzcd}$$

A better algorithm in this restriction: The merge chase

The variable z is mergeable on the term t:

$$
\begin{tikzcd}
t : A, B & & z : A \\
& u : A \arrow[ul, "{R, S}"'] \arrow[ur, "R"'] &
\end{tikzcd}
$$

Theorem

If the variable z is mergeable on the term t, the facts generated on z can be generated on t.

A better algorithm in this restriction: The merge chase

The variable z is mergeable on the term t:

$$A(x) \rightarrow \exists y.R(x,y) \wedge B(y)$$

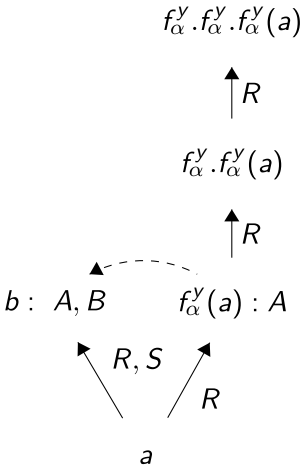
$$\begin{tikzcd}
t : A, B & & z : A \\
& u : A \arrow[ul, "{R, S}"] \arrow[ur, "R"'] &
\end{tikzcd}$$

Theorem

If the variable z is mergeable on the term t, the facts generated on z can be generated on t.

A better algorithm in this restriction: The merge chase

The atomic merge



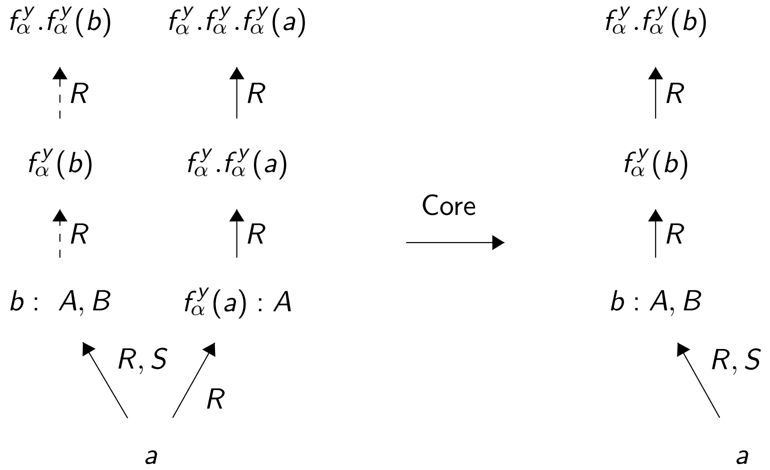
A better algorithm in this restriction: The merge chase

The atomic merge

$$
\begin{tikzcd}
f^y_\alpha . f^y_\alpha (b) & & f^y_\alpha . f^y_\alpha . f^y_\alpha (a) \\
f^y_\alpha (b) \arrow[u, dashed, "R"'] & & f^y_\alpha . f^y_\alpha (a) \arrow[u, "R"'] \\
b : \ A, B \arrow[u, dashed, "R"'] & & f^y_\alpha (a) : A \arrow[u, "R"'] \\
& a \arrow[ul, "{R, S}"'] \arrow[ur, "R"'] &
\end{tikzcd}
$$

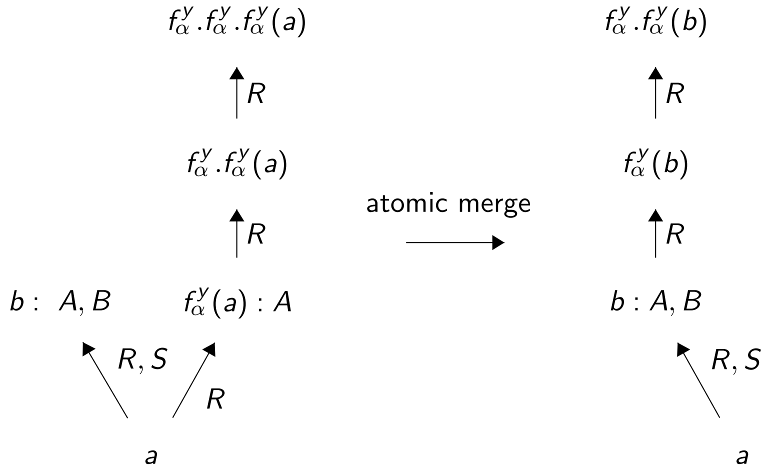
A better algorithm in this restriction: The merge chase

