A Coinductive Axiomatisation of Regular Expressions under Bisimulation

Clemens Grabmayer

Dept. of Computer Science
Vrije Universiteit Amsterdam

CMCS 2006 Short Contributions
March 27, 2005

Overview

- The process interpretation of regular expressions.
- Milner's 3 questions (1984) concerning the process interpretation; i.p. the (still open) axiomatisation question.
- Relationship of the proc. int. with Antimirov's partial derivatives.
- A coinductively motivated proof system for regular expressions under bisimulation.
- A partial answer to the axiomatisation question.

Language Interpretation L

$\mathbf{0} \stackrel{L}{\longmapsto}$ empty set $\emptyset$

$\mathbf{1} \stackrel{L}{\longmapsto} \{\lambda\}$ (λ the empty word)

$a \stackrel{L}{\longmapsto} \{a\}$

$e + f \stackrel{L}{\longmapsto}$ union of $L(e)$ and $L(f)$

$e \cdot f \stackrel{L}{\longmapsto}$ element-wise concatenation of $L(e)$ and $L(f)$

$e^* \stackrel{L}{\longmapsto}$ set of “words over of $L(e)$”

Process Interpretation P

$\mathbf{0} \quad \stackrel{P}{\longmapsto} \quad$ deadlock δ

$\mathbf{1} \quad \stackrel{P}{\longmapsto} \quad$ empty process ϵ

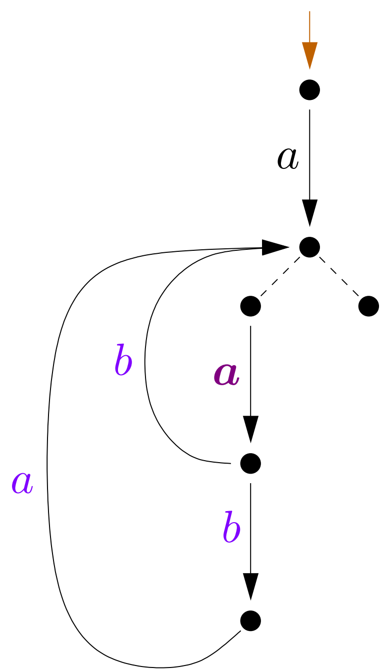
$a \quad \stackrel{P}{\longmapsto} \quad$ atomic action a

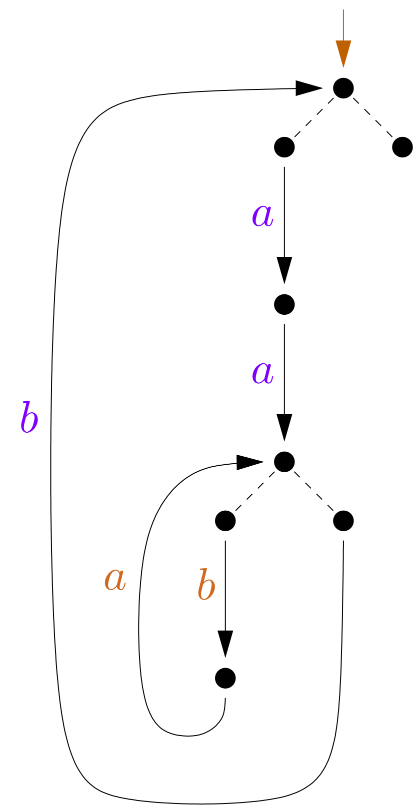
$e + f \quad \stackrel{P}{\longmapsto} \quad$ alternative composition between $P(e)$ and $P(f)$

$e \cdot f \quad \stackrel{P}{\longmapsto} \quad$ sequential composition of $P(e)$ and $P(f)$

