# A decision procedure for bisimilarity of generalized regular expressions

M. Bonsangue<sup>1,2</sup> G. Caltais<sup>5</sup> E. Goriac<sup>5</sup> D. Lucanu<sup>4</sup> J. Rutten<sup>1,3</sup> A. Silva<sup>1</sup>

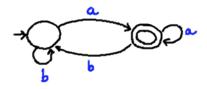
<sup>1</sup>Centrum Wiskunde & Informatica, The Netherlands
 <sup>2</sup>LIACS - Leiden University, The Netherlands
 <sup>3</sup>Radboud Universiteit Nijmegen, The Netherlands
 <sup>4</sup>Faculty of Computer Science - Alexandru Ioan Cuza University, Romania
 <sup>5</sup>School of Computer Science - Reykjavik University, Iceland

SBMF'10, November 2010



#### **Deterministic automata (DA)**

- Widely used model in Computer Science.
- Acceptors of languages

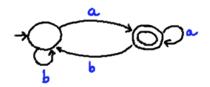


#### Regular expressions

- User-friendly alternative to DA notation.
- Many applications: pattern matching (grep), specification of circuits, . . .

#### **Deterministic automata (DA)**

- Widely used model in Computer Science.
- Acceptors of languages



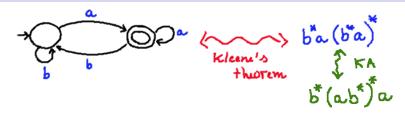
#### Regular expressions

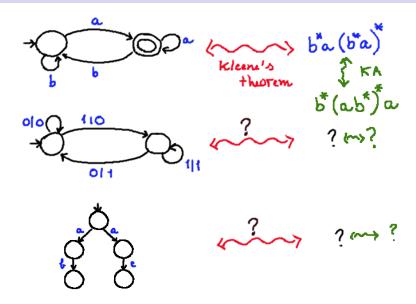
- User-friendly alternative to DA notation.
- Many applications: pattern matching (grep), specification of circuits, . . .

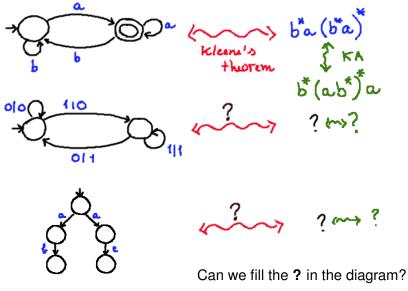
#### Kleene's Theorem

Let  $A \subseteq \Sigma^*$ . The following are equivalent.

- **1** A = L(A), for some finite automaton A.
- 2 A = L(r), for some regular expression r.







#### In previous work ...

#### We presented:

- a generalized notion of regular expressions;
- an analogue of Kleene's theorem;
- and sound and complete axiomatizations with respect to bisimilarity

for a large class of systems (labelled transition systems, Mealy machines, probabilistic automata).

All the above was derived modularly from the type of each system.

Question: Can we automate the reasoning on equivalence of expressions, also in a modular way?

#### In previous work ...

#### We presented:

- a generalized notion of regular expressions;
- an analogue of Kleene's theorem;
- and sound and complete axiomatizations with respect to bisimilarity

for a large class of systems (labelled transition systems, Mealy machines, probabilistic automata).

All the above was derived modularly from the type of each system.

Question: Can we automate the reasoning on equivalence of expressions, also in a modular way?

#### In previous work ...

#### We presented:

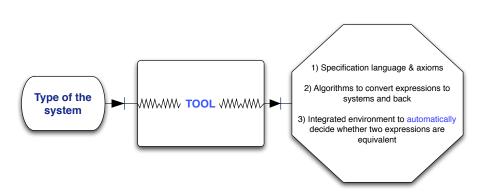
- a generalized notion of regular expressions;
- an analogue of Kleene's theorem;
- and sound and complete axiomatizations with respect to bisimilarity

for a large class of systems (labelled transition systems, Mealy machines, probabilistic automata).

All the above was derived modularly from the type of each system.

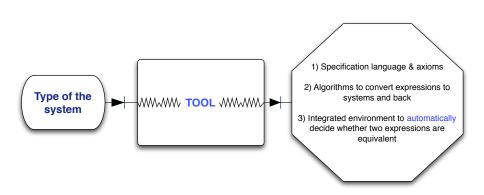
Question: Can we automate the reasoning on equivalence of expressions, also in a modular way?

