# Vérification de systèmes et réécriture de plus en plus efficace 

Yohan Boichut<br>LIFO - Université d'Orléans

JIRC 08, Orléans

9 octobre 2008

## Context

Java Bytecode analysis

- Rewriting semantics for the Java Bytecode
- Static analysis from reachability analysis in rewriting
- Tree automata technique [RTA98]
- Timbuk tool (http://www.irisa.fr/lande/genet/timbuk)


## Context

Java Bytecode analysis

- Rewriting semantics for the Java Bytecode
- Static analysis from reachability analysis in rewriting
- Tree automata technique [RTA98]
- Timbuk tool (http://www.irisa.fr/lande/genet/timbuk)


## Rewriting Semantics for the Java Bytecode

## JVM Execution Trace



For a given program P

- JVM states as Terms
- Rewrite rules for
- the Bytecode instructions interpretation (generic rules)
- the program


## Reachability Analysis in Rewriting



E : initial terms
$\longrightarrow J V M$ initial state
R : term rewriting system
$\longrightarrow$ JVM Transition relation for the given program
Bad : forbidden terms
$\longrightarrow$ Forbidden JVM states

## Tree automata completion



- A set of terms is represented by a tree automaton language
- $\mathcal{A}_{i+1}=\mathcal{A}_{i}+$ new transitions and states
- Completion stops when a fix point automaton is found


## Computing substitutions for a completion step



## Computing substitutions for a completion step



|  | a | $\rightarrow$ | $q_{a}$ |
| :---: | :---: | :---: | :---: |
|  | $b$ | $\rightarrow$ | $q_{b}$ |
|  | c | $\rightarrow$ | $q_{c}$ |
| $\mathcal{A}$ | $g\left(q_{a}, q_{b}\right)$ | $\rightarrow$ | $q_{g 1}$ |
|  | $f\left(q_{c}\right)$ | $\rightarrow$ | $q_{f}$ |
|  | $g\left(q_{g 1}, q_{f}\right)$ | $\rightarrow$ | $q_{g} 2$ |

## Computing substitutions for a completion step



|  | $a$ | $\rightarrow$ | $q_{a}$ |
| :--- | :--- | :--- | :--- |
|  | $b$ | $\rightarrow$ | $q_{b}$ |
|  | $c$ | $\rightarrow$ | $q_{c}$ |
| $\mathcal{A}$ | $g\left(q_{a}, q_{b}\right)$ | $\rightarrow$ | $q_{g 1}$ |
|  | $f\left(q_{c}\right)$ | $\rightarrow$ | $q_{f}$ |
|  | $g\left(q_{g 1}, q_{f}\right)$ | $\rightarrow$ | $q_{g 2}$ |



## Computing substitutions for a completion step



|  | $a$ |  |  |
| :--- | :--- | :--- | :--- |
|  | $b$ |  | $q_{a}$ |
|  | $c$ | $\rightarrow$ | $q_{b}$ |
| $\mathcal{A}$ | $g\left(q_{a}, q_{b}\right)$ | $\rightarrow$ | $q_{c}$ |
|  | $f\left(q_{c}\right)$ | $\rightarrow$ | $q_{g 1}$ |
|  | $g\left(q_{g 1}, q_{f}\right)$ | $\rightarrow$ | $q_{f}$ |
|  |  |  | $q_{g 2}$ |



## Computing substitutions for a completion step



|  | $a$ |  | $q_{a}$ |
| :--- | :--- | :--- | :--- |
|  | $b$ | $\rightarrow$ | $q_{b}$ |
|  | $c$ | $\rightarrow$ | $q_{c}$ |
| $\mathcal{A}$ | $g\left(q_{a}, q_{b}\right)$ | $\rightarrow$ | $q_{g 1}$ |
|  | $f\left(q_{c}\right)$ | $\rightarrow$ | $q_{f}$ |
|  | $g\left(q_{g 1}, q_{f}\right)$ | $\rightarrow$ | $q_{g 2}$ |