$e^* \quad \stackrel{P}{\longmapsto} \quad$ unbounded iteration of $P(e)$

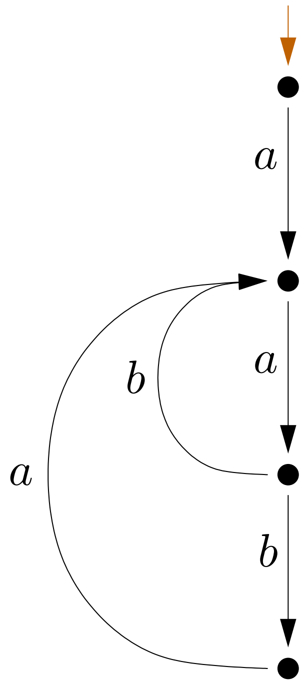
Process Interpretation P



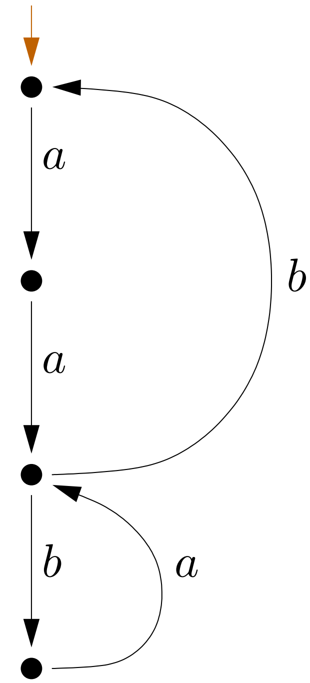


$$P\Big(\, a\big(\boldsymbol{a}(b+ba)\big)^{*}.0\,\Big) \qquad\qquad P\Big(\, \big(aa(ba)^{*}b\big)^{*}.0\,\Big)$$

