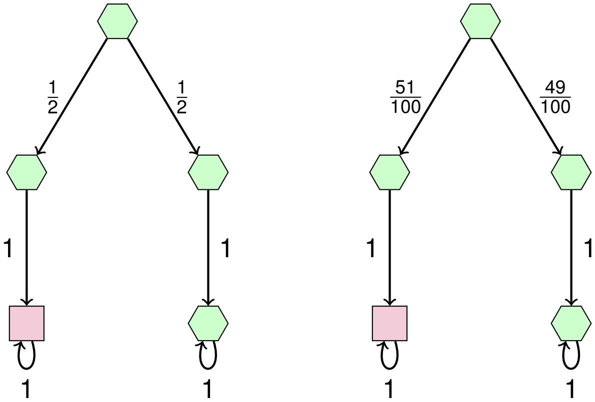
How to Explain Probabilistic Bisimilarity Distances?

Franck van Breugel

York University, Toronto

Probabilistic bisimilarity is *not* robust



Probabilistic bisimilarity is *not* robust



Giacalone, Jou, and Smolka, PROCOMET 1990.

Probabilistic bisimilarity is *not* robust

Fundamental problem

Probabilistic bisimilarity is *not* robust.

Probabilistic bisimilarity is *not* robust

Fundamental problem

Probabilistic bisimilarity is *not* robust.

Robust alternative

Instead of an equivalence relation

$$\sim : S \times S \to \{\text{true}, \text{false}\}$$

use distances

$$d : S \times S \to [0, 1].$$

Probabilistic bisimilarity is *not* robust

Fundamental problem

Probabilistic bisimilarity is *not* robust.

Robust alternative

Instead of an equivalence relation

$$\sim : S \times S \to \{\text{true}, \text{false}\}$$

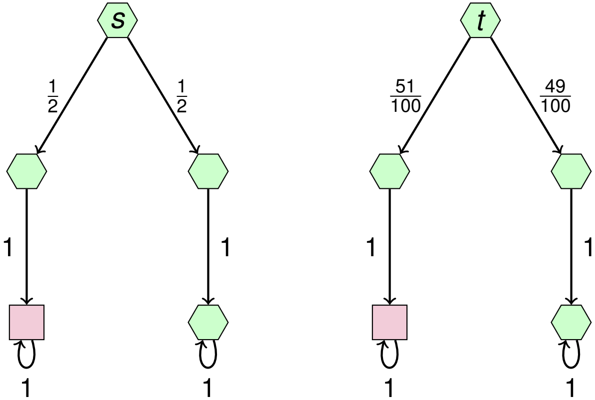
use distances

$$d : S \times S \to [0, 1].$$

Generalization

For all states s and t, $s \sim t$ iff $d(s, t) = 0$.

Probabilistic bisimilarity distances



Question

How do we explain that the distance of s and t is $\frac{1}{100}$?

Explainability

Question

How do we explain why states of a labelled transition system are *not* bisimilar?

Explainability

Question

How do we explain why states of a labelled transition system are *not* bisimilar?

- Game
  Ehrenfeucht and Fraïssé, 1950 and 1961.
- Logic
  Hennessy and Milner, 1980.
- ...

Bisimulation game

Theorem

States s and t of a labelled transition system are bisimilar iff duplicator has a winning strategy from (s, t).

Bisimulation game

Theorem

States s and t of a labelled transition system are bisimilar iff duplicator has a winning strategy from (s, t).

Question

How do we explain that states of a labelled transition system are *not* bisimilar?

Bisimulation game

Theorem

States s and t of a labelled transition system are bisimilar iff duplicator has a winning strategy from (s, t).

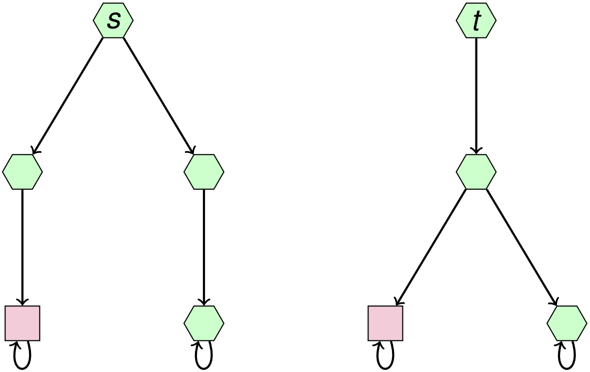
Question

How do we explain that states of a labelled transition system are *not* bisimilar?