## Computing substitutions for a completion step



|  | $a$ |  |  |
| :--- | :--- | :--- | :--- |
|  | $b$ |  | $q_{a}$ |
|  | $c$ | $\rightarrow$ | $q_{b}$ |
| $\mathcal{A}$ | $g\left(q_{a}, q_{b}\right)$ | $\rightarrow$ | $q_{c}$ |
|  | $f\left(q_{c}\right)$ | $\rightarrow$ | $q_{g 1}$ |
|  | $g\left(q_{g 1}, q_{f}\right)$ | $\rightarrow$ | $q_{f}$ |
|  |  |  | $q_{g 2}$ |



## Computing substitutions for a completion step



|  | $a$ |  |  |
| :--- | :--- | :--- | :--- |
|  | $b$ |  | $q_{a}$ |
|  | $c$ | $\rightarrow$ | $q_{b}$ |
| $\mathcal{A}$ | $g\left(q_{a}, q_{b}\right)$ | $\rightarrow$ | $q_{c}$ |
|  | $f\left(q_{c}\right)$ | $\rightarrow$ | $q_{g 1}$ |
|  | $g\left(q_{g 1}, q_{f}\right)$ | $\rightarrow$ | $q_{f}$ |
|  |  |  | $q_{g 2}$ |



## Computing substitutions for a completion step



|  | $a$ |  |  |
| :--- | :--- | :--- | :--- |
|  | $b$ |  | $q_{a}$ |
|  | $c$ | $\rightarrow$ | $q_{b}$ |
| $\mathcal{A}$ | $g\left(q_{a}, q_{b}\right)$ | $\rightarrow$ | $q_{c}$ |
|  | $f\left(q_{c}\right)$ | $\rightarrow$ | $q_{g 1}$ |
|  | $g\left(q_{g 1}, q_{f}\right)$ | $\rightarrow$ | $q_{f}$ |
|  |  |  | $q_{g 2}$ |



## Computing substitutions for a completion step



|  | $a$ |  | $q_{a}$ |
| :--- | :--- | :--- | :--- |
|  | $b$ | $\rightarrow$ | $q_{b}$ |
|  | $c$ | $\rightarrow$ | $q_{c}$ |
| $\mathcal{A}$ | $g\left(q_{a}, q_{b}\right)$ | $\rightarrow$ | $q_{g 1}$ |
|  | $f\left(q_{c}\right)$ | $\rightarrow$ | $q_{f}$ |
|  | $g\left(q_{g 1}, q_{f}\right)$ |  | $\rightarrow$ |
| $q_{g 2}$ |  |  |  |



## Computing substitutions for a completion step



|  | $a$ |  |  |
| :--- | :--- | :--- | :--- |
|  | $b$ |  | $q_{a}$ |
|  | $c$ | $\rightarrow$ | $q_{b}$ |
| $\mathcal{A}$ | $g\left(q_{a}, q_{b}\right)$ | $\rightarrow$ | $q_{c}$ |
|  | $f\left(q_{c}\right)$ | $\rightarrow$ | $q_{g 1}$ |
|  | $g\left(q_{g 1}, q_{f}\right)$ | $\rightarrow$ | $q_{f}$ |
|  |  |  | $q_{g 2}$ |



## Computing substitutions for a completion step



|  | $a$ |  |  |
| :--- | :--- | :--- | :--- |
|  | $b$ |  | $q_{a}$ |
|  | $c$ | $\rightarrow$ | $q_{b}$ |
| $\mathcal{A}$ | $g\left(q_{a}, q_{b}\right)$ | $\rightarrow$ | $q_{c}$ |
|  | $f\left(q_{c}\right)$ | $\rightarrow$ | $q_{g 1}$ |
|  | $g\left(q_{g 1}, q_{f}\right)$ | $\rightarrow$ | $q_{f}$ |
|  |  |  | $q_{g 2}$ |



## Computing substitutions for a completion step



|  | $a$ |  | $q_{a}$ |
| :--- | :--- | :--- | :--- |
|  | $b$ | $\rightarrow$ | $q_{b}$ |
|  | $c$ | $\rightarrow$ | $q_{c}$ |
| $\mathcal{A}$ | $g\left(q_{a}, q_{b}\right)$ | $\rightarrow$ | $q_{g 1}$ |
|  | $f\left(q_{c}\right)$ | $\rightarrow$ | $q_{f}$ |
|  | $g\left(q_{g 1}, q_{f}\right)$ | $\rightarrow$ | $q_{g 2}$ |