Process Interpretation $\boldsymbol{P}$

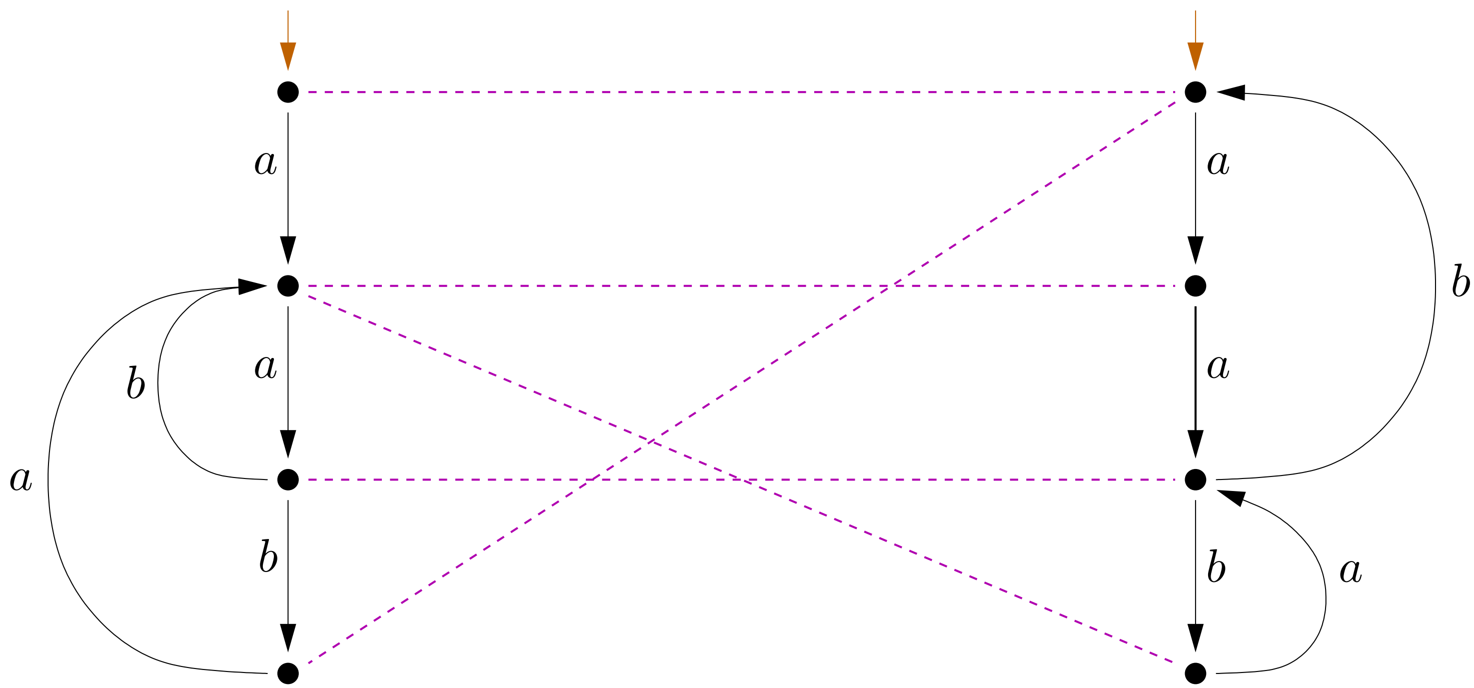


$$\boldsymbol{P}\big(\, a(a(b+ba))^*.0 \,\big)$$



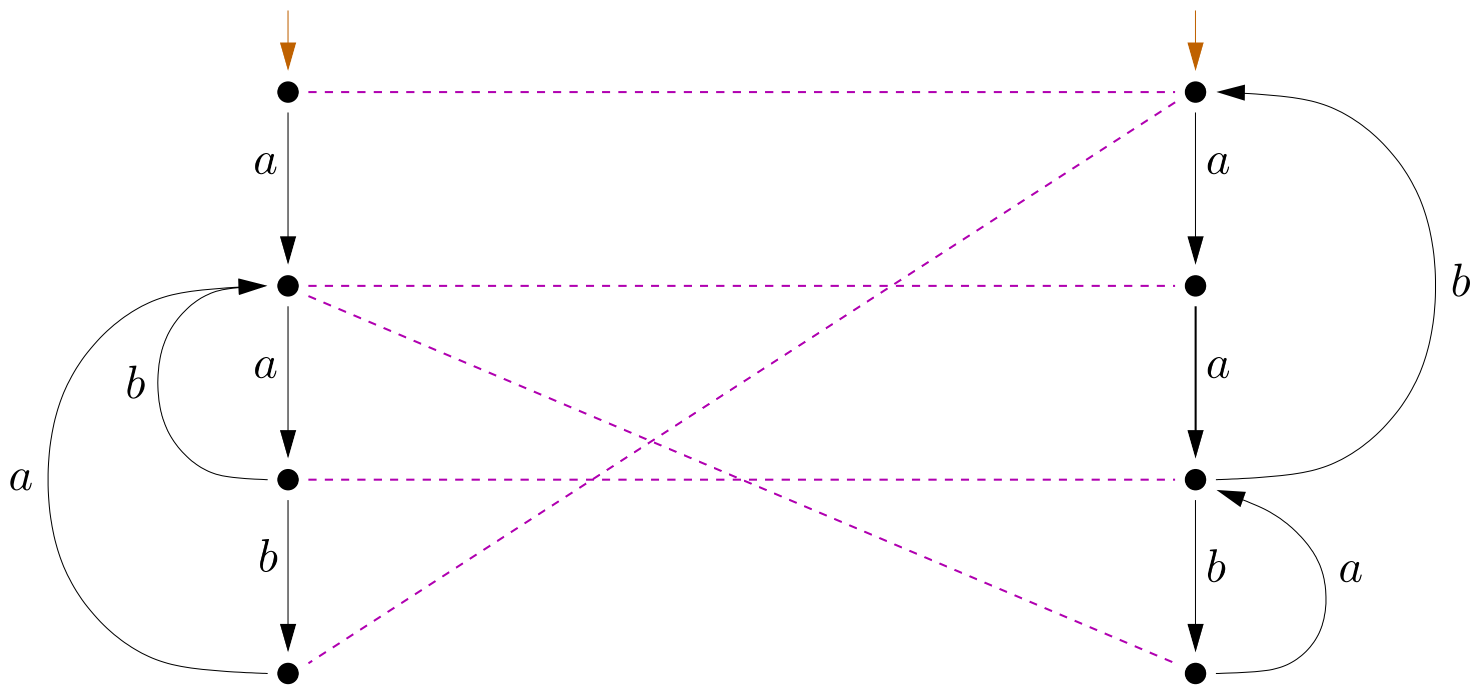
$$\boldsymbol{P}\big(\, (aa(ba)^*a)^*.0 \,\big)$$