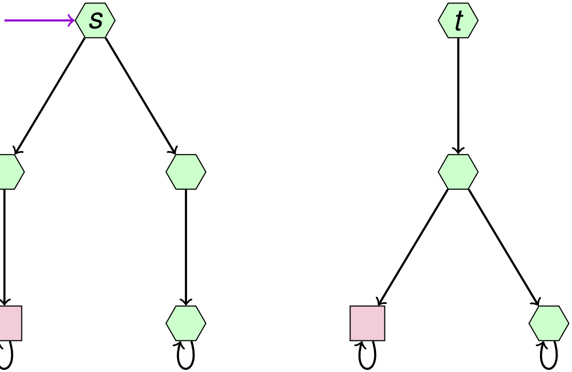
Answer

Find a strategy for spoiler that is winning from (s, t).

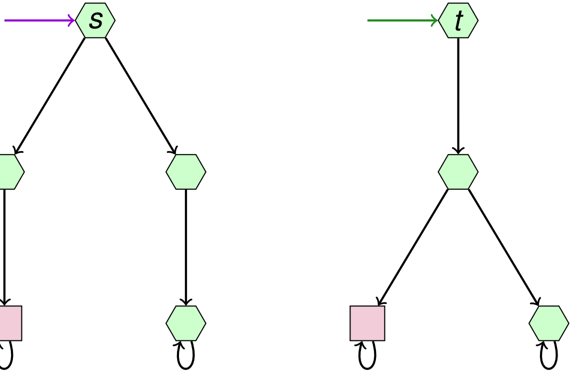
Bisimilarity



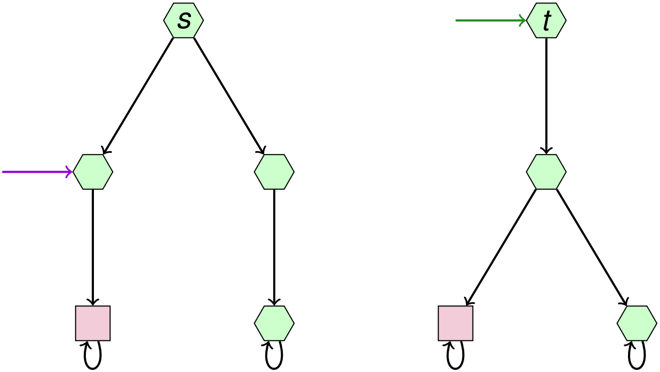
Bisimilarity



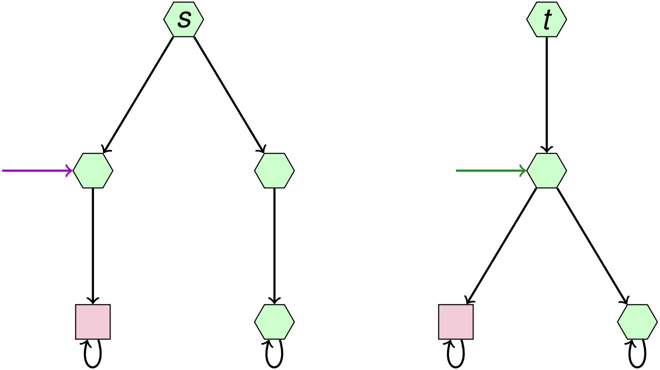
Bisimilarity



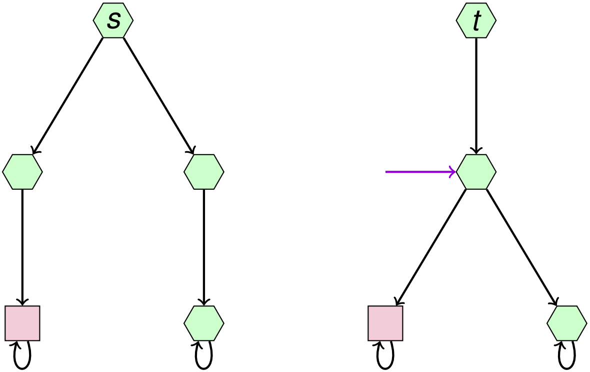
Bisimilarity



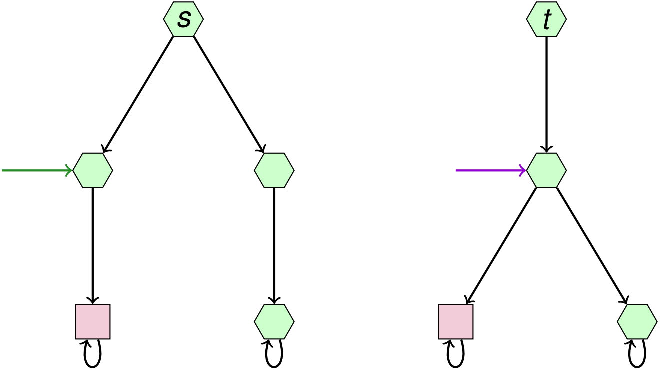
Bisimilarity



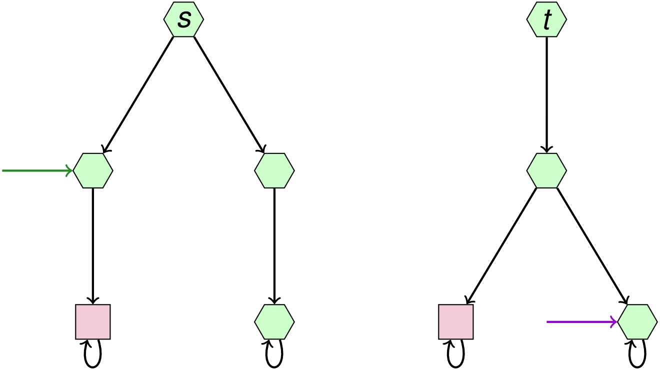
Bisimilarity



Bisimilarity



Bisimilarity



Hennessy-Milner theorem

Theorem

States of a labelled transition system are bisimilar iff they satisfy the same formulas of the Hennessy-Milner logic.

Hennessy and Milner, ICALP 1980.

Explainability

Question

How do we explain that states of a labelled transition system are *not* bisimilar?

Explainability

Question

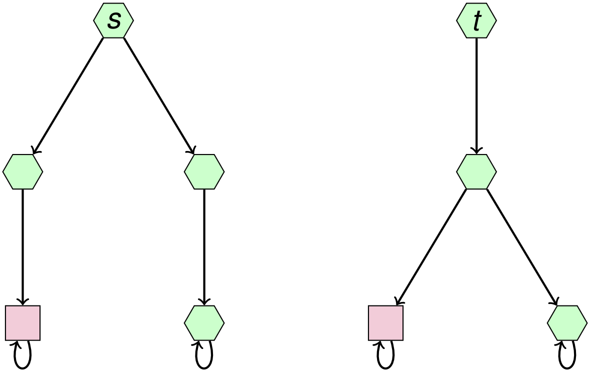