## Computing substitutions for a completion step



|  | $a$ |  | $q_{a}$ |
| :--- | :--- | :--- | :--- |
|  | $b$ | $\rightarrow$ | $q_{b}$ |
|  | $c$ | $\rightarrow$ | $q_{c}$ |
| $\mathcal{A}$ | $g\left(q_{a}, q_{b}\right)$ | $\rightarrow$ | $q_{g 1}$ |
|  | $f\left(q_{c}\right)$ | $\rightarrow$ | $q_{f}$ |
|  | $g\left(q_{g 1}, q_{f}\right)$ | $\rightarrow$ | $q_{g 2}$ |



## Computing substitutions for a completion step



|  |  |  |  |
| :--- | :--- | :--- | :--- |
|  |  |  |  |
|  | $b$ |  | $q_{a}$ |
|  | $c$ | $\rightarrow$ | $q_{b}$ |
| $\mathcal{A}$ | $g\left(q_{a}, q_{b}\right)$ | $\rightarrow$ | $q_{c}$ |
|  |  | $\rightarrow$ | $q_{g 1}$ |
|  | $g\left(q_{c}\right)$ | $\rightarrow$ | $q_{f}$ |
|  | $g\left(q_{g 1}, q_{f}\right)$ |  | $\rightarrow$ |
| $q_{g 2}$ |  |  |  |



## Computing substitutions for a completion step



|  | $a$ |  |  |
| :--- | :--- | :--- | :--- |
|  | $b$ |  | $q_{a}$ |
|  | $c$ | $\rightarrow$ | $q_{b}$ |
| $\mathcal{A}$ | $g\left(q_{a}, q_{b}\right)$ | $\rightarrow$ | $q_{c}$ |
|  | $f\left(q_{c}\right)$ | $\rightarrow$ | $q_{g 1}$ |
|  | $g\left(q_{g 1}, q_{f}\right)$ | $\rightarrow$ | $q_{f}$ |
|  |  |  | $q_{g 2}$ |



## Computing substitutions for a completion step



|  | $a$ |  | $q_{a}$ |
| :--- | :--- | :--- | :--- |
|  | $b$ | $\rightarrow$ | $q_{b}$ |
|  | $c$ | $\rightarrow$ | $q_{c}$ |
| $\mathcal{A}$ | $g\left(q_{a}, q_{b}\right)$ | $\rightarrow$ | $q_{g 1}$ |
|  | $f\left(q_{c}\right)$ | $\rightarrow$ | $q_{f}$ |
|  | $g\left(q_{g 1}, q_{f}\right)$ | $\rightarrow$ | $q_{g 2}$ |



## Computing substitutions for a completion step



|  | $a$ |  | $q_{a}$ |
| :--- | :--- | :--- | :--- |
|  | $b$ | $\rightarrow$ | $q_{b}$ |
|  | $c$ | $\rightarrow$ | $q_{c}$ |
| $\mathcal{A}$ | $g\left(q_{a}, q_{b}\right)$ | $\rightarrow$ | $q_{g 1}$ |
|  | $f\left(q_{c}\right)$ | $\rightarrow$ | $q_{f}$ |
|  | $g\left(q_{g 1}, q_{f}\right)$ |  | $\rightarrow$ |
| $l_{g 2}$ |  |  |  |



$$
\rightarrow
$$

$$
/_{x\left[q_{b}\right]}^{g\left[q_{g 2}\right]}
$$

## Computing substitutions for a completion step



|  | $a$ |  | $q_{a}$ |
| :--- | :--- | :--- | :--- |
|  | $b$ | $\rightarrow$ | $q_{b}$ |
|  | $c$ | $\rightarrow$ | $q_{c}$ |
| $\mathcal{A}$ | $g\left(q_{a}, q_{b}\right)$ | $\rightarrow$ | $q_{g 1}$ |
|  | $f\left(q_{c}\right)$ | $\rightarrow$ | $q_{f}$ |
|  | $g\left(q_{g 1}, q_{f}\right)$ | $\rightarrow$ | $q_{g 2}$ |
|  | $g\left(q_{b}, q_{c}\right)$ | $\rightarrow$ | $q_{g 2}$ |



$$
\longrightarrow
$$

$$
/_{x\left[q_{b}\right]}^{g\left[q_{g 2}\right]} \backslash_{y\left[q_{c}\right]}
$$

