The Probabilistic Rabin Tree Theorem

Damian Niwiński, **Paweł Parys**, Michał Skrzypczak
University of Warsaw

FICS 2023 (paper submitted to LICS 2023)

Theorem

We can compute the probability that a random infinite tree belongs to a given regular language L.

Theorem

We can compute the probability that a random infinite tree belongs to a given regular language L.

given by, e.g.

- an MSO formula
- a nondeterministic parity automaton

full binary tree,
each label chosen
independently in random

- the result is an algebraic number
- can be computed in 3-EXPTIME
- can be compared with a given rational q in 2-EXPSPACE

Context

Decidable

- some results for ω-words (probability always rational)
- infinite trees: the probability exists (not clear because regular languages of infinite trees need not to be Borel) *[Gogacz, Michalewski, Mio, Skrzypczak 2017]*
- determ. top-down parity autom. *[Chen, Dräger, Kiefer 2012]*
- game automata *[Michalewski, Mio 2015]*
- weak MSO *[Niwiński, Przybyłko, Skrzypczak 2020]*

Undecidable

- nonemptiness for probabilistic automata (exists a finite word accepted with probability >0.5)
- value-1 for probabilistic automata (exists a sequence of finite words where acceptance probability tends to 1)
- exists a ω-word accepted by a probabilistic Büchi automaton with probability >0.

Open

- Satisfiability of PCTL*

Two worlds

Languages $\longrightarrow$ Probabilities

Two worlds

$$\text{Languages} \longrightarrow \text{Probabilities}$$

Key difficulty:

$$\begin{tikzcd} A \arrow[dr] & & B \arrow[dl] \\ & A \cup B & \end{tikzcd} \qquad\qquad \begin{tikzcd} \mathbb{P}(A) \arrow[dr] & & \mathbb{P}(B) \arrow[dl] \\ & \mathbb{P}(A \cup B) & \end{tikzcd}$$

Two worlds

Languages $\longrightarrow$ Probabilities

Key difficulty:

$$
\begin{tikzcd}[column sep=small]
A \arrow[dr] & & B \arrow[dl] \\
& A \cup B &
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}[column sep=small]
\mathbb{P}(A) \arrow[dr] & & \mathbb{P}(B) \arrow[dl] \\
& \cancel{\mathbb{P}(A \cup B)} &
\end{tikzcd}
$$

Another aspect:

(random variables)

$$
\begin{tikzcd}[column sep=small]
X \arrow[dr] & & Y \arrow[dl] \\
& F(X,Y) &
\end{tikzcd}
$$

distribution of X, distribution of Y

~~distribution of $F(X,Y)$~~

$$
\begin{tikzcd}
\text{distribution of } X \times Y \arrow[d] \\
\text{distribution of } F(X,Y)
\end{tikzcd}
$$

✔

Step 1

~~Nondeterministic~~ automata $\longrightarrow$ μ-calculus / powersets

Step 1

~~Nondeterministic automata~~ $\longrightarrow$ μ-calculus / powersets

Basic objects: profiles

τ : trees $\rightarrow$ P(Q)

Profile τ_A corresponding to automaton A:

$\tau_A(t)$ = states from which t can be accepted

Proposition: $\tau_A(t) = \mu x_1.\nu x_2.\mu x_3.\nu x_4\ldots\mu x_{d-1}.\nu x_d.\delta(x_1,x_2,\ldots,x_d)$

where $\delta(x_1,x_2,\ldots,x_d)$ applies transition function once

(transitions of priority i go to x_i)

Goal: compute probability distribution of the random variable τ_A

Step 2

$\cancel{\delta(\tau_1,\tau_2,\ldots,\tau_d)}$ $\longrightarrow$ $\bar{\tau}=(\tau_1,\tau_2,\ldots,\tau_d)$

$\Delta(\bar{\tau})$

Step 2

$\cancel{\delta(\tau_1,\tau_2,\ldots,\tau_d)} \longrightarrow \bar{\tau}=(\tau_1,\tau_2,\ldots,\tau_d)$

$\Delta(\bar{\tau})$

Convenient to take:

$\Delta(\bar{\tau})=(\delta(\tau_1,\tau_1,\tau_1,\ldots,\tau_1), \delta(\tau_1,\tau_2,\tau_2,\ldots,\tau_2), \delta(\tau_1,\tau_2,\tau_3,\ldots,\tau_3), \ldots, \delta(\tau_1,\tau_2,\ldots,\tau_d))$

Previously: $\tau_A(t) = \mu x_1.\nu x_2.\mu x_3.\nu x_4\ldots\mu x_{d-1}.\nu x_d.\delta(x_1,x_2,\ldots,x_d)$

Now: convenient to write things like: $\mu x.F(x\vee y)$, $\nu x.F(x\wedge y)$

(but $\vee,\wedge$ does not translate to probabilities)

Step 3

$\cancel{\mu x.F(x \vee y)} \longrightarrow F\!\uparrow\!(y)$

Intuition behind $\mu x.F(x \vee y)$ (but not precise meaning):
least fixed point of F above y

We define: $F\!\uparrow\!(y)$ = least fixed point of F above y

Unary μ-calculus

Syntax: H, $F_1;F_2$, $F\uparrow$, $F\downarrow$ (defines a one-argument function $V \to V$)

composition

fixed base functions

$F\downarrow(y)$ = greatest fixed point of F below y

$F\uparrow(y)$ = least fixed point of F above y

Unary μ-calculus

partial function

Syntax: H, $F_1;F_2$, $F\uparrow$, $F\downarrow$ (defines a one-argument ~~function~~ $V \to V$)

composition

fixed base functions

$F\downarrow(y)$ = greatest fixed point of F below y

$F\uparrow(y)$ = least fixed point of F above y

Problem: $F\uparrow(y)$ may be undefined

- maybe there are no fixed points above y
- maybe there are many incomparable fixed points above y

So: $F\uparrow$ is a partial function

Unary μ-calculus – type system

How to prove that a formula of unary μ-calculus has a defined value?

Type system: statements $F :: A \to B$ (F is defined on A and has values in B)

$$\frac{\forall x \in A . H(x) \in B}{H :: A \to B}$$

$$\frac{F_1 :: A \to C \qquad F_2 :: C \to B}{F_1 ; F_2 :: A \to B}$$

Unary μ-calculus – type system

How to prove that a formula of unary μ-calculus has a defined value?

Type system: statements $F :: A \to B$ (F is defined on A and has values in B)

$$\frac{}{H :: A \to B}\ \forall x \in A.\, H(x) \in B \qquad\qquad \frac{F_1 :: A \to C \qquad F_2 :: C \to B}{F_1 ; F_2 :: A \to B}$$

$$\frac{F :: A \to A}{F\uparrow :: A \to B}\ A \text{ chain complete, } \forall x \in A.\ F(x) \geq x, \text{ Fix}(F) \cap A \subseteq B$$

$$\frac{F :: A \to A}{F\downarrow :: A \to B}\ A \text{ chain complete, } \forall x \in A.\ F(x) \leq x, \text{ Fix}(F) \cap A \subseteq B$$

every chain of elements of A has infimum and supremum in A

Why? $F\uparrow(x)$ / $F\downarrow(x)$ will be reached by:

- applying F
- taking limits of chains

Unary μ-calculus – the formula

How to define τ_A in unary μ-calculus?

Base functions:

- $\Delta(\tau_1,\tau_2,\ldots,\tau_d)=(\delta(\tau_1,\tau_1,\tau_1,\ldots,\tau_1),\ \delta(\tau_1,\tau_2,\tau_2,..,\tau_2),\ \delta(\tau_1,\tau_2,\tau_3,\ldots,\tau_3),\ \ldots,\ \delta(\tau_1,\tau_2,\ldots,\tau_d))$
- $\mathrm{Bid}_n(\tau_1,\tau_2,\ldots,\tau_d)=(\tau_1,\ldots,\tau_{n\text{-}2},\tau_{n\text{-}1},\tau_{n\text{-}2},\tau_{n\text{-}2},\ldots,\tau_{n\text{-}2})$ for $n=1,\ldots,d$; $\tau_{-1}=\bot$; $\tau_0=\top$
- $\mathrm{Cut}_n(\tau_1,\tau_2,\ldots,\tau_d)=(\tau_1,\ldots,\tau_{n\text{-}2},\tau_{n\text{-}1},\tau_{n+1},\tau_{n+1},\ldots,\tau_{n+1})$ for $n=1,\ldots,d\text{-}1$