How do we explain that states of a labelled transition system are *not* bisimilar?

Answer

Find a formula of the Hennessy-Milner logic that is satisfied in one of the states but not the other.

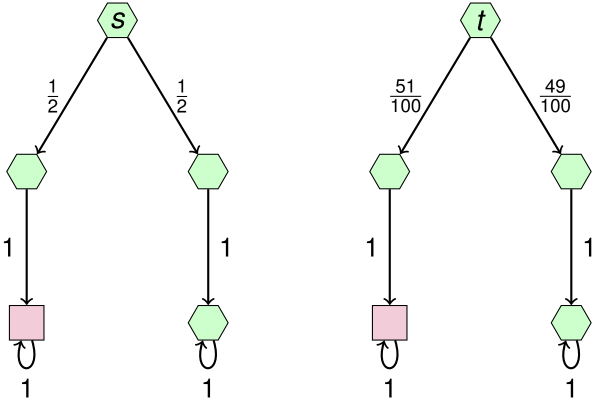
Cleaveland, CAV 1990.

Bisimilarity



State *t* satisfies $\bigcirc((\bigcirc$  $) \wedge (\bigcirc$  $))$, but state *s* does not.

Probabilistic bisimilarity distances



Question

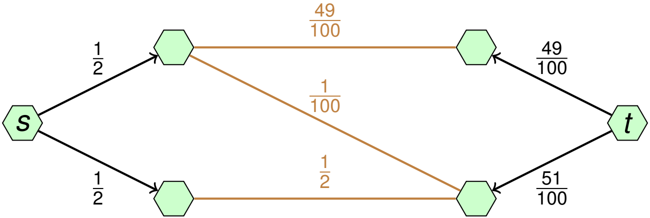
How do we explain that the distance of s and t is $\frac{1}{100}$?

Probabilistic bisimilarity distances

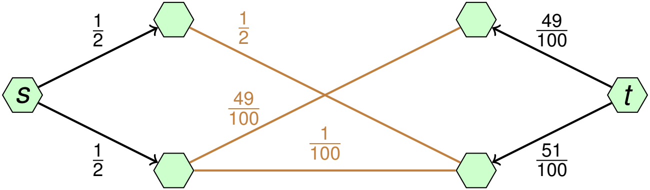
1. Game.
2. Logic.

Matching probabilities

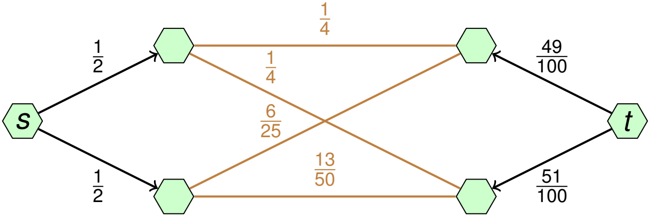
Matching probabilities



Matching probabilities

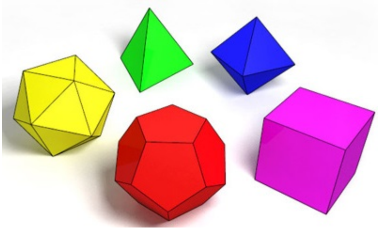


Matching probabilities



Matching probabilities = couplings

The couplings form a convex polytope



Doeblin, 1938.