## Don't forget that it is used for Verification!

## - Security protocols: [CADE00,WITS03,CAV05,TFIT06,ICTAC06] - Java program verification: [RTA 07]

## Don't forget that it is used for Verification!

- Security protocols: [CADE00,WITS03,CAV05,TFIT06,ICTAC06]
- Java program verification: [RTA 07]


## Don't forget that it is used for Verification!

- Security protocols: [CADE00,WITS03,CAV05,TFIT06,ICTAC06]
- Java program verification: [RTA 07]


## Don't forget that it is used for Verification!

- Security protocols: [CADE00,WITS03,CAV05,TFIT06,ICTAC06]
- Java program verification: [RTA 07]

But for the verification of Java programs...

- TRS are huge (more than 600 rules for a bubble sort program)
- Computation times with Timbuk may exceed 4 days!


## Don't forget that it is used for Verification!

- Security protocols: [CADE00,WITS03,CAV05,TFIT06,ICTAC06]
- Java program verification: [RTA 07]

But for the verification of Java programs...

- TRS are huge (more than 600 rules for a bubble sort program)
- Computation times with Timbuk may exceed 4 days!

Goal: Propose practical techniques to solve scalability issues

## Reachability Preserving TRS Transformation

Fact: Collecting all possible ground instances of a deep pattern may be expensive

Idea: Transform TRS into simpler TRS

- A simple form for the left hand-side of rules (depth max=2)
- Flat: $f\left(x_{1}, \ldots, x_{n}\right)$ or $c$
- $f\left(t_{1}, \ldots, t_{n}\right)$ where each $t_{i}$ is flat
- Reachability preserving (Terms computed with the original TRS must be also computed by the resulting TRS)


## Reachability Preserving TRS Transformation

Fact: Collecting all possible ground instances of a deep pattern may be expensive

Idea: Transform TRS into simpler TRS

- A simple form for the left hand-side of rules (depth max=2)
- Flat: $f\left(x_{1}, \ldots, x_{n}\right)$ or $c$
- $f\left(t_{1}, \ldots, t_{n}\right)$ where each $t_{i}$ is flat
- Reachability preserving (Terms computed with the original TRS must be also computed by the resulting TRS)


## Reachability Preserving TRS Transformation

Fact: Collecting all possible ground instances of a deep pattern may be expensive

Idea: Transform TRS into simpler TRS

- A simple form for the left hand-side of rules (depth max=2)
- $f\left(t_{1}, \ldots, t_{n}\right)$ where each $t_{i}$ is flat
- Reachability preserving (Terms computed with the original TRS must be also computed by the resulting TRS)


## Reachability Preserving TRS Transformation

Fact: Collecting all possible ground instances of a deep pattern may be expensive

Idea: Transform TRS into simpler TRS

- A simple form for the left hand-side of rules (depth max=2)
- Flat: $f\left(x_{1}, \ldots, x_{n}\right)$ or $c$
- $f\left(t_{1}, \ldots, t_{n}\right)$ where each $t_{i}$ is flat
- Reachability preserving (Terms computed with the original TRS must be also computed by the resulting TRS)


## Reachability Preserving TRS Transformation

Fact: Collecting all possible ground instances of a deep pattern may be expensive

Idea: Transform TRS into simpler TRS

- A simple form for the left hand-side of rules (depth max=2)
- Flat: $f\left(x_{1}, \ldots, x_{n}\right)$ or $c$
- $f\left(t_{1}, \ldots, t_{n}\right)$ where each $t_{i}$ is flat
- Reachability preserving (Terms computed with the original TRS must be also computed by the resulting TRS)


## Transformation of the example rule

$\mathcal{R}$


## Transformation of the example rule



## Transformation of the example rule



## Transformation of the example rule



## Transformation of the example rule



$$
\begin{array}{lll}
a & \rightarrow & C_{1} \\
g\left(C_{1}, x\right) & \rightarrow & C_{2}(x)
\end{array}
$$