#### The ultimate goal...



In this talk, we will be focusing on 1) and 3).

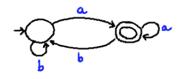
#### The ultimate goal...



In this talk, we will be focusing on 1) and 3).

#### Outline

- Generalized regular expressions
- Equivalence of expressions
- Snapshot of the tool

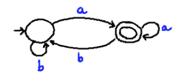




$$(S,\delta:S\to 2\times S^A)$$

$$(S, \delta : S \to (B \times S)^A)$$

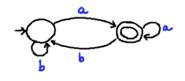
$$(S, \delta: S \to 1 + (\mathcal{P}S)^A)$$



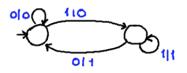
$$(S, \delta: S \to 2 \times S^A)$$

$$(S, \delta : S \to (B \times S)^A)$$

$$(S, \delta: S \to 1 + (\mathcal{P}S)^A)$$

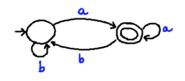


$$(S, \delta: S \rightarrow 2 \times S^A)$$

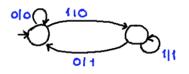


$$(S, \delta: S \rightarrow (B \times S)^A)$$

$$(S, \delta: S \rightarrow 1 + (\mathcal{P}S)^A)$$

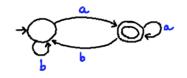


$$(S, \delta: S \rightarrow 2 \times S^A)$$

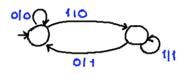


$$(S, \delta: S \to (B \times S)^A)$$

$$(S, \delta: S \rightarrow 1 + (\mathcal{P}S)^A)$$

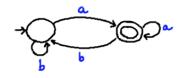


$$(S, \delta: S \rightarrow 2 \times S^{A})$$

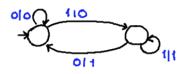


$$(S, \delta: S \to (B \times S)^A)$$

$$(S, \delta: S \rightarrow 1 + (\mathcal{P}S)^A)$$



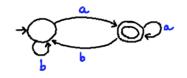
$$(S, \delta: S \rightarrow 2 \times S^{A})$$



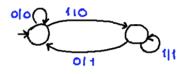
$$(S, \delta: S \to (B \times S)^A)$$

$$(S, \delta: S \rightarrow \mathbf{1} + (\mathcal{P}S)^{\mathbf{A}})$$

 $(S, \delta : S \rightarrow GS)$ 



$$(S, \delta: S \rightarrow 2 \times S^A)$$



$$(S, \delta: S \to (B \times S)^A)$$

$$(S, \delta: S \rightarrow \mathbf{1} + (\mathcal{P}S)^{\mathbf{A}})$$

 $(S, \delta: S \rightarrow \S S)$   $\S$ -coalgebras

## Coalgebras

- Generalizations of deterministic automata
- Set of states S and a transition function t : S → GS where G encodes the type of the system:

$$\mathfrak{G}::= Id \mid B \mid \mathfrak{G} \times \mathfrak{G} \mid \mathfrak{G} + \mathfrak{G} \mid \mathfrak{G}^{A} \mid \mathfrak{P}\mathfrak{G}$$

P finite

#### Examples

• 
$$9 = 2 \times Id^A$$

• 
$$G = (B \times Id)^A$$

• 
$$9 = 1 + (PId)^A$$

Deterministic automata

Mealy machines

Mealy machines

LTS (with explicit termination)

## Coalgebras

- Generalizations of deterministic automata
- Set of states S and a transition function t : S → GS where G encodes the type of the system:

$$\mathfrak{G}::= Id \mid B \mid \mathfrak{G} \times \mathfrak{G} \mid \mathfrak{G} + \mathfrak{G} \mid \mathfrak{G}^{A} \mid \mathfrak{P}\mathfrak{G}$$

 $\mathcal{P}$  finite

#### Examples

• 
$$9 = 2 \times Id^A$$

• 
$$\mathfrak{G} = (B \times Id)^A$$

• 
$$\mathfrak{G} = 1 + (\mathfrak{P}Id)^A$$

...