Regular Expressions under Bisimulation



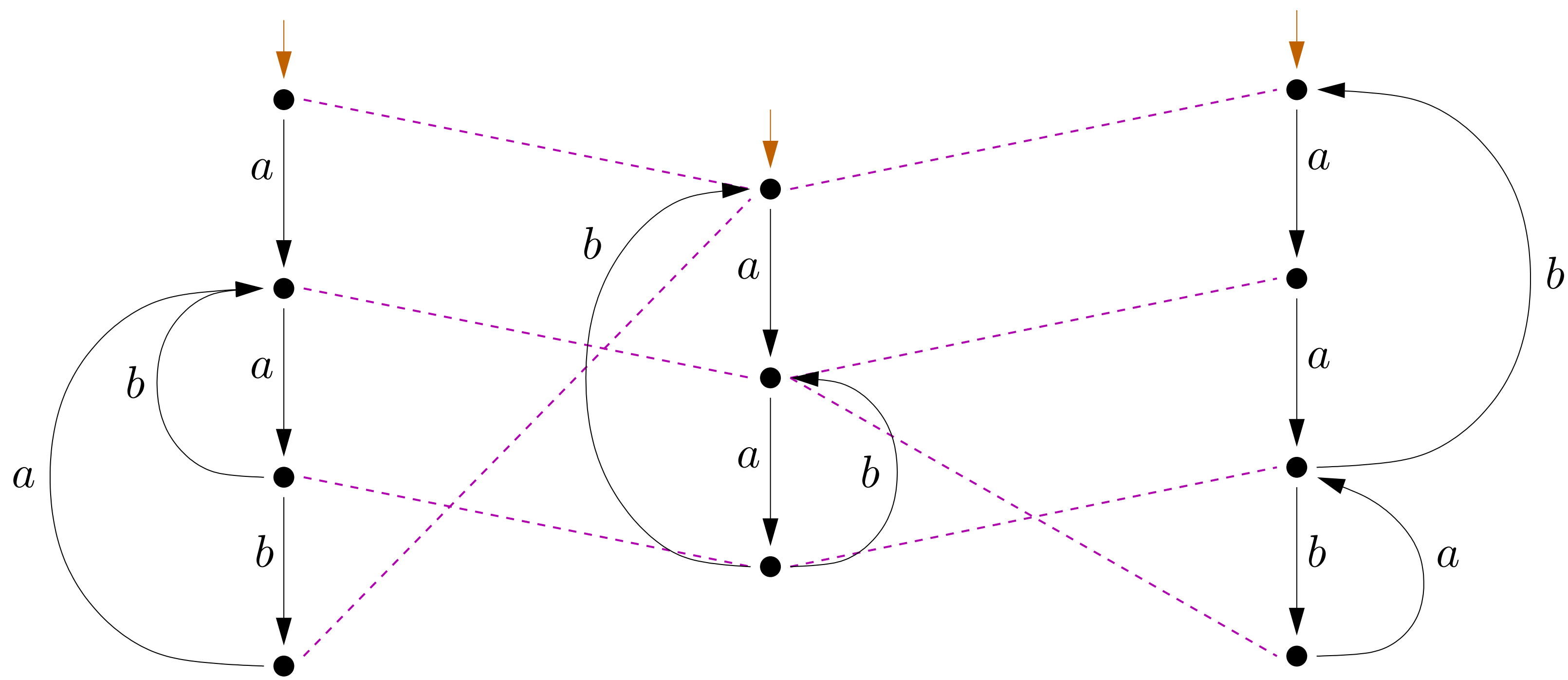
$$\boldsymbol{P}\big(\, a(a(b+ba))^*.0 \,\big) \quad \Leftrightarrow \quad \boldsymbol{P}\big(\, (aa(ba)^*a)^*.0 \,\big)$$

Regular Expressions under Bisimulation



$$\big(\, a(a(b + ba))^*.0 \qquad \underline{\leftrightarrow}_{\boldsymbol{P}} \qquad (aa(ba)^*a)^*.0$$

Regular Expressions under Bisimulation

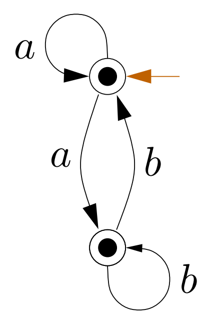


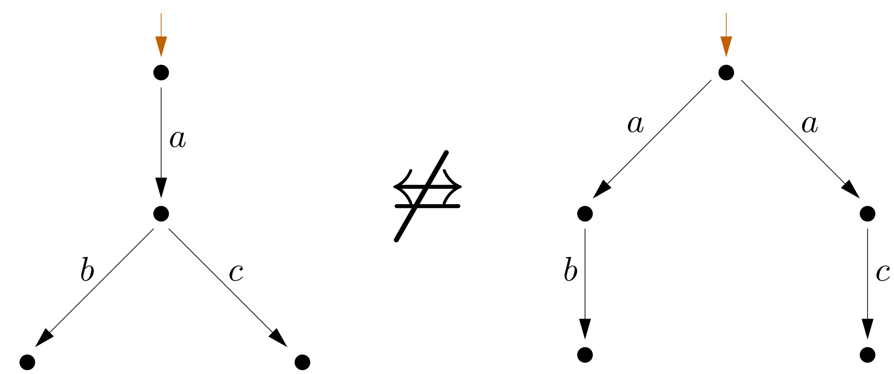
"*Two-exit iteration*"

$$\notin im(\boldsymbol{P})$$

Properties of P

- Not every finite-state process is the process interpretation $P(e)$ of a regular expression e.
- What is more: not every finite-state process is bisimilar to the process interpretation $P(e)$ of a regular expression e.
- Fewer identities hold w.r.t. $\Leftrightarrow_P$ than w.r.t. $=_L$.





$$a.(b+c) \quad \not\Leftrightarrow_P \quad a.b + a.c$$

The Axiom System REG for $=_L$ (Salomaa's axiomatisation F_1 reversed)