## Transformation of the example rule



| $a$ | $\rightarrow$ | $C_{1}$ |
| :--- | :--- | :--- |
| $g\left(C_{1}, x\right)$ | $\rightarrow$ | $C_{2}(x)$ |
| $f(y)$ | $\rightarrow$ | $C_{3}(y)$ |

## Transformation of the example rule



| $a$ | $\rightarrow$ | $C_{1}$ |
| :--- | :--- | :--- |
| $g\left(C_{1}, x\right)$ | $\rightarrow$ | $C_{2}(x)$ |
| $f(y)$ | $\rightarrow$ | $C_{3}(y)$ |

## Transformation of the example rule



| $a$ | $\rightarrow$ | $C_{1}$ |
| :--- | :--- | :--- |
| $g\left(C_{1}, x\right)$ | $\rightarrow$ | $C_{2}(x)$ |
| $f(y)$ | $\rightarrow$ | $C_{3}(y)$ |
| $g\left(C_{2}(x), C_{3}(y)\right)$ | $\rightarrow$ | $C_{4}(x, y)$ |

## Transformation of the example rule



$$
\begin{array}{lll}
a & \rightarrow & C_{1} \\
g\left(C_{1}, x\right) & \rightarrow & C_{2}(x) \\
f(y) & \rightarrow & C_{3}(y) \\
g\left(C_{2}(x), C_{3}(y)\right) & \rightarrow & C_{4}(x, y)
\end{array}
$$

## Transformation of the example rule



$$
\begin{array}{lll}
a & \rightarrow & C_{1} \\
g\left(C_{1}, x\right) & \rightarrow & C_{2}(x) \\
f(y) & \rightarrow & C_{3}(y) \\
g\left(C_{2}(x), C_{3}(y)\right) & \rightarrow & C_{4}(x, y) \\
C_{4}(x, y) & \rightarrow & g(x, y)
\end{array}
$$

## Transformation of the example rule

$\mathcal{R}$


$$
\begin{array}{lll}
a & \rightarrow & C_{1} \\
g\left(C_{1}, x\right) & \rightarrow & C_{2}(x) \\
f(y) & \rightarrow & C_{3}(y) \\
g\left(C_{2}(x), C_{3}(y)\right) & \rightarrow & C_{4}(x, y) \\
C_{4}(x, y) & \rightarrow & g(x, y)
\end{array}
$$

## Transformation of the example rule

$\mathcal{R}$


$$
\begin{array}{llll} 
& a & \rightarrow & C_{1} \\
& g\left(C_{1}, x\right) & \rightarrow & C_{2}(x) \\
b(\mathcal{R}) & f(y) & \rightarrow & C_{3}(y) \\
& g\left(C_{2}(x), C_{3}(y)\right) & \rightarrow & C_{4}(x, y) \\
& C_{4}(x, y) & \rightarrow & g(x, y) \\
\forall t, t^{\prime} \in \mathcal{T}(\mathcal{F}) \cdot t \rightarrow \mathcal{R} t^{\prime} & \longrightarrow t \rightarrow_{\phi(\mathcal{R})}^{*} t^{\prime}
\end{array}
$$

## Main Result



## Main Result

An over-approximation computed for $\phi(\mathcal{R})$ is also an over-approximation for $\mathcal{R}$


## Dedicated completion algorithm (1)

## Facts

- For each $I \rightarrow r \in \phi(\mathcal{R})$, I does not exceed a depth of 2

$$
g\left(C_{2}(x), C_{3}(y)\right) \rightarrow C_{4}(x, y)
$$

- Very close to a direct pattern-matching on transitions

$$
g\left(q_{g 1}, q_{f}\right) \rightarrow q_{g 2}
$$

- For this transition, the current matching algorithm computes all possible instances from $g\left(q_{g 1}, q_{f}\right)$

Can we reduce the substitution computation to a simple pattern-matching problem?

## Dedicated completion algorithm (1)

## Facts

- For each $I \rightarrow r \in \phi(\mathcal{R})$, I does not exceed a depth of 2

$$
g\left(C_{2}(x), C_{3}(y)\right) \rightarrow C_{4}(x, y)
$$

- Very close to a direct pattern-matching on transitions

$$
g\left(q_{g 1}, q_{f}\right) \rightarrow q_{g 2}
$$

- For this transition, the current matching algorithm computes all possible instances from $g\left(q_{g 1}, q_{f}\right)$

Can we reduce the substitution computation to a simple pattern-matching problem?

