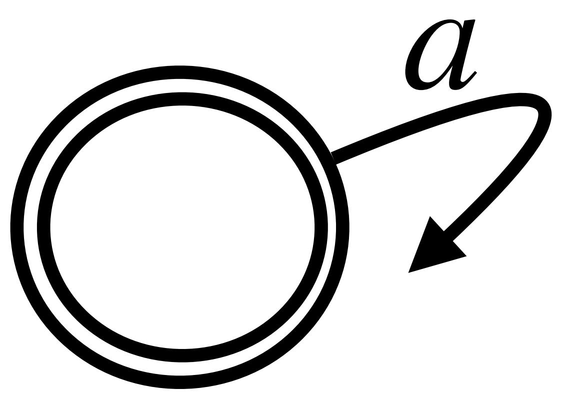
A Completeness Theorem for Probabilistic Regular Expressions

Wojciech Różowski (University College London) and Alexandra Silva (Cornell University)

Regular expressions

$$a^*$$

(Non)deterministic
finite automata



Regular languages

$$\{\varepsilon, a, aa, aaa, \ldots\} \in A^* \to 2$$

This talk:
Probabilistic Regular Expressions (PRE)

$$e, f \in \mathsf{Exp} ::= 0 \mid 1 \mid a \in A \mid e + f \mid e ; f \mid e^*$$

do e with probability p or do f with probability $1-p$

$$e, f \in \mathsf{Exp} ::= 0 \mid 1 \mid a \in A \mid e \oplus_p f \mid e;f \mid e^{[p]}$$

do e with probability p and then start again
or terminate with probability $1-p$

$\mathsf{Sem}: A* \to [0,1]$

$$\mathsf{Sem}\left(a^{\left[\frac{1}{2}\right]}\right)(w) = \begin{cases} \frac{1}{2} \times \left(\frac{1}{2}\right)^{n} & w = a^{n} \\ 0 & \text{otherwise} \end{cases}$$

The word aaa is accepted with probability $\dfrac{1}{16}$

Contribution:
Complete axiomatisation of probabilistic language equivalence of PREs