A convex polytope is fully determined by its vertices

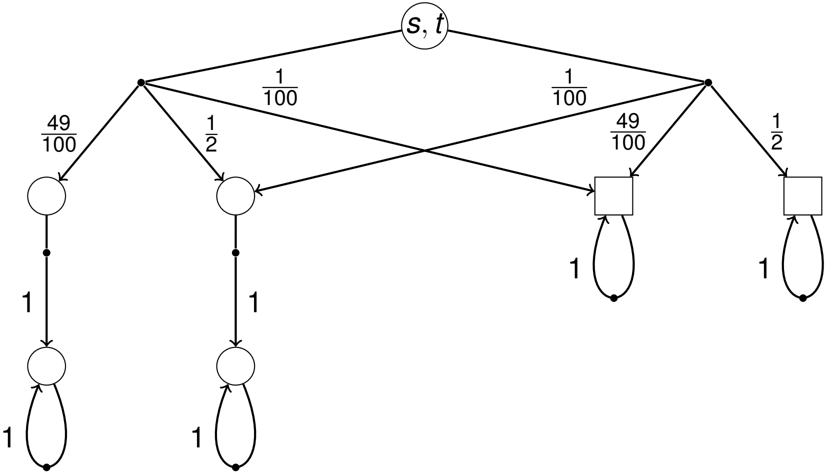
Probabilistic bisimilarity distance game

Labelled Markov chain $\langle S, \ldots \rangle$

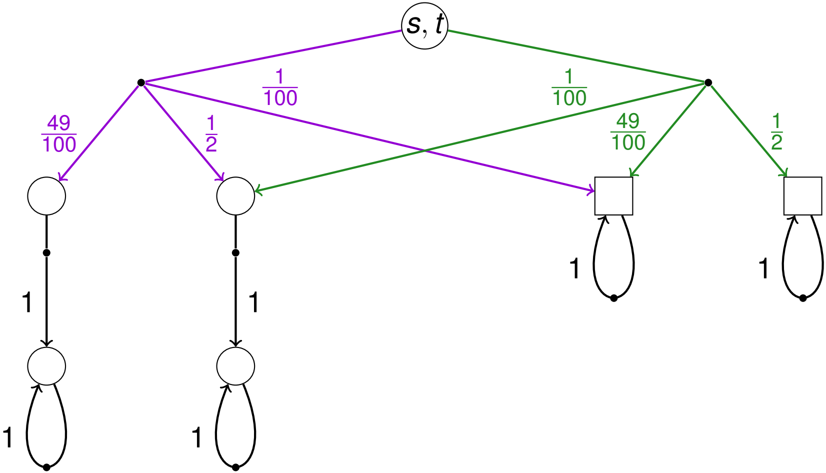
$\Downarrow$

Markov decision process $\langle S \times S, \ldots \rangle$

Markov decision process



Markov decision process



Probabilistic bisimilarity distances game

Theorem

$$d(s, t) = \min_{\text{strategy}} \text{probability of reaching } \square$$

Chen, Worrell, and vB, FoSSaCS 2012.

Explainability

Question

How do we explain the probabilistic bisimilarity distances?

Explainability

Question

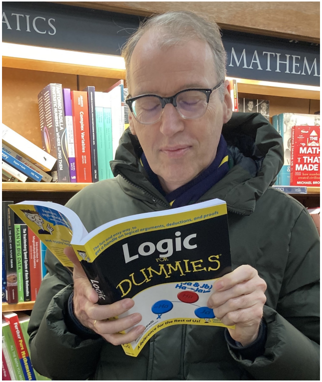
How do we explain the probabilistic bisimilarity distances?

Answer

Find a strategy that minimizes the probability of reaching $\Box$.

Probabilistic bisimilarity distances

1. Game.
2. Logic.



A quantitative Hennessy-Milner theorem

Theorem

For all states s, t,

$$d(s,t) = \sup_{f \in \mathcal{L}} [\![f]\!](s) - [\![f]\!](t).$$

A quantitative Hennessy-Milner theorem

Theorem

For all states s, t,

$$d(s,t) = \sup_{f \in \mathcal{L}} [\![f]\!](s) - [\![f]\!](t).$$

The logic $\mathcal{L}$ is defined by

$$f ::= a \mid \neg f \mid \bigcirc f \mid f \ominus q \mid f \wedge f$$

where a is a label and $q \in \mathbb{Q} \cap [0,1]$.