Axioms :

(B1)	$x + (y + z) = (x + y) + z$	(B7)	$x.1 = x$
(B2)	$(x.y).z = x.(y.z)$	(B8)	$x.0 = 0$
(B3)	$x + y = y + x$	(B9)	$x + 0 = x$
(B4)	$(x + y).z = x.z + y.z$	(B10)	$x^* = 1 + x.x^*$
(B5)	$x.(y + z) = x.y + x.z$	(B11)	$x^* = (1 + x)^*$
(B6)	$x + x = x$		

Inference rules : equational logic *plus*

$$\frac{e = f.e + g}{e = f^*.g}\,\mathrm{FIX}\ (\text{if } \lambda \notin \boldsymbol{L}(f))$$

Sound and Unsound Axioms of REG w.r.t. $\leftrightarrow_P$

(B1)	$x + (y + z) = (x + y) + z$	(B7)	$x.1 = x$
(B2)	$(x.y).z = x.(y.z)$	(B8)	$x.0 = 0$
(B3)	$x + y = y + x$	(B9)	$x + 0 = x$
(B4)	$(x + y).z = x.z + y.z$	(B10)	$x^* = 1 + x.x^*$
(B5)	$x.(y + z) = x.y + x.z$	(B11)	$x^* = (1 + x)^*$
(B6)	$x + x = x$		

Also sound are:

$$0.x = 0 \qquad\qquad \frac{e = f.e + g}{e = f^*.g}\,\mathsf{FIX}\ (\text{if } \lambda \notin L(f))$$

Milner's Adaptation for $\underline{\leftrightarrow}_{\boldsymbol{P}}$: $\mathbf{BPA}^{*}_{0,1}$ $+$1-$\mathbf{RSP}^{*}_{0,1}$

Axioms :

(B1)	$x + (y + z) = (x + y) + z$	(B7)	$x.1 = x$
(B2)	$(x.y).z = x.(y.z)$	(B8)′	$0.x = 0$
(B3)	$x + y = y + x$	(B9)	$x + 0 = x$
(B4)	$(x + y).z = x.z + y.z$	(B10)	$x^{*} = 1 + x.x^{*}$
		(B11)	$x^{*} = (1 + x)^{*}$
(B6)	$x + x = x$		

Inference rules : equational logic *plus*

$$\frac{e = f.e + g}{e = f^{*}.g}\ \text{1-RSP}^{*}_{0,1}\ (\text{if } \lambda \notin \boldsymbol{L}(f))$$

Milner's Questions (1984)

(1) *Is a variant of Salomaa's axiomatisation for language equality complete for* $\leftrightarrow_{\boldsymbol{P}}$ *?*

– To my knowledge: Yet unsolved. (But: partial & related results.)

(2) *What structural property characterises the finite-state processes that are bisimilar to processes in the image of* $\boldsymbol{P}$ *?*

– Definiability by "well-behaved" specifications (Baeten, Corradini, 2005); this is decidable (Baeten, Corradini, and C.G., 2005).

(3) *Does "minimal star height" over single-letter alphabets define a hierarchy modulo* $\leftrightarrow_{\boldsymbol{P}}$*?*

– Yes! (Hirshfeld and Moller, 1999).

Antimirov and Brzozowki Derivatives

Brzozowski derivative (1963) Antimirov's partial derivatives (1995)

$$(\cdot)_{\cdot} : \mathcal{R}(\Sigma) \times \Sigma \to \mathcal{R}(\Sigma) \qquad \partial : \mathcal{R}(\Sigma) \times \Sigma \to \mathcal{P}_{\mathsf{f}}(\mathcal{R}(\Sigma))$$

$$\langle e, a \rangle \mapsto e_a \qquad \langle e, a \rangle \mapsto \partial_a(e)$$

– Brzozowski der's mimic language derivatives on a synatactic level:

$$\boldsymbol{L}(e_a) \;=\; (\boldsymbol{L}(e))_a \; (=_{\mathsf{def}} \{v \mid a.v \in \boldsymbol{L}(e)\})\,.$$

– Partial derivatives are mathematically motivated refinements.

– Both defined syntactically by induction on the size of reg. expr's.

– Relationship: For all $e \in \mathcal{R}(\Sigma)$, $e_a \equiv_{\mathbf{ACI}} \sum_{e' \in \partial_a(e)} e'$.