In the case of RE, problem studied by Salomaa, Krob, Boffa and Kozen

Generative Probabilistic Transition Systems

Transition function

$$\big(X, \alpha : X \to \mathcal{D}_\omega(\checkmark + A \times X)\big)$$

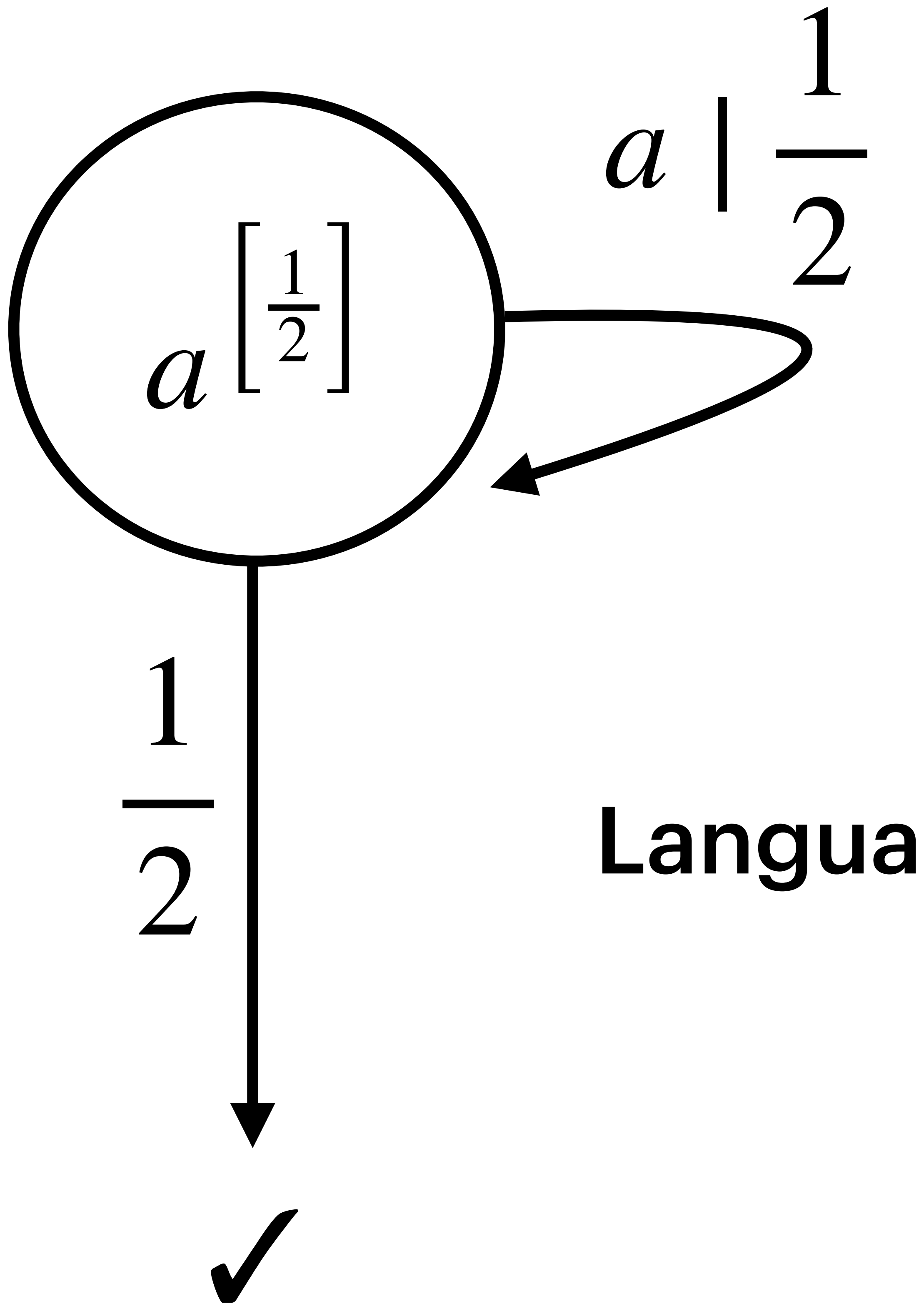
Set of states

Finitely supported subprobability distribution

Termination or labelled transitions

Operational semantics

Transition system whose states are expressions (similar to Antimirov derivatives)



$$\Big(\mathsf{Exp}, \partial : \mathsf{Exp} \to \mathscr{D}_\omega(\checkmark + A \times \mathsf{Exp})\Big)$$

$$\mathsf{Sem}(e) := \mathsf{PLan}_{(\mathsf{Exp},\partial)}(e)$$

Language semantics obtained through a construction from coalgebra

Silva, A., Bonchi, F., Bonsangue, M. and Rutten, J. (2013). Generalizing determinization from automata to coalgebras. *Logical Methods in Computer Science*

The axioms

Probabilistic choice

$$e \oplus_p e \equiv e$$

$$e \oplus_1 f \equiv e$$

$$e \oplus_p f \equiv f \oplus_{1-p} e$$

$$(e \oplus_p f) \oplus_q g \equiv e \oplus_{pq} (f \oplus_{\frac{(1-p)q}{1-pq}} g)$$

Can be generalised to an n-ary subprobalistic choice and becomes Positive Convex Algebra

Stone, M.H. (1949). Postulates for the barycentric calculus. *Annali di matematica pura ed applicata*

Sequential composition

$$1; e \equiv e \equiv e; 1$$

$$0; e \equiv 0 \equiv e; 0$$

$$e; (f; g) \equiv (e; f); g$$

$$e; (f \oplus_p g) \equiv e; g \oplus_p e; g$$

$$(e \oplus_p f); g \equiv e; g \oplus_p f; g$$

Loops

$$e^{[p]} \equiv e; e^{[p]} \oplus_p 1$$

$$(e \oplus_p 1)^{[q]} \equiv e^{\left[\frac{pq}{1-(1-p)q}\right]}$$

$$1^{[1]} \equiv 0$$

$$\frac{g \equiv eg \oplus_p f \qquad e \text{ accepts the empty word with probability } 0}{g \equiv e^{[p]}; f}$$

Soundness: If $e \equiv f$ then $\mathsf{Sem}(e) = \mathsf{Sem}(f)$ ✓

Completeness: If $\mathsf{Sem}(e) = \mathsf{Sem}(f)$ then $e \equiv f$

$$\Updownarrow$$

Equivalent to establishing a certain universal property

Using a series of very hard results in convex algebra

(Due to Sokolova, Woracek, Milius)

Milius, S. (2018). Proper Functors and Fixed Points for Finite Behaviour. *Logical Methods in Computer Science*