A logical characterization

Theorem

For all states s, t,

$$d(s,t) = \sup_{f \in \mathcal{L}} [\![f]\!](s) - [\![f]\!](t).$$

$$[\![a]\!](s) = \begin{cases} 1 & \text{if } a \text{ is label of } s \\ 0 & \text{otherwise} \end{cases}$$

A logical characterization

Theorem

For all states s, t,

$$d(s,t) = \sup_{f \in \mathcal{L}} [\![f]\!](s) - [\![f]\!](t).$$

$$\begin{aligned} [\![a]\!](s) &= \begin{cases} 1 & \text{if } a \text{ is label of } s \\ 0 & \text{otherwise} \end{cases} \\ [\![\neg f]\!](s) &= 1 - [\![f]\!](s) \end{aligned}$$

A logical characterization

Theorem

For all states s, t,

$$d(s,t) = \sup_{f \in \mathcal{L}} [\![f]\!](s) - [\![f]\!](t).$$

$$\begin{aligned}
[\![a]\!](s) &= \begin{cases} 1 & \text{if } a \text{ is label of } s \\ 0 & \text{otherwise} \end{cases} \\
[\![\neg f]\!](s) &= 1 - [\![f]\!](s) \\
[\![\bigcirc f]\!](s) &= \textstyle\sum_{t \in S} P(s,t)\, [\![f]\!](t)
\end{aligned}$$

A logical characterization

Theorem

For all states s, t,

$$d(s,t) = \sup_{f \in \mathcal{L}} [\![f]\!](s) - [\![f]\!](t).$$

$$\begin{array}{rcl}
[\![a]\!](s) & = & \begin{cases} 1 & \text{if } a \text{ is label of } s \\ 0 & \text{otherwise} \end{cases} \\
[\![\neg f]\!](s) & = & 1 - [\![f]\!](s) \\
[\![\bigcirc f]\!](s) & = & \sum_{t \in S} P(s,t)\, [\![f]\!](t) \\
[\![f \ominus q]\!](s) & = & [\![f]\!](s) - q
\end{array}$$

A logical characterization

Theorem

For all states s, t,

$$d(s,t) = \sup_{f \in \mathcal{L}} \llbracket f \rrbracket(s) - \llbracket f \rrbracket(t).$$

$$\begin{aligned}
\llbracket a \rrbracket(s) &= \begin{cases} 1 & \text{if } a \text{ is label of } s \\ 0 & \text{otherwise} \end{cases} \\
\llbracket \neg f \rrbracket(s) &= 1 - \llbracket f \rrbracket(s) \\
\llbracket \bigcirc f \rrbracket(s) &= \textstyle\sum_{t \in S} P(s,t) \llbracket f \rrbracket(t) \\
\llbracket f \ominus q \rrbracket(s) &= \max(\llbracket f \rrbracket(s) - q, 0)
\end{aligned}$$

A logical characterization

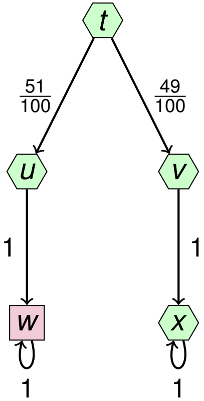
Theorem

For all states s, t,

$$d(s,t) = \sup_{f \in \mathcal{L}} [\![f]\!](s) - [\![f]\!](t).$$

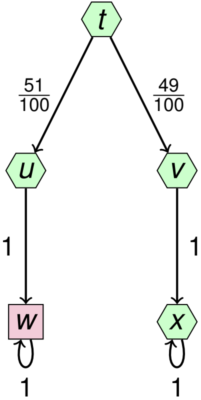
$$\begin{array}{rcl}
[\![a]\!](s) & = & \begin{cases} 1 & \text{if } a \text{ is label of } s \\ 0 & \text{otherwise} \end{cases} \\
[\![\neg f]\!](s) & = & 1 - [\![f]\!](s) \\
[\![\bigcirc f]\!](s) & = & \sum_{t \in S} P(s,t)\,[\![f]\!](t) \\
[\![f \ominus q]\!](s) & = & \max([\![f]\!](s) - q, 0) \\
[\![f \wedge g]\!](s) & = & \min([\![f]\!](s), [\![g]\!](s))
\end{array}$$