– *Every regular expression has only finitely many Brzozowski (Antimirov) derivatives.*

The Coalgebra Induced by Partial Derivatives

Antimirov's partial derivatives induce an F-coalgebra $(\mathcal{R}(\Sigma), \langle o, t\rangle)$, for the functor $F(X) = 2 \times \mathcal{P}_{\mathsf{f}}(\Sigma \times X)$, by:

$$\langle o, t\rangle : \mathcal{R}(\Sigma) \longmapsto 2 \times \mathcal{P}_{\mathsf{f}}(\Sigma \times \mathcal{R}(\Sigma)), \text{ where}$$

$$\begin{aligned} o :\ & \mathcal{R}(\Sigma) \longrightarrow 2 \\ & e \longmapsto o(e) =_{\mathsf{def}} \begin{cases} 0 & \ldots \boldsymbol{P}(e)\not\downarrow \quad (\lambda \notin \boldsymbol{L}(e)) \\ 1 & \ldots \boldsymbol{P}(e)\downarrow \quad (\lambda \in \boldsymbol{L}(e)) \end{cases} \end{aligned}$$

$$\begin{aligned} t :\ & \mathcal{R}(\Sigma) \longrightarrow \mathcal{P}_{\mathsf{f}}(\Sigma \times \mathcal{R}(\Sigma)) \\ & e \longmapsto t(e) =_{\mathsf{def}} \left\{ \langle a, e'\rangle \mid a \in \Sigma,\ e' \in \partial_a e \right\} . \end{aligned}$$

$\sim$: bisimilarity on this coalgebra;

$e \sim_{\mathsf{fin}} f$: there is a finite bisimulation between e and f.

Relationship with the Process Interpretation

Lemma. *For all* $e, f \in \mathcal{R}(\Sigma)$*:* $\left[\, \boldsymbol{P}(e) \xrightarrow{a} \boldsymbol{P}(f) \iff f \in \partial_a(e) \,\right]$.

Lemma. *For all* $e, f \in \mathcal{R}(\Sigma)$*:*

$$e \leftrightarrows_{\boldsymbol{P}} f \iff e \sim f \;\; in \; (\mathcal{R}(\Sigma), \langle o, t \rangle) \; .$$

As a refinement we get a *finitary coinduction principle* (*finite bisimulation principle*).

Theorem. *For all* $e, f \in \mathcal{R}(\Sigma)$*:*

$$e \leftrightarrows_{\boldsymbol{P}} f \iff e \sim_{\mathit{fin}} f \;\; in \; (\mathcal{R}(\Sigma), \langle o, t \rangle) \; .$$

The Proof System c-BPA$^*_{0,1}$

Inference rule in **c-BPA$^*_{0,1}$** : (Given $\Sigma = \{a_1, \ldots, a_n\}$).

$$\dfrac{\ldots \quad \begin{array}{c}[e = f]^{\boldsymbol{u}} \\ \mathcal{D}_1^{(i)} \\ e_1^{(i)} = f_1^{(i)}\end{array} \quad \ldots \quad \begin{array}{c}[e = f]^{\boldsymbol{u}} \\ \mathcal{D}_{m_i}^{(i)} \\ e_{m_i}^{(i)} = f_{m_i}^{(i)}\end{array} \quad \ldots}{e = f}\ \text{COMP/FIX}, \boldsymbol{u} \text{ (if (*))}$$

where (*) demands:

– $o(e) = o(f)$ holds, and

– $\partial_{a_i} e = \{e_1^{(i)}, \ldots, e_{m_i}^{(i)}\}$ and $\partial_{a_i} f = \{f_1^{(i)}, \ldots, f_{m_i}^{(i)}\}$ (for all $i \in \{1, \ldots, n\}$).

A Derivation in c-BPA$^*_{\mathbf{0,1}}$

For $e =_{\text{def}} 1.(a.a.(b.a)^*.b)^*.0$ and $f =_{\text{def}} a.(a.(b + b.a)^*).0$, for which $e \leftrightarrows_{\boldsymbol{P}} f$ holds, we find the following proof in **c-BPA**$^*_{\mathbf{0,1}}$:

$$
\dfrac{\dfrac{\dfrac{\dfrac{(e_1 = f_1)^{\boldsymbol{u}}}{e = f_3}\,\text{COMP} \qquad \dfrac{(e_2 = f_2)^{\boldsymbol{v}}}{e_3 = f_1}\,\text{COMP}}{e_2 = f_2}\,\text{COMP/FIX}, \boldsymbol{v}}{e_1 = f_1}\,\text{COMP/FIX}, \boldsymbol{u}}{e = f}\,\text{COMP}
$$