Sokolova, A. and Harald Woracek (2015). Congruences of convex algebras. *Journal of pure and applied algebra*

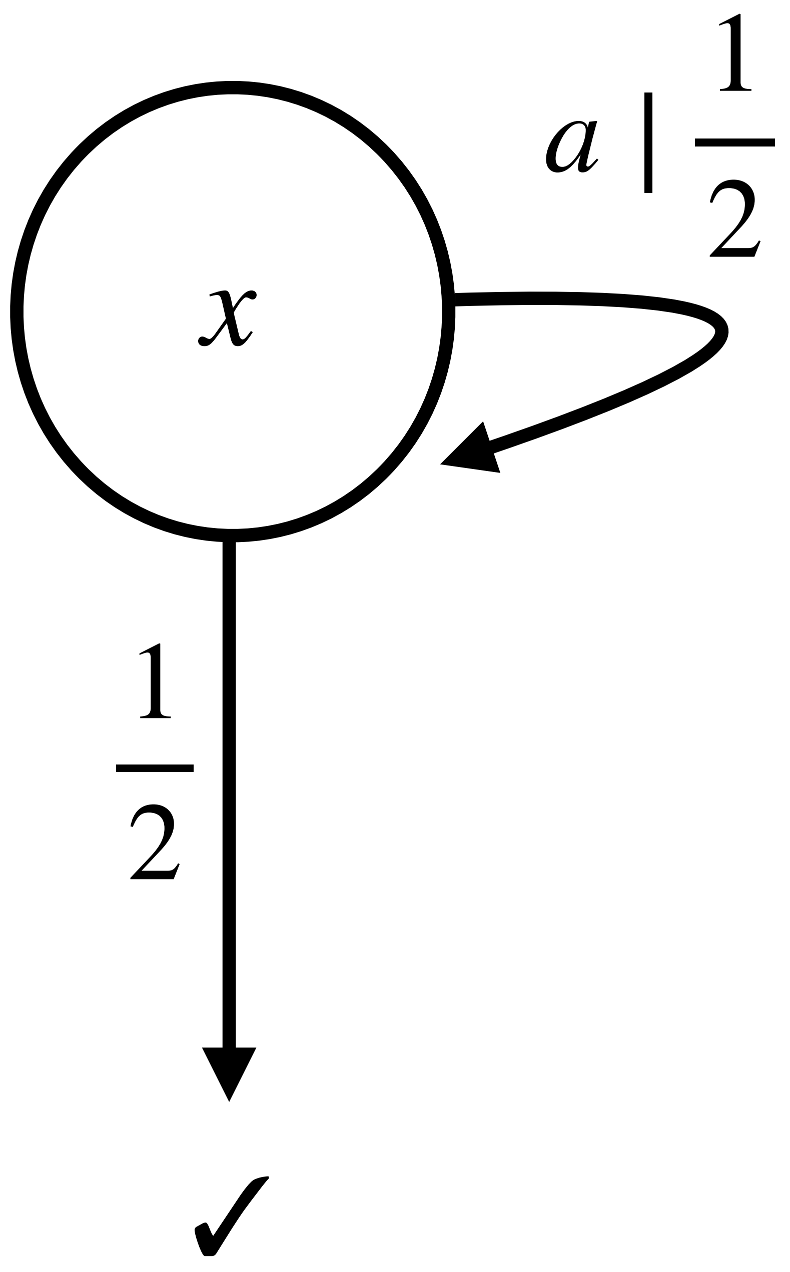
Sokolova, A. and Woracek, H. (2018). Proper Semirings and Proper Convex Functors. *Foundations of Software Science and Computation Structures (FoSSaCS)*

…and through a reduction which is the most technical bit of this work

It suffices to show:

Every finite-state automaton can be uniquely converted to an expression (up to the axioms of $\equiv$)

Automaton



System of equations

$$x \equiv ax \oplus_{\frac{1}{2}} 1$$

Solution

$$x \to a^{[\frac{1}{2}]}; 1$$

Thank you!

- Generalisation of classic automata theory through heavy tools from coalgebra and convex algebra
- Previous work by Silva and Sokolova relied on extending the complete axiomatisation modulo bismilarity of a probabilistic process calculus
- Future work: quantitative axiomatisations, algebraic Kozen-style axiomatisation

Check out the full paper :)