Deterministic automata

Mealy machines

LTS (with explicit termination)

## The power of 9

#### The functor 9 determines:

- notion of observational equivalence (coalg. bisimulation)
- behaviour (final coalgebra)
- set of expressions describing finite systems
- axioms to prove bisimulation equivalence of expressions

## The power of 9

#### The functor 9 determines:

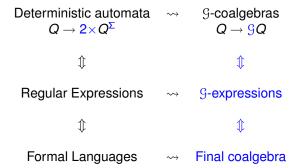
- notion of observational equivalence (coalg. bisimulation)
- behaviour (final coalgebra)
- set of expressions describing finite systems
- axioms to prove bisimulation equivalence of expressions

## The power of 9

#### The functor 9 determines:

- notion of observational equivalence (coalg. bisimulation)
- behaviour (final coalgebra)
- set of expressions describing finite systems
- axioms to prove bisimulation equivalence of expressions
- 1 + 2 are standard universal coalgebra; 1 + 1 are [BRS10]

## In a nutshell — beyond deterministic automata



## **9-expressions**

$$E ::= \underline{\emptyset} \mid \epsilon \mid E \cdot E \mid E + E \mid E^*$$

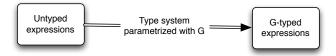
$$E_{\mathfrak{G}}$$
 ::= ?

## **9-expressions**

$$E \quad ::= \quad \underline{\emptyset} \mid \epsilon \mid E \cdot E \mid E + E \mid E^*$$

## $E_{\mathfrak{G}}$ ::= ?

#### How do we define $E_{\mathfrak{G}}$ ?



#### Deterministic automata expressions – $9 = 2 \times Id^A$

$$\varepsilon ::= \underbrace{\emptyset \mid \varepsilon \oplus \varepsilon \mid \mu \mathbf{X}.\gamma}_{\mathfrak{G}} \mid$$

#### Deterministic automata expressions – $9 = 2 \times Id^A$

$$\varepsilon ::= \underbrace{\emptyset \mid \varepsilon \oplus \varepsilon \mid \mu \mathbf{X}. \gamma}_{\mathfrak{S}} \mid \underbrace{\qquad \qquad }_{\times}$$

#### Deterministic automata expressions – $9 = 2 \times Id^A$

$$\varepsilon ::= \underbrace{\emptyset \mid \varepsilon \oplus \varepsilon \mid \mu x. \gamma}_{\mathcal{G}} \mid \underbrace{\frac{1}{2} \mid \underbrace{0}_{2} \mid \underbrace{a(\varepsilon)}_{Id^{A}}}_{\times}$$

#### Deterministic automata expressions – $\mathcal{G} = \mathbf{2} \times Id^A$

$$\varepsilon ::= \underbrace{\emptyset \mid \varepsilon \oplus \varepsilon \mid \mu x. \gamma}_{S} \mid \underbrace{\frac{1}{2} \mid \underbrace{0}_{2} \mid \underbrace{a(\varepsilon)}_{Id^{A}}}_{\times}$$

## LTS expressions – $\mathfrak{G} = 1 + (\mathfrak{P}Id)^A$

$$\varepsilon$$
 ::=  $\underline{\emptyset} \mid \varepsilon \oplus \varepsilon \mid \mu \mathbf{x}.\gamma$ 

#### Deterministic automata expressions – $9 = 2 \times Id^A$

$$\varepsilon ::= \underbrace{\emptyset \mid \varepsilon \oplus \varepsilon \mid \mu x. \gamma}_{\mathcal{G}} \mid \underbrace{\frac{1}{2} \mid \underbrace{0}_{2} \mid \underbrace{a(\varepsilon)}_{Id^{A}}}_{\times}$$

#### LTS expressions – $\mathfrak{G} = 1 + (\mathfrak{P}Id)^A$

$$\varepsilon ::= \underline{\emptyset} \mid \varepsilon \oplus \varepsilon \mid \mu \mathbf{X}.\gamma \mid \underbrace{\checkmark}_{1} \mid \underbrace{\partial}_{(\mathfrak{P}Id)^{A}} \mid \underbrace{\mathbf{a}.\varepsilon}_{(\mathfrak{P}Id)^{A}}$$

The set of G-expressions has a coalgebraic structure given by

$$\delta_{\mathcal{G}} : \mathsf{Exp}_{\mathcal{G}} \to \mathcal{G}(\mathsf{Exp}_{\mathcal{G}})$$

 $\delta g \dots$ 

- ... provides an operational semantics for the set of expressions
- ... defines the dynamics of the system
- ... is used for observing the behaviour of the system

The set of  $\mathcal{G}$ -expressions has a coalgebraic structure given by

$$\delta_{\mathcal{G}} : \mathsf{Exp}_{\mathcal{G}} \to \mathcal{G}(\mathsf{Exp}_{\mathcal{G}})$$