## Dedicated completion algorithm (1)

## Facts

- For each $I \rightarrow r \in \phi(\mathcal{R})$, I does not exceed a depth of 2

$$
g\left(C_{2}(x), C_{3}(y)\right) \rightarrow C_{4}(x, y)
$$

- Very close to a direct pattern-matching on transitions

$$
g\left(q_{g 1}, q_{f}\right) \rightarrow q_{g 2}
$$

- For this transition, the current matching algorithm computes all possible instances from $g\left(q_{g 1}, q_{f}\right)$

Can we reduce the substitution computation to a simple pattern-matching problem?

## Dedicated completion algorithm (2)

| $a$ | $\rightarrow$ | $q_{a}$ |  |  |  |
| :--- | :--- | :--- | :--- | :--- | :--- |
| $b$ | $\rightarrow$ | $q_{b}$ |  |  |  |
| $c$ | $\rightarrow$ | $q_{c}$ | $a$ |  | $C_{1}$ |
| $g\left(q_{a}, q_{b}\right)$ | $\rightarrow$ | $q_{g 1}$ | $g\left(C_{1}, x\right)$ |  | $\rightarrow$ |
| $C_{2}(x)$ |  |  |  |  |  |
| $f\left(q_{c}\right)$ | $\rightarrow$ | $q_{f}$ | $f(y)$ | $\rightarrow$ | $C_{3}(y)$ |
| $g\left(q_{g 1}, q_{f}\right)$ | $\rightarrow$ | $q_{g 2}$ | $g\left(C_{2}(x), C_{3}(y)\right)$ | $\rightarrow$ | $C_{4}(x, y)$ |
| $g\left(q_{b}, q_{c}\right)$ | $\rightarrow$ | $q_{g 2}$ | $C_{4}(x, y)$ | $\rightarrow$ | $g(x, y)$ |
| $C_{1}$ | $\rightarrow$ | $q_{a}$ |  |  |  |
| $C_{3}\left(q_{c}\right)$ | $\rightarrow$ | $q_{f}$ |  |  |  |
| $C_{2}\left(q_{b}\right)$ | $\rightarrow$ | $q_{g 1}$ |  |  |  |

## Dedicated completion algorithm (2)

| $a$ | $\rightarrow$ | $q_{a}$ |  |  |  |
| :--- | :--- | :--- | :--- | :--- | :--- |
| $b$ | $\rightarrow$ | $q_{b}$ |  |  |  |
| $c$ | $\rightarrow$ | $q_{c}$ | $a$ |  | $C_{1}$ |
| $g\left(q_{a}, q_{b}\right)$ | $\rightarrow$ | $q_{g 1}$ | $g\left(C_{1}, x\right)$ |  | $\rightarrow$ |
| $C_{2}(x)$ |  |  |  |  |  |
| $f\left(q_{c}\right)$ | $\rightarrow$ | $q_{f}$ | $f(y)$ | $\rightarrow$ | $C_{3}(y)$ |
| $g\left(q_{g 1}, q_{f}\right)$ | $\rightarrow$ | $q_{g 2}$ | $g\left(C_{2}(x), C_{3}(y)\right)$ | $\rightarrow$ | $C_{4}(x, y)$ |
| $g\left(q_{b}, q_{c}\right)$ | $\rightarrow$ | $q_{g 2}$ | $C_{4}(x, y)$ | $\rightarrow$ | $g(x, y)$ |
| $C_{1}$ | $\rightarrow$ | $q_{a}$ |  |  |  |
| $C_{3}\left(q_{c}\right)$ | $\rightarrow$ | $q_{f}$ |  |  |  |
| $C_{2}\left(q_{b}\right)$ | $\rightarrow$ | $q_{g 1}$ |  |  |  |