  (Bid_n and Cut_n only swap coordinates)

Unary μ-calculus – the formula

How to define τ_A in unary μ-calculus?

Base functions:

- $\Delta(\tau_1,\tau_2,\ldots,\tau_d)=(\delta(\tau_1,\tau_1,\tau_1,\ldots,\tau_1), \delta(\tau_1,\tau_2,\tau_2,\ldots,\tau_2), \delta(\tau_1,\tau_2,\tau_3,\ldots,\tau_3), \ldots, \delta(\tau_1,\tau_2,\ldots,\tau_d))$
- $\mathrm{Bid}_n(\tau_1,\tau_2,\ldots,\tau_d)=(\tau_1,\ldots,\tau_{n-2},\tau_{n-1},\tau_{n-2},\tau_{n-2},\ldots,\tau_{n-2})$ for $n=1,\ldots,d$; $\tau_{-1}=\bot$; $\tau_0=\top$
- $\mathrm{Cut}_n(\tau_1,\tau_2,\ldots,\tau_d)=(\tau_1,\ldots,\tau_{n-2},\tau_{n-1},\tau_{n+1},\tau_{n+1},\ldots,\tau_{n+1})$ for $n=1,\ldots,d-1$

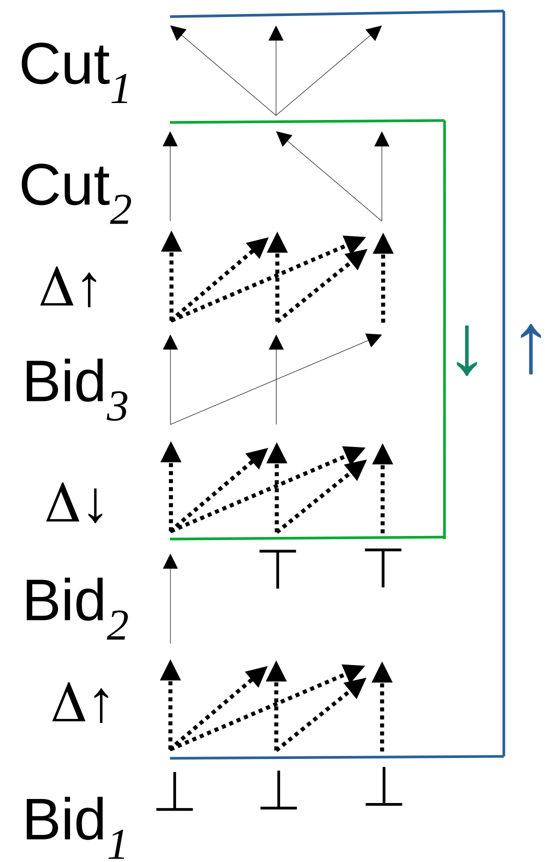
  (Bid$_n$ and Cut$_n$ only swap coordinates)

$\Phi_d = \mathrm{Bid}_d;\Delta\uparrow$ (odd d)

$\Phi_d = \mathrm{Bid}_d;\Delta\downarrow$ (even d)

$\Phi_n = \mathrm{Bid}_n;(\Delta\uparrow;\Phi_{n+1};\mathrm{Cut}_n)\uparrow$ (odd $n<d$)

$\Phi_n = \mathrm{Bid}_n;(\Delta\downarrow;\Phi_{n+1};\mathrm{Cut}_n)\downarrow$ (even $n<d$)