- $\delta g \dots$ 
  - ...provides an operational semantics for the set of expressions
  - ... defines the dynamics of the system
  - ...is used for observing the behaviour of the system

The set of G-expressions has a coalgebraic structure given by

$$\delta_{\mathcal{G}} : \mathsf{Exp}_{\mathcal{G}} \to \mathcal{G}(\mathsf{Exp}_{\mathcal{G}})$$

 $\delta g \dots$ 

- ... provides an operational semantics for the set of expressions
- ... defines the dynamics of the system
- ...is used for observing the behaviour of the system

The set of  $\mathcal{G}$ -expressions has a coalgebraic structure given by

$$\delta_{\mathcal{G}} : \mathsf{Exp}_{\mathcal{G}} \to \mathcal{G}(\mathsf{Exp}_{\mathcal{G}})$$

 $\delta g \dots$ 

- ...provides an operational semantics for the set of expressions
- ...defines the dynamics of the system
- ...is used for observing the behaviour of the system

## The goal: proving equivalence

- Automatically proving equivalence (bisimilarity) of expressions;
- Tool: Circ
- Meta-language implemented in Maude
- Input: Algebraic specification + coalgebraic structure (dynamics)
- Engine: circular coinduction constructing bisimulation
- Output: Yes / No / ?

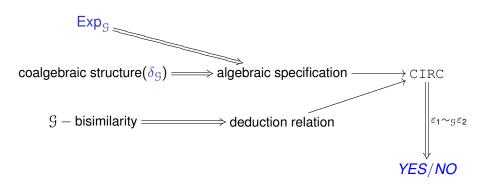
## The goal: proving equivalence

- Automatically proving equivalence (bisimilarity) of expressions;
- Tool: Circ
- Meta-language implemented in Maude
- Input: Algebraic specification + coalgebraic structure (dynamics)
- Engine: circular coinduction constructing bisimulation
- Output: Yes / No / ?

## The goal: proving equivalence

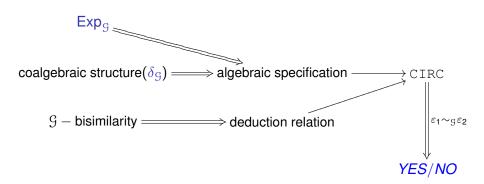
- Automatically proving equivalence (bisimilarity) of expressions;
- Tool: Circ
- Meta-language implemented in Maude
- Input: Algebraic specification + coalgebraic structure (dynamics)
- Engine: circular coinduction constructing bisimulation
- Output: Yes / No / ?

# Our challenge



How to guarantee that the prover always says YES/NO?

# Our challenge

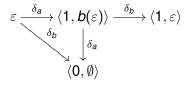


How to guarantee that the prover always says YES/NO?

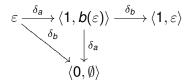
$$\varepsilon = \mu x.a(b(x)) \oplus 1$$

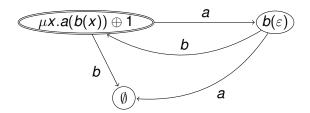
$$\varepsilon \xrightarrow{\delta_a} \langle \mathbf{1}, \mathbf{b}(\varepsilon) \rangle \xrightarrow{\delta_b} \langle \mathbf{1}, \varepsilon \rangle$$

$$\varepsilon = \mu x.a(b(x)) \oplus 1$$



$$\varepsilon = \mu x.a(b(x)) \oplus 1$$





But ...

$$\varepsilon = \mu x.a(x \oplus x)$$

But ...

$$\varepsilon = \mu x.a(x \oplus x)$$

$$\varepsilon \overset{\delta}{\longmapsto} \langle \mathbf{0}, \varepsilon \!\oplus\! \varepsilon \rangle$$

But . . .

$$\varepsilon = \mu x.a(x \oplus x)$$

$$\varepsilon \stackrel{\delta}{\longmapsto} \langle 0, \varepsilon \oplus \varepsilon \rangle \stackrel{\delta}{\longmapsto} \langle 0, (\varepsilon \oplus \varepsilon) \oplus (\varepsilon \oplus \varepsilon) \rangle \stackrel{\delta}{\longmapsto} \langle 0, (\varepsilon \oplus \varepsilon) \oplus (\varepsilon \oplus \varepsilon) \oplus (\varepsilon \oplus \varepsilon) \rangle \dots$$

But ...

$$\varepsilon = \mu x.a(x \oplus x)$$

$$\varepsilon \stackrel{\delta}{\longmapsto} \langle 0, \varepsilon \oplus \varepsilon \rangle \stackrel{\delta}{\longmapsto} \langle 0, (\varepsilon \oplus \varepsilon) \oplus (\varepsilon \oplus \varepsilon) \rangle \stackrel{\delta}{\longmapsto} \langle 0, (\varepsilon \oplus \varepsilon) \oplus (\varepsilon \oplus \varepsilon) \oplus (\varepsilon \oplus \varepsilon) \rangle \dots$$

We need ACI!