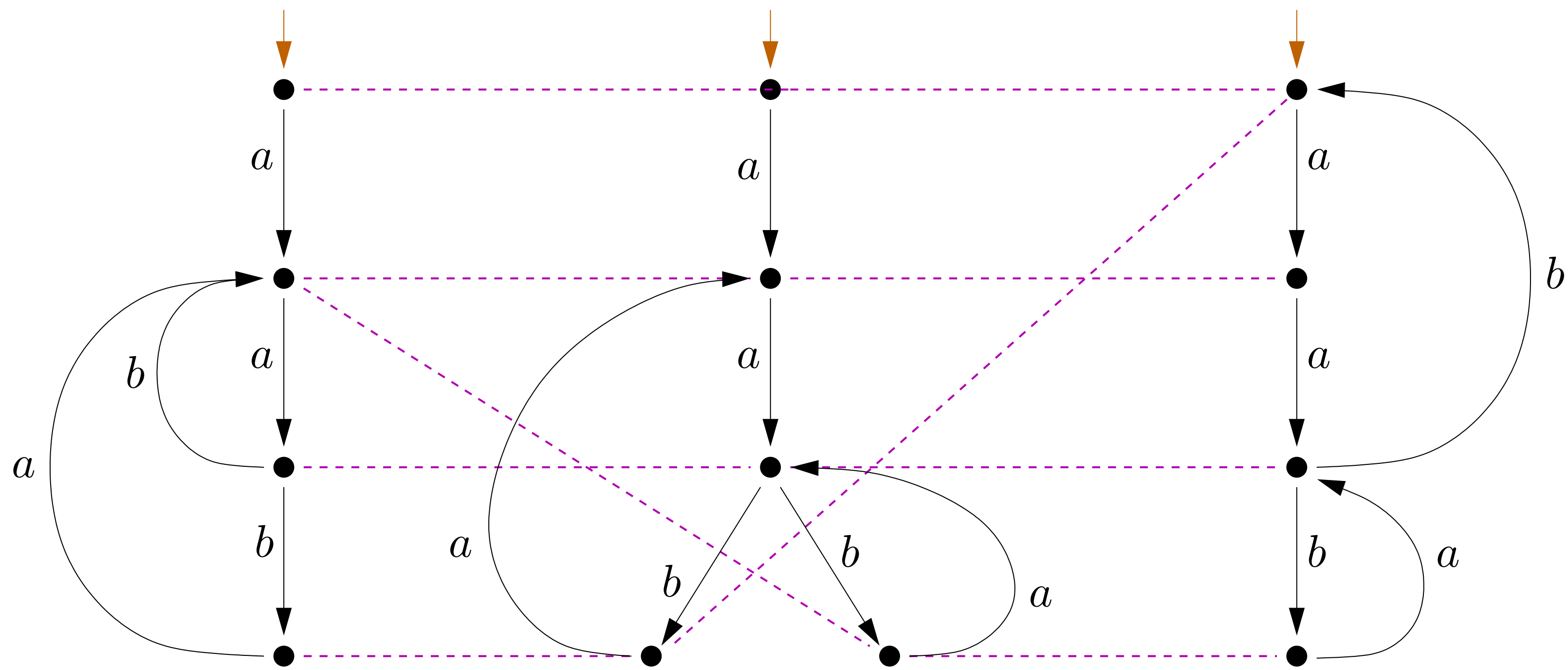
where, in particular,

$$e_2 \equiv 1.(b.a)^*.b.(a.a.(b.a)^*.b)^*.0\,,$$

$$f_2 \equiv 1.(b + b.a).(a.(b + b.a))^*.0\,,$$

$\partial_b(e_2) = \{e, e_3\}$ and $\partial_b(f_2) = \{f_1, f_3\}$.

A Derivation in c-BPA$^*_{0,1}$ (the Intuition)



"*Two-exit iteration*"

$$\notin im(\boldsymbol{P})$$

Completeness of c-BPA$^*_{0,1}$

Theorem. **c-BPA$^*_{0,1}$** *is sound and complete w.r.t.* $\underline{\leftrightarrow}_{\boldsymbol{P}}$:

$$(\forall e, f \in \mathcal{R}(\Sigma)) \left[\; \vdash_{\textbf{c-BPA}^*_{0,1}} e = f \quad \Longleftrightarrow \quad e \underline{\leftrightarrow}_{\boldsymbol{P}} f \; \right].$$

Proof. By the finitary coinduction principle for $\underline{\leftrightarrow}_{\boldsymbol{P}}$. □

Reconstructing Regular Expressions from Partial Derivatives

Let $\Sigma = \{a_1, \ldots, a_n\}$.