A logical characterization



$$[\![\bigcirc(⬢ \wedge \bigcirc ⬢)]\!](t)$$

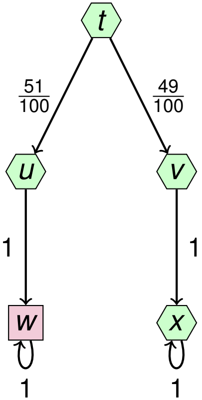
A logical characterization



$$\llbracket \bigcirc(⬡ \wedge \bigcirc ⬡) \rrbracket(t)$$

$$= \frac{51}{100} \llbracket ⬡ \wedge \bigcirc ⬡ \rrbracket(u) + \frac{49}{100} \llbracket ⬡ \wedge \bigcirc ⬡ \rrbracket$$

A logical characterization

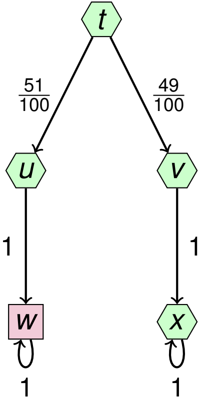


$$[\![\bigcirc(\hexagon \wedge \bigcirc \hexagon)]\!](t)$$

$$= \frac{51}{100} [\![\hexagon \wedge \bigcirc \hexagon]\!](u) + \frac{49}{100} [\![\hexagon \wedge \bigcirc \hexagon]\!]$$

$$= \frac{51}{100} [\![\bigcirc \hexagon]\!](u) + \frac{49}{100} [\![\bigcirc \hexagon]\!](v)$$

A logical characterization

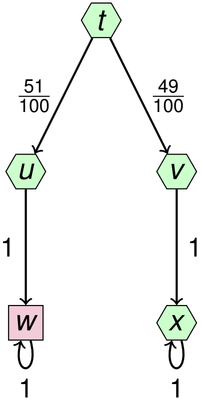


$$\llbracket \bigcirc(\hexagon \wedge \bigcirc \hexagon) \rrbracket (t)$$

$$= \frac{51}{100} \llbracket \hexagon \wedge \bigcirc \hexagon \rrbracket (u) + \frac{49}{100} \llbracket \hexagon \wedge \bigcirc \hexagon \rrbracket$$

$$= \frac{51}{100} \llbracket \bigcirc \hexagon \rrbracket (u) + \frac{49}{100} \llbracket \bigcirc \hexagon \rrbracket (v)$$

$$= \frac{51}{100} \llbracket \hexagon \rrbracket (w) + \frac{49}{100} \llbracket \hexagon \rrbracket (x)$$



$$[\![\bigcirc(\hexagon \wedge \bigcirc \hexagon)]\!](t)$$

$$= \frac{51}{100} [\![\hexagon \wedge \bigcirc \hexagon]\!](u) + \frac{49}{100} [\![\hexagon \wedge \bigcirc \hexagon]\!]$$

$$= \frac{51}{100} [\![\bigcirc \hexagon]\!](u) + \frac{49}{100} [\![\bigcirc \hexagon]\!](v)$$

$$= \frac{51}{100} [\![\hexagon]\!](w) + \frac{49}{100} [\![\hexagon]\!](x)$$

$$= \frac{49}{100}$$

Explainability

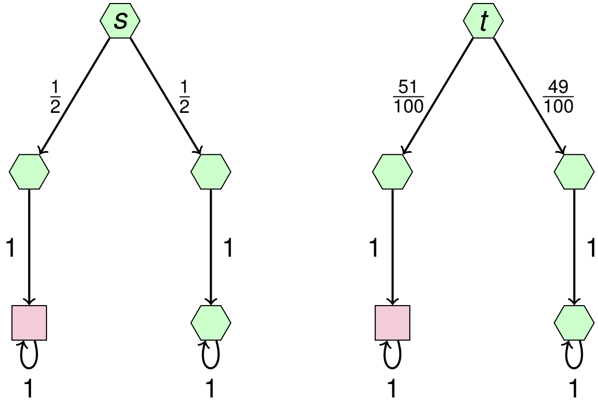
Question

How do we explain the probabilistic bisimilarity distances?

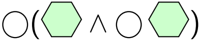
Attempt

Find a formula f of the logic $\mathcal{L}$ with $d(s, t) = [\![f]\!](s) - [\![f]\!](t)$.

Explainability



The distance of states *s* and *t* can be explained by the formula



Question

How do we explain the probabilistic bisimilarity distances?

Attempt fails

Find a formula f of the logic $\mathcal{L}$ with $d(s, t) = [\![f]\!](s) - [\![f]\!](t)$.

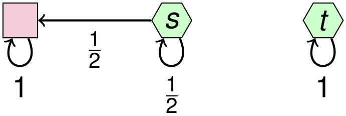
Explainability

Question

How do we explain the probabilistic bisimilarity distances?

Attempt fails

Find a formula f of the logic $\mathcal{L}$ with $d(s, t) = [\![f]\!](s) - [\![f]\!](t)$.