But ...

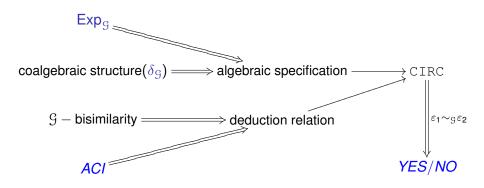
$$\varepsilon = \mu x.a(x \oplus x)$$

$$\varepsilon \overset{\delta}{\longmapsto} \langle 0, \varepsilon \oplus \varepsilon \rangle \overset{\delta}{\longmapsto} \langle 0, (\varepsilon \oplus \varepsilon) \oplus (\varepsilon \oplus \varepsilon) \rangle \overset{\delta}{\longmapsto} \langle 0, (\varepsilon \oplus \varepsilon) \oplus (\varepsilon \oplus \varepsilon) \oplus (\varepsilon \oplus \varepsilon) \rangle \dots$$

We need ACI!

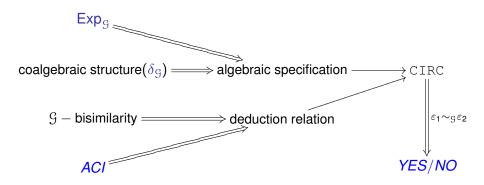
$$(\mu x.a(x \oplus x))$$
 a

# A decision procedure for bisimilarity



- Adding equations is possible because Circ is an extension of
- With ACI, the bisimulation game is decidable!

# A decision procedure for bisimilarity

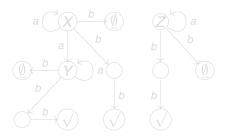


- Adding equations is possible because Circ is an extension of Maude
- With ACI, the bisimulation game is decidable!

#### Algebra meets coalgebra

## Example

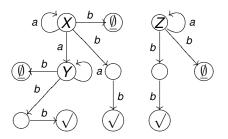
$$\begin{array}{lll} \varepsilon_{\mathcal{X}} & = & \mu x.a.x \oplus b.\underline{\emptyset} \oplus a.(\mu y.a.y \oplus b.\underline{\emptyset} \oplus b.b.\sqrt{)} \oplus b.b.\sqrt{} \\ \varepsilon_{\mathcal{Z}} & = & \mu z.a.z \oplus b.\underline{\emptyset} \oplus b.b.\sqrt{} \end{array}$$



**Question** Are  $\varepsilon_X$  and  $\varepsilon_Z$  equivalent?

## Example

$$\begin{array}{lll} \varepsilon_{\mathcal{X}} & = & \mu x.a.x \oplus b.\underline{\emptyset} \oplus a.(\mu y.a.y \oplus b.\underline{\emptyset} \oplus b.b.\sqrt{)} \oplus b.b.\sqrt{} \\ \varepsilon_{\mathcal{Z}} & = & \mu z.a.z \oplus b.\underline{\emptyset} \oplus b.b.\sqrt{} \end{array}$$



**Question** Are  $\varepsilon_X$  and  $\varepsilon_Z$  equivalent?

#### Maude> in protofunctorizer.maude

Proof succeeded.

$$\mathfrak{G} = B + \mathfrak{P}(Id)^A, A = \{a, b\}, B = \{\sqrt{a}\}$$



```
Maude> in protofunctorizer.maude
(alph A is a b endalph)
```

$$\mathfrak{G} = B + \mathfrak{P}(Id)^A, A = \{a, b\}, B = \{\sqrt{a}\}$$



```
Maude> in protofunctorizer.maude
(alph A is a b endalph)
(alph B is tick endalph)
```

$$\mathfrak{G} = B + \mathfrak{P}(Id)^A, A = \{a, b\}, B = \{\sqrt{a}\}$$



```
Maude> in protofunctorizer.maude
(alph A is a b endalph)
(alph B is tick endalph)
(functor B + P(Id)^A.)
```