Lemma 1. *For all $e \in \mathcal{R}(\Sigma)$ it holds:*

$$\vdash_{\mathbf{BPA}^*_{0,1}} e = o(e) + \sum_{i=1}^{n} \sum_{e' \in \partial_{a_i}(e)} a_i.e' \,.$$

(This is reminiscent of the *fundamental theorem of calculus* that links differentiation and integration.)

Unique Solvability Principle(s)

$$\text{1-RSP}^{*}_{0,1}\ \frac{x = f.x + g}{x = f^{*}.g}\ (\text{if } \lambda \notin \boldsymbol{L}(f))$$

$$\text{1-USP}^{*}_{0,1}\ \frac{x = f.x + g \qquad y = f.y + g}{x = y}\ (\text{if } \lambda \notin \boldsymbol{L}(f))$$

$$\text{USP}^{*}_{0,1}\ \frac{\{\ x_j = E_j(x_1, \ldots, x_m)\ \}_{j=1}^{m} \qquad \{\ y_j = E_j(y_1, \ldots, y_m)\ \}_{j=1}^{m}}{x_i = y_i}$$

where, for all $i \in \{1, \ldots, m\}$,

$E_j(x_1, \ldots, x_m)$ is of the form $[1+] \sum_{k=1}^{m_j} a_{l_k}.x_{l_{j,k}}$.

Transforming into $\mathbf{BPA}^{*}_{0,1}+$ $\mathsf{USP}^{*}_{0,1}$-der's (Example)

Using the "expression reconstruction lemma", one finds that in the example the vectors $\langle e, e_1, e_2, e, e_3\rangle$ and $\langle f, f_1, f_2, f_3, f_1\rangle$ of regular expressions satisfy the same system of equations. This enables to extract from the proof in **c-BPA**$^{*}_{0,1}$ the following proof in **BPA**$^{*}_{0,1}$+$\mathsf{USP}^{*}_{0,1}$:

$$\dfrac{\dfrac{e \overset{\vdots}{=} a.e_1 \qquad e_3 \overset{\vdots}{=} a.e_2}{\dfrac{e_2 \overset{\vdots}{=} b.e + b.e_3}{\dfrac{e_1 \overset{\vdots}{=} a.e_2}{e \overset{\vdots}{=} a.e_1}}} \qquad \dfrac{f_3 \overset{\vdots}{=} a.f_1 \qquad f_1 \overset{\vdots}{=} a.f_2}{\dfrac{f_2 \overset{\vdots}{=} b.f_3 + b.f_1}{\dfrac{f_1 \overset{\vdots}{=} a.f_2}{f \overset{\vdots}{=} a.f_1}}}}{e = f}\ \mathsf{USP}^{*}_{0,1}$$

Completeness of $\mathbf{BPA}^{*}_{\mathbf{0,1}}+\mathsf{USP}^{*}_{0,1}$

Theorem. $\mathbf{BPA}^{*}_{\mathbf{0,1}}+\mathsf{USP}^{*}_{0,1}$ *is sound and complete w.r.t.* $\leftrightarrows_{\boldsymbol{P}}$ *:*

$$(\forall e, f \in \mathcal{R}(\Sigma)) \left[\; \vdash_{\mathbf{BPA}^{*}_{\mathbf{0,1}}+\mathsf{USP}^{*}_{0,1}} e = f \quad \Longleftrightarrow \quad e \leftrightarrows_{\boldsymbol{P}} f \;\right].$$

Remaining Question (equivalent to Milner's first question):

Is $\mathbf{BPA}^{*}_{\mathbf{0,1}}+1\text{-}\mathsf{USP}^{*}_{0,1}$ complete for $\leftrightarrows_{\boldsymbol{P}}$?

Observations and Results

- Antimirov's partial derivatives guide the operational behaviour of regular expressions under the process interpretation.

- A finitary coinduction principle for $\leftrightarrows_{\boldsymbol{P}}$.

- The coind. motivated, complete proof system $\textbf{c-BPA}^{*}_{\mathbf{0,1}}$ for $\leftrightarrows_{\boldsymbol{P}}$.

- Replacing the rule $\text{1-RSP}^{*}_{0,1}$ in Milner's system $\textbf{BPA}^{*}_{\mathbf{0,1}}+\text{1-RSP}^{*}_{0,1}$ by the *unique solvability principle* $\text{USP}^{*}_{0,1}$ gives a complete axiomatisation for $\leftrightarrows_{\boldsymbol{P}}$: the system $\textbf{BPA}^{*}_{\mathbf{0,1}}+\text{USP}^{*}_{0,1}$.

My thanks to Jan Rutten for a number of discussions and suggestions!

Thanks for your attention!