For every formula f, $[\![f]\!](s) - [\![f]\!](t) < d(s, t) = 1$.

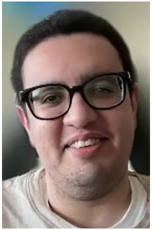
Explainability

Question

How do we explain the probabilistic bisimilarity distances?

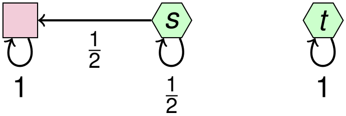
Attempt succeeds

Find a sequence $(f_n)_n$ of formulas of the logic $\mathcal{L}$ with $d(s,t) = \lim_n [\![f_n]\!](s) - [\![f_n]\!](t)$.



Rady and vB, FoSSaCS 2023.

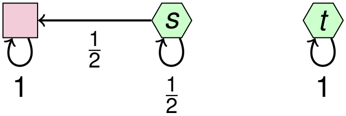
Explainability



The distance of states *s* and *t* can be explained by the sequence of formulas

- false
- $\bigcirc \square$
- $\bigcirc (\square \vee \bigcirc \square)$
- $\bigcirc (\square \vee \bigcirc (\square \vee \bigcirc \square))$
- $\ldots$

Explainability



The distance of states *t* and *s* can be explained by the sequence of formulas

- false
- $(\bigcirc ⬡) \ominus \frac{1}{2}$
- $(\bigcirc (⬡ \wedge ((\bigcirc ⬡) \ominus \frac{1}{2} \oplus \frac{1}{2}))) \ominus \frac{1}{4}$
- $(\bigcirc (⬡ \wedge ((\bigcirc (⬡ \wedge ((\bigcirc ⬡) \ominus \frac{1}{2} \oplus \frac{1}{2}))) \ominus \frac{1}{4} \oplus \frac{1}{4}))) \ominus \frac{1}{8}$
- $\ldots$

Some technical details

Objective

Construct a sequence $(f_{st}^n)_n$ of formulas of the logic $\mathcal{L}$ with $d(s,t) = \lim_n [\![f_{st}^n]\!](s) - [\![f_{st}^n]\!](t)$.

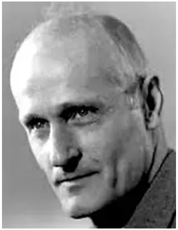
Some technical details

Objective

Construct a sequence $(f^n_{st})_n$ of formulas of the logic $\mathcal{L}$ with $d(s,t) = \lim_n [\![f^n_{st}]\!](s) - [\![f^n_{st}]\!](t)$.

Kleene fixed point theorem

$d = \lim_n d_n$.



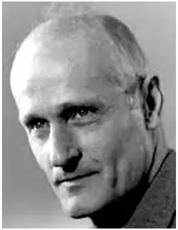
Some technical details

Objective

Construct a sequence $(f_{st}^n)_n$ of formulas of the logic $\mathcal{L}$ with $d(s,t) = \lim_n [\![f_{st}^n]\!](s) - [\![f_{st}^n]\!](t)$.

Kleene fixed point theorem

$d = \lim_n d_n$.



Objective

Construct a formula f_{st}^n of the logic $\mathcal{L}$ with $d_n(s,t) = [\![f_{st}^n]\!](s) - [\![f_{st}^n]\!](t)$.

Some technical details

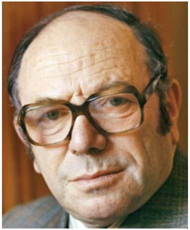
Objective

Construct a formula f^n_{st} of the logic $\mathcal{L}$ with
$d_n(s,t) = [\![f^n_{st}]\!](s) - [\![f^n_{st}]\!](t)$.

Kantorovich-Rubinstein duality theorem

There exists $F^n_{st} : S \to [0,1]$ such that

$$d_n(s,t) = \left(\sum_{u \in S} P(s,u)\, F^n_{st}(u)\right) - \left(\sum_{u \in S} P(t,u)\, F^n_{st}(u)\right).$$



Some technical details

The formula g^n_{stuv} is defined by

$$\begin{cases} \text{false} \oplus F^n_{st}(u) & \text{if } F^n_{st}(u) = F^n_{st}(v) \\ (f^n_{uv} \ominus (d_n(u,v) - (F^n_{st}(u) - F^n_{st}(v))) \oplus F^n_{st}(v) & \text{if } F^n_{st}(u) > F^n_{st}(v) \\ (f^n_{vu} \ominus (d_n(u,v) - (F^n_{st}(v) - F^n_{st}(u))) \oplus F^n_{st}(u) & \text{otherwise.} \end{cases}$$

Stone-Weierstrass approximation theorem

$$f^n_{st} = \left[\!\left[\bigwedge_{u \in S} \bigvee_{v \in S} g^n_{stuv} \right]\!\right]$$

The formula f_{st}^{n}

Some technical details

The formula f^n_{st}

If s and t have a different label then f^n_{st}= label of s.