$$\mathfrak{G} = B + \mathfrak{P}(Id)^A, A = \{a, b\}, B = \{\sqrt{a}\}$$



```
Maude> in protofunctorizer.maude
(alph A is a b endalph)
(alph B is tick endalph)
(functor B + P(Id)^A.)
(E1 = mu X. ...)
(E2 = mu Z. ...)
(set goal E1 = E2 .)
```

$$\mathfrak{G} = B + \mathfrak{P}(Id)^A, A = \{a, b\}, B = \{\sqrt{a}\}$$



```
Maude> in protofunctorizer.maude
(alph A is a b endalph)
(alph B is tick endalph)
(functor B + P(Id)^A.)
(E1 = mu X. ...)
(E2 = mu Z. ...)
(set goal E1 = E2.)
(generate coalgebra .)
```

$$\mathfrak{G} = B + \mathfrak{P}(Id)^A, A = \{a, b\}, B = \{\sqrt{a}\}$$



```
Maude> in protofunctorizer.maude
(alph A is a b endalph)
(alph B is tick endalph)
(functor B + P(Id)^A.)
(E1 = mu X. ...)
(E2 = mu Z. ...)
(set goal E1 = E2.)
(generate coalgebra .)
296 lines of CIRC specification!
```

$$\mathfrak{G} = \mathbf{B} + \mathfrak{P}(\mathbf{Id})^{\mathbf{A}}, \mathbf{A} = \{\mathbf{a}, \mathbf{b}\}, \mathbf{B} = \{\sqrt{a}\}$$



```
Maude> in protofunctorizer.maude
(alph A is a b endalph)
(alph B is tick endalph)
(functor B + P(Id)^A.)
(E1 = mu X. ...)
(E2 = mu Z. ...)
(set goal E1 = E2.)
(generate coalgebra .)
296 lines of CIRC specification!
Maude> in circ.maude Xand7.maude
```

$$\mathfrak{G} = \mathbf{B} + \mathfrak{P}(\mathbf{Id})^{\mathbf{A}}, \mathbf{A} = \{\mathbf{a}, \mathbf{b}\}, \mathbf{B} = \{\sqrt{3}\}$$



```
Maude> in protofunctorizer.maude
(alph A is a b endalph)
(alph B is tick endalph)
(functor B + P(Id)^A.)
(E1 = mu X. ...)
(E2 = mu Z. ...)
(set goal E1 = E2.)
(generate coalgebra .)
296 lines of CIRC specification!
Maude> in circ.maude Xand7.maude
(coinduction .)
```

$$g = B + \mathcal{P}(Id)^A, A = \{a, b\}, B = \{\sqrt{a}\}$$



```
Maude> in protofunctorizer.maude
(alph A is a b endalph)
(alph B is tick endalph)
(functor B + P(Id)^A.)
(E1 = mu X. ...)
(E2 = mu Z. ...)
(set goal E1 = E2 .)
(generate coalgebra .)
296 lines of CIRC specification!
Maude> in circ.maude Xand7.maude
(coinduction .)
```

#### Proof succeeded.

$$\mathfrak{G} = \mathbf{B} + \mathfrak{P}(\mathbf{Id})^{\mathbf{A}}, \mathbf{A} = \{\mathbf{a}, \mathbf{b}\}, \mathbf{B} = \{\sqrt{3}\}$$



#### Conclusions and Future work

#### Conclusions

- Generic framework to uniformly derive language and axioms for a large class of systems
- Generalization of Kleene theorem and Kleene algebra, parametric on the functor.
- Automation in Circ: decision procedure for equivalence of expressions.

#### Future work

- Making the tool more user friendly;
- Extending the class of systems to include quantitative systems (probabilistic automata, etc)
- Apply it to a serious case study (circuit design, compiler optimization, ...)

#### Conclusions and Future work

#### Conclusions

- Generic framework to uniformly derive language and axioms for a large class of systems
- Generalization of Kleene theorem and Kleene algebra, parametric on the functor.
- Automation in Circ: decision procedure for equivalence of expressions.

#### Future work

- Making the tool more user friendly;
- Extending the class of systems to include quantitative systems (probabilistic automata, etc)
- Apply it to a serious case study (circuit design, compiler optimization, ...)

## Thank you!

- for more details see:
  - "Non-deterministic Kleene coalgebras"
     A.Silva, M. Bonsangue, J. Rutten
  - "Circular coinduction A proof theoretical foundation"
     G. Roşu, D. Lucanu
- QUESTIONS?