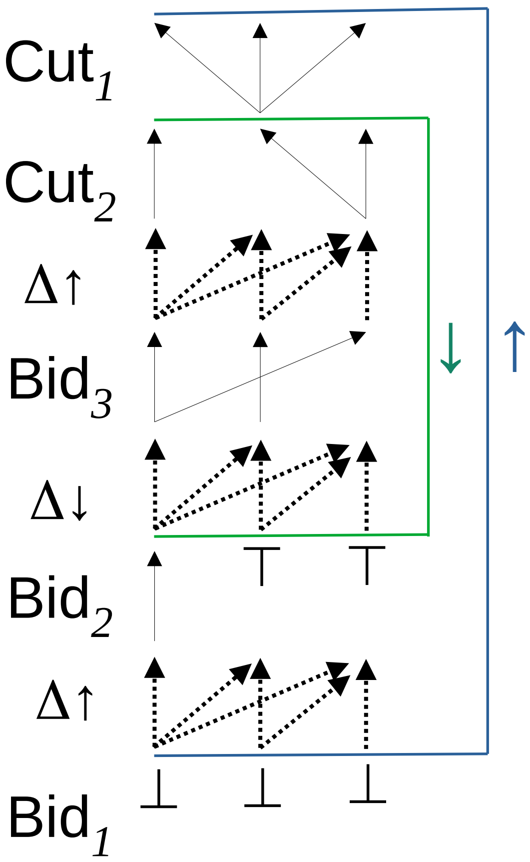
Unary μ-calculus – the formula

Φ_d = Bid$_d$;$\Delta\uparrow$ (odd d)

Φ_d = Bid$_d$;$\Delta\downarrow$ (even d)

Φ_n = Bid$_n$;($\Delta\uparrow$;Φ_{n+1};Cut$_n$)$\uparrow$ (odd $n<d$)

Φ_n = Bid$_n$;($\Delta\downarrow$;Φ_{n+1};Cut$_n$)$\downarrow$ (even $n<d$)



What has to be shown?

1) $\Phi_1(\cdot)$ is defined (using the type system)

2) $\Phi_1(\cdot)$ computes τ_A

3) all intermediate profiles used while computing $\Phi_1(\cdot)$ are measurable

4) the same computation can be done on distributions

Why the value is well defined?

We define sets S_n – we have $(\tau_1,\tau_2,\dots,\tau_d) \in S_n$ if

- $\tau_1 \le \tau_3 \le \tau_5 \le \dots \le \tau_6 \le \tau_4 \le \tau_2$
- $\tau_n = \tau_{n+1} = \tau_{n+2} = \dots = \tau_d$
- first n-1 coordinates of $\Delta(\tau_1,\tau_2,\dots,\tau_d)$ are $(\tau_1,\tau_2,\dots,\tau_{n-1})$
- $\Delta(\tau_1,\tau_2,\dots,\tau_d) \ge (\tau_1,\tau_2,\dots,\tau_d)$ if n odd, and
- $\Delta(\tau_1,\tau_2,\dots,\tau_d) \ge (\tau_1,\tau_2,\dots,\tau_d)$ if n even.

For the base functions we derive:

- $\Delta :: S_n \rightarrow S_n$
- $\text{Bid}_n :: S_n \rightarrow S_n$
- $\text{Cut}_n :: S_{n+2} \rightarrow S_n$

Then we show (using the type system) that:

- $\Phi_n :: S_n \rightarrow S_{n+1}$

Why is the value correct?

(why $\Phi_1(\cdot)$ computes τ_A?)

Why is the value correct?

(why $\Phi_1(\cdot)$ computes τ_A?)

Step 1:

Recall the intuition: $F\uparrow(y)$ was introduced to simulate $\mu x.F(x \vee y)$.

The typing rule says:

$$\frac{F :: A \to A}{F\uparrow :: A \to B} \quad A \text{ chain complete, } \forall x \in A .\ F(x) \geq x, \text{ Fix}(F) \cap A \subseteq B$$

so $F\uparrow(y) = \mu x.F(x \vee y)$ for $y \in A$.

Why is the value correct?

(why $\Phi_1(\cdot)$ computes τ_A?)