## Dedicated completion algorithm (2)

| $a$ | $\rightarrow$ | $q_{a}$ |
| :--- | :--- | :--- |
| $b$ | $\rightarrow$ | $q_{b}$ |
| $c$ | $\rightarrow$ | $q_{c}$ |
| $g\left(q_{a}, q_{b}\right)$ | $\rightarrow$ | $q_{g 1}$ |
| $f\left(q_{c}\right)$ | $\rightarrow$ | $q_{f}$ |
| $g\left(q_{g 1}, q_{f}\right)$ | $\rightarrow$ | $q_{g 2}$ |
| $g\left(q_{b}, q_{c}\right)$ | $\rightarrow$ | $q_{g 2}$ |
| $C_{1}$ | $\rightarrow$ | $q_{a}$ |
| $C_{3}\left(q_{c}\right)$ | $\rightarrow$ | $q_{f}$ |
| $C_{2}\left(q_{b}\right)$ | $\rightarrow$ | $q_{g 1}$ |



## Dedicated completion algorithm (2)

| $a$ | $\rightarrow$ | $q_{a}$ |
| :--- | :--- | :--- |
| $b$ | $\rightarrow$ | $q_{b}$ |
| $c$ | $\rightarrow$ | $q_{c}$ |
| $g\left(q_{a}, q_{b}\right)$ | $\rightarrow$ | $q_{g 1}$ |
| $f\left(q_{c}\right)$ | $\rightarrow$ | $q_{f}$ |
| $g\left(q_{g 1}, q_{f}\right)$ | $\rightarrow$ | $q_{g 2}$ |
| $g\left(q_{b}, q_{c}\right)$ | $\rightarrow$ | $q_{g 2}$ |
| $C_{1}$ | $\rightarrow$ | $q_{a}$ |
| $C_{3}\left(q_{c}\right)$ | $\rightarrow$ | $q_{f}$ |
| $C_{2}\left(q_{b}\right)$ | $\rightarrow$ | $q_{g 1}$ |



## Dedicated completion algorithm (2)

| $a$ | $\rightarrow$ | $q_{a}$ |  |
| :--- | :--- | :--- | :--- |
| $b$ | $\rightarrow$ | $q_{b}$ |  |
| $c$ | $\rightarrow$ | $q_{c}$ |  |
| $g\left(q_{a}, q_{b}\right)$ | $\rightarrow$ | $q_{g 1}$ |  |
| $f\left(q_{c}\right)$ | $\rightarrow$ | $q_{f}$ |  |
| $g\left(q_{g 1}, q_{f}\right)$ | $\rightarrow$ | $q_{g 2}$ |  |
| $g\left(q_{b}, q_{c}\right)$ | $\rightarrow$ | $q_{g 2}$ |  |
| $C_{1}$ | $\rightarrow$ | $q_{a}$ | $q_{a}$ |
| $C_{3}\left(q_{c}\right)$ | $\rightarrow$ | $q_{f}$ | $q_{b}$ |
| $C_{2}\left(q_{b}\right)$ | $\rightarrow$ | $q_{g 1}$ |  |
| $q_{b}$ |  |  |  |

## Dedicated completion algorithm (2)

| $a$ | $\rightarrow$ | $q_{a}$ |
| :--- | :--- | :--- |
| $b$ | $\rightarrow$ | $q_{b}$ |
| $c$ | $\rightarrow$ | $q_{c}$ |
| $g\left(q_{a}, q_{b}\right)$ | $\rightarrow$ | $q_{g 1}$ |
| $f\left(q_{c}\right)$ | $\rightarrow$ | $q_{f}$ |
| $g\left(q_{g 1}, q_{f}\right)$ | $\rightarrow$ | $q_{g 2}$ |
| $g\left(q_{b}, q_{c}\right)$ | $\rightarrow$ | $q_{g 2}$ |
| $C_{1}($ | $\rightarrow$ | $q_{a}$ |
| $C_{3}\left(q_{c}\right)$ | $\rightarrow$ | $q_{f}$ |
| $C_{2}\left(q_{b}\right)$ | $\rightarrow$ | $q_{g 1}$ |