The atomic merge



A better algorithm in this restriction: The merge chase

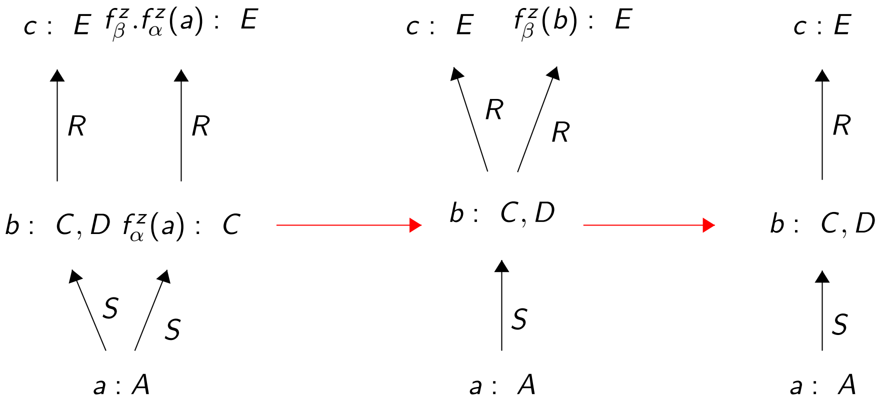
The atomic merge



The merge

Definition

For a factbase F, $Merge(F)$ is obtained by applying all the possible atomic mergings on F.



The merge chase

Theorem

For a factbase F, $Merge(F)$ is a core.

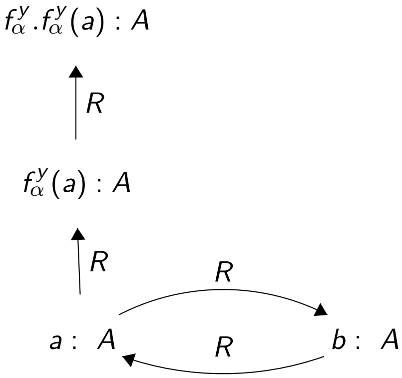
Definition

The merge chase for a **Horn-**$\mathcal{ALCH}$ knowledge base K is the core chase where we replace the computation of a core by the merge operation.

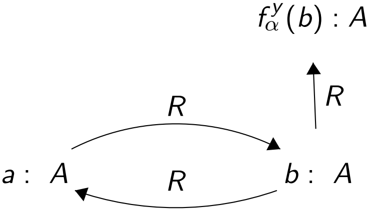
Corollary

*The **Horn-**$\mathcal{ALCH}$ knowledge base K admits a finite universal model if and only if the merge chase terminates on K. In this case, the merge chase computes a universal model.*

The atomic merging operation has to be applied exhaustively



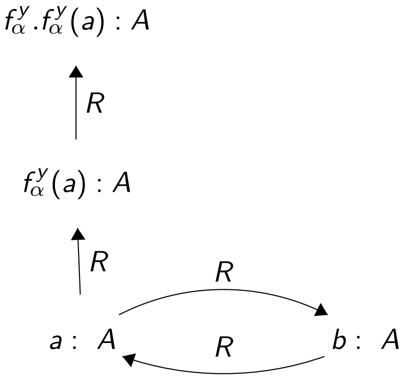
The atomic merging operation has to be applied exhaustively



The atomic merging operation has to be applied exhaustively

$$\begin{tikzcd}
 & f_\alpha^y.f_\alpha^y.f_\alpha^y(b) : A \\
 & f_\alpha^y.f_\alpha^y(b) : A \arrow[u, "R"'] \\
 & f_\alpha^y(b) : A \arrow[u, "R"'] \\
a : A \arrow[r, bend left, "R"] & b : A \arrow[u, "R"'] \arrow[l, bend left, "R"]
\end{tikzcd}$$

The atomic merging operation has to be applied exhaustively



The atomic merging operation has to be applied exhaustively

$$\begin{tikzcd} & f^{y}_{\alpha}(b) : A \\ a : A \arrow[r, bend left, "R"] & b : A \arrow[l, bend left, "R"] \arrow[u, "R"'] \end{tikzcd}$$

The atomic merging operation has to be applied exhaustively

$$\begin{tikzcd} a:\ A \arrow[r, bend left, "R"] & b:\ A \arrow[l, bend left, "R"] \end{tikzcd}$$

Conclusion

- Theoretical study of an alternative of the core chase in **Horn-**$\mathcal{ALCH}$
  - define the merging operation;
  - prove the correctness and completness of the merge chase;

Conclusion

- Theoretical study of an alternative of the core chase in **Horn**-$\mathcal{ALCH}$
  - define the merging operation;
  - prove the correctness and completness of the merge chase;
- Furtherwork before publication
  - Extend our work to **Horn**-$\mathcal{ALCHI}$: $R(x,y) \wedge S(x,y) \rightarrow T(\textcolor{red}{y},\textcolor{red}{x})$;
  - Experimentation on the Graal platform using standard benchmark.