Some technical details

The formula f^n_{st}

If s and t have a different label then f^n_{st}= label of s.

Otherwise,

$$f^n_{st} = \left(\bigcirc \bigwedge_{u \in S} \bigvee_{v \in S} g^n_{stuv} \right) \ominus \left(\sum_{u \in S} P(t,u)\, F^n_{st}(u) \right).$$

Open problems

- How to extend the logic $\mathcal{L}$ and, for states s and t, find a formula $f_{st} \in \mathcal{L}$ such that

$$d(s,t) = [\![f_{st}]\!](s) - [\![f_{st}]\!](t)?$$

- How to extend the logic $\mathcal{L}$ and, for state s, find a formula $f_s \in \mathcal{L}$ such that for all states t

$$d(s,t) = [\![f_s]\!](s) - [\![f_s]\!](t)?$$

- How to extend the logic $\mathcal{L}$ and, for state s, find a formula $f_s \in \mathcal{L}$ such that for all states t

$$d(s,t) = [\![f_s]\!](t)?$$

References

Slide 3

- Photo of Smolka:
  www.concur2016.ulaval.ca/fr/program/speakers_html/
- Alessandro Giacalone, Chi-Chang Jou, and Scott Smolka. Algebraic reasoning for probabilistic concurrent systems. In *Proceedings of the IFIP WG 2.2/2.3 Working Conference on Programming Concepts and Methods*, pages 443–458, Sea of Gallilee, Israel, April 1990. North-Holland.

References

Slide 16

- Photo of Hennessy: www.scss.tcd.ie/Matthew.Hennessy/
- Photo of Milner: www.cl.cam.ac.uk/archive/rm135/
- Matthew Hennessy and Robin Milner. On observing nondeterminism and concurrency. In Jaco de Bakker and Jan van Leeuwen, editors, *Proceedings of the 7th Colloquium on Automata, Languages and Programming*, volume 85 of *Lecture Notes in Computer Science*, pages 299–309, Noordwijkerhout, The Netherlands, July 1980. Springer-Verlag.

References

Slide 17

- Photo of Cleaveland: cmns.umd.edu/people/rance-cleaveland/
- Rance Cleaveland. On automatically distinguishing inequivalent processes. In Edmund Clarke and Robert Kurshan, editors, *Proceedings of a DIMACS Workshop on Computer Aided Verification*, volume 3 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 463–476, New Brunswick, NJ, USA, June 1990. DIMACS/AMS.

Slide 25

- Photo of Doeblin: mathshistory.st-andrews.ac.uk/Biographies/Doeblin/
- Wolfgang Doeblin. Exposé de la théorie des chaînes simples constantes de Markoff à un nombre fini d'états. *Revue Mathématiques de l'Union Interbalkanique*, 2: 77-105, 1938.
- Image of polytopes: www.greatlittleminds.com/pages/maths/3d-platonic-solids.html

Slide 29

- Photo of Worrell: www.cs.ox.ac.uk/people/james.worrell/home.html
- Di Chen, Franck van Breugel, and James Worrell. On the complexity of computing probabilistic bisimilarity. In Lars Birkedal, editor, *Proceedings of the 15th International Conference on Foundations of Software Science and Computational Structures*, volume 7213 of *Lecture Notes in Computer Science*, pages 437–451, Tallinn, Estonia, March/April 2012. Springer-Verlag.

Slide 38

- Amgad Rady and Franck van Breugel. Explainability of probabilistic bisimilarity distances for labelled Markov chains. In Orna Kupferman and Pawel Sobocinski, editors, *Proceedings of the 26th International Conference on Foundations of Software Science and Computational Structures*, volume 13992 of *Lecture Notes in Computer Science*, pages 285–307, Paris, France, April 2023. Springer-Verlag.

References

Slide 41

- Photo of Kleene: www.computerhope.com/people/stephen_kleene.htm
- Stephen Kleene. *Introduction to Metamathematics*.

References

Slide 42

- Photo of Kantorovich: armstrongeconomics.wordpress.com/research/economic-thought/by-author/kantorovich-leonid/
- Leonid Kantorovich and Gennadi Rubinstein. On the space of completely additive functions (in Russian). *Vestnik Leningradskogo Universiteta*, 3(2):52–59, 1958.

Slide 43

- Photo of Stone: wildpeaches.xyz/blog/curve-fitting-with-julia/
- Photo of Weierstrass: wildpeaches.xyz/blog/curve-fitting-with-julia/
- Marshall Stone. Applications of the theory of Boolean rings to general topology. *Transactions of the American Mathematical Society*, 41(3): 375–481, 1937.