## Dedicated completion algorithm (2)

| $a$ | $\rightarrow$ | $q_{a}$ |
| :--- | :--- | :--- |
| $b$ | $\rightarrow$ | $q_{b}$ |
| $c$ | $\rightarrow$ | $q_{c}$ |
| $g\left(q_{a}, q_{b}\right)$ | $\rightarrow$ | $q_{g 1}$ |
| $f\left(q_{c}\right)$ | $\rightarrow$ | $q_{f}$ |
| $g\left(q_{g 1}, q_{f}\right)$ | $\rightarrow$ | $q_{g 2}$ |
| $g\left(q_{b}, q_{c}\right)$ | $\rightarrow$ | $q_{g 2}$ |
| $C_{1}$ |  |  |
| $C_{3}\left(q_{c}\right)$ | $\rightarrow$ | $q_{a}$ |
| $C_{2}\left(q_{b}\right)$ | $\rightarrow$ | $q_{g 1}$ |



## Dedicated completion algorithm (3)

We want to do a completion step with the rule $g\left(C_{2}(x), C_{3}(y)\right) \rightarrow C_{4}(x, y)$.


The particular form of the rules allows us to replace the substitution calculus by associative pattern-matching

## Dedicated completion algorithm (3)

We want to do a completion step with the rule $g\left(C_{2}(x), C_{3}(y)\right) \rightarrow C_{4}(x, y)$.


The particular form of the rules allows us to replace the substitution calculus by associative pattern-matching

## Dedicated completion algorithm (3)

We want to do a completion step with the rule $g\left(C_{2}(x), C_{3}(y)\right) \rightarrow C_{4}(x, y)$.


The particular form of the rules allows us to replace the substitution calculus by associative pattern-matching

## Dedicated completion algorithm (3)

We want to do a completion step with the rule $g\left(C_{2}(x), C_{3}(y)\right) \rightarrow C_{4}(x, y)$.


The particular form of the rules allows us to replace the substitution calculus by associative pattern-matching

## Dedicated completion algorithm (3)

We want to do a completion step with the rule $g\left(C_{2}(x), C_{3}(y)\right) \rightarrow C_{4}(x, y)$.


The particular form of the rules allows us to replace the substitution calculus by associative pattern-matching

## General schema of the implementation

The Tom language [RTA'07]: Piggybacking Rewriting on top of Java

- Efficient support for algebraic terms (hash-consing),
- Pattern-matching (AU theory, variadic operators),
- Expressive strategy language (à la ELAN, Stratego).


## General schema of the implementation

The Tom language [RTA'07]: Piggybacking Rewriting on top of Java

- Efficient support for algebraic terms (hash-consing),
- Pattern-matching (AU theory, variadic operators),
- Expressive strategy language (à la ELAN, Stratego).

Generator of dedicated completion programs written in Tom


## Experimental results

|  | NSPK <br> protocol | View-Only <br> protocol | Java program <br> (chained lists) |
| :--- | :---: | :---: | :---: |
| TRS size (nb of rules) | 13 | 15 | 303 |
| Timbuk: <br> Time (secs) | 19.7 | 6420 | 37387 |
| Tom: <br> Time (secs) | 5.9 | 150 | 303 |
| Timbuk/Tom | 3 | 40 | 120 |

## Experimental results

|  | NSPK <br> protocol | View-Only <br> protocol | Java program <br> (chained lists) |
| :--- | :---: | :---: | :---: |
| TRS size (nb of rules) | 13 | 15 | 303 |
| Timbuk: <br> Time (secs) | 19.7 | 6420 | 37387 |
| Tom: <br> Time (secs) | 5.9 | 150 | 303 |
| Timbuk/Tom | 3 | 40 | 120 |

In practice, conclusive analyses with Timbuk are also conclusive with Tom

## Conclusion

Main results:

- Definition of a reachability preserving transformation on TRS
- Computations of over-approximations using associative pattern-matching
- Implementation in Tom/Java
- A factor 10 in general, and up to 100 on Java examples

Future work:

- Verification of MIDlets
- A better control of approximations
- Using threads to parallelize the completion procedure


## Conclusion

Main results:

- Definition of a reachability preserving transformation on TRS
- Computations of over-approximations using associative pattern-matching
- Implementation in Tom/Java
- A factor 10 in general, and up to 100 on Java examples

Future work:

- Verification of MIDlets
- A better control of approximations
- Using threads to parallelize the completion procedure