Step 1:

Recall the intuition: $F\uparrow(y)$ was introduced to simulate $\mu x.F(x \vee y)$.

The typing rule says:

$$\frac{F :: A \to A}{F\uparrow :: A \to B} \quad A \text{ chain complete, } \forall x \in A\,.\, F(x) \geq x,\ \mathrm{Fix}(F) \cap A \subseteq B$$

so $F\uparrow(y) = \mu x.F(x \vee y)$ for $y \in A$.

Step 2:

Change Φ_1 into $\mu x_1.\nu x_2.\mu x_3.\nu x_4 \ldots \mu x_{d-1}.\nu x_d.\delta(x_1, x_2, \ldots, x_d)$ using some laws of μ-calculus, like

$\mu x.\nu y.F(x, x \vee y) = \mu x.\nu y.F(x, y)$

$\mu x.\nu y.\mu z.F(x, y, x \vee z) = \mu x.\nu y.\mu z.F(x, y, z)$

$\mu x.\nu y.F(\mu z.F(x \vee z, x \vee z), y) = \mu x.\nu y.F(x, y)$

Measurability

Recall that $F{\uparrow}(x)$ / $F{\downarrow}(x)$ can be reached from x by:

- applying F
- taking limits of chains

Difficulty:
We need to know that all intermediate values in this computation are measurable (so it makes sense to consider their probability distribution)

Measurability

Recall that $F{\uparrow}(x)$ / $F{\downarrow}(x)$ can be reached from x by:

- applying F
- taking limits of chains

Difficulty:
We need to know that all intermediate values in this computation are measurable (so it makes sense to consider their probability distribution)

Solution:
Similar proof as for showing that every regular language is measurable
[Gogacz, Michalewski, Mio, Skrzypczak 2017], [Lusin, Sierpiński 1918]

Moreover:
(in this case) probability of the limit of a chain is the limit of probabilities.

Probability distributions

Profiles $\longrightarrow$ Distributions

$\tau : \text{trees} \rightarrow \mathrm{P}(Q \times \{1,\ldots,d\})$

$\hat{\tau} : \mathbb{D}(\mathrm{P}(Q \times \{1,\ldots,d\}))$

$\hat{\tau}(R) = \mathbb{P}(\{t \mid \tau(t) = R\})$

Probability distributions

Profiles $\longrightarrow$ Distributions

τ : trees $\rightarrow$ P$(Q \times \{1,\ldots,d\})$

$\hat{\tau}$: $\mathbb{D}$(P$(Q \times \{1,\ldots,d\})$)

$\hat{\tau}(R) = \mathbb{P}(\{t \mid \tau(t) = R\})$

coordinatewise order $\longrightarrow$ probabilistic powerdomain order

[Jones, Plotkin 1989]

for each upward-closed $U \subseteq$ P$(Q \times \{1,\ldots,d\})$

$$\sum_{R \in U} \alpha(R) \leq \sum_{R \in U} \beta(R)$$

Probability distributions

Profiles $\longrightarrow$ Distributions

$\tau : \text{trees} \to \mathrm{P}(Q \times \{1,\dots,d\})$

$\hat{\tau} : \mathbb{D}(\mathrm{P}(Q \times \{1,\dots,d\}))$

$\hat{\tau}(R) = \mathbb{P}(\{t \mid \tau(t) = R\})$

coordinatewise order $\longrightarrow$ probabilistic powerdomain order

[Jones, Plotkin 1989]

for each upward-closed $U \subseteq \mathrm{P}(Q \times \{1,\dots,d\})$

$$\sum_{R \in U} \alpha(R) \leq \sum_{R \in U} \beta(R)$$

Δ, Bid_n, Cut_n $\longrightarrow$ Δ, Bid_n, Cut_n

Φ_1 $\longrightarrow$ Φ_1

Φ_1 can be expressed in first-order logic over reals – decidable by Tarski

(the formula is of exponential size)

Conclusions

- We shown how to compute the probability that a random infinite tree belongs to a given regular language.
- We introduced unary μ-calculus, which works well for orders without $\vee$ and $\wedge$ (e.g. probability distributions)

Thank you