Deciding Probabilistic Bisimilarity Distance One for Labelled Markov Chains

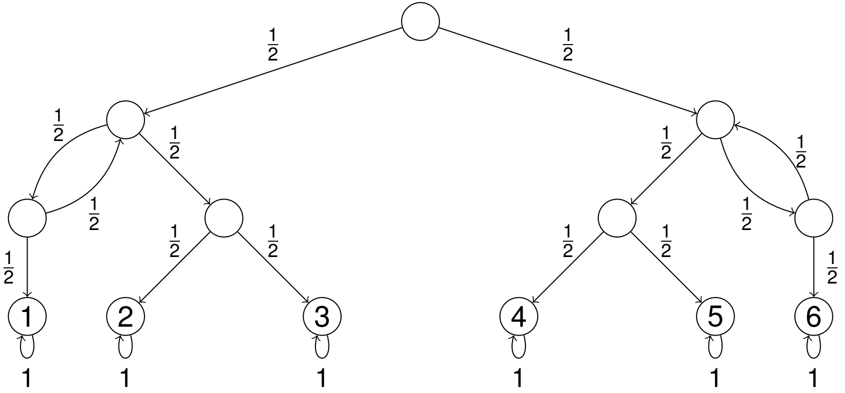
Franck van Breugel

Joint work with Qiyi Tang

December 12, 2017

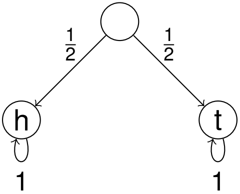
Table of Contents

Labelled Markov Chain

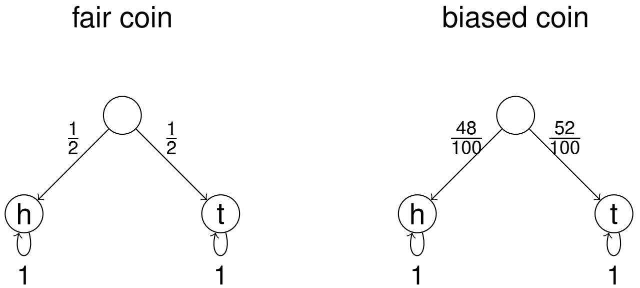


Probabilistic Bisimilarity is not Robust

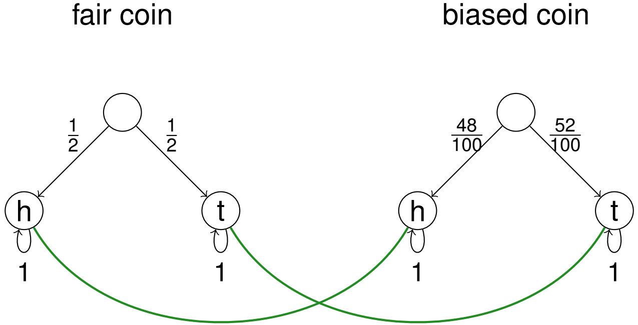
fair coin



Probabilistic Bisimilarity is not Robust

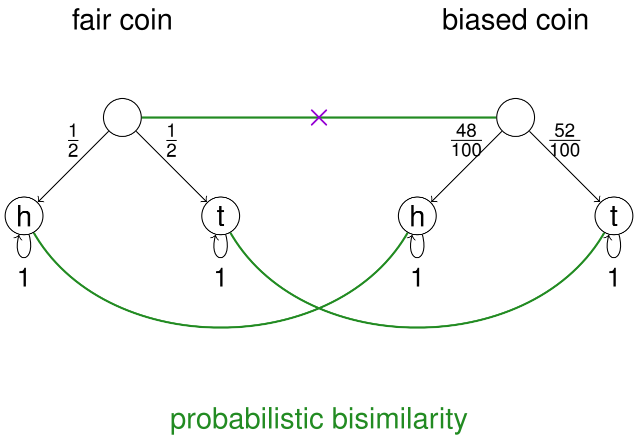


Probabilistic Bisimilarity is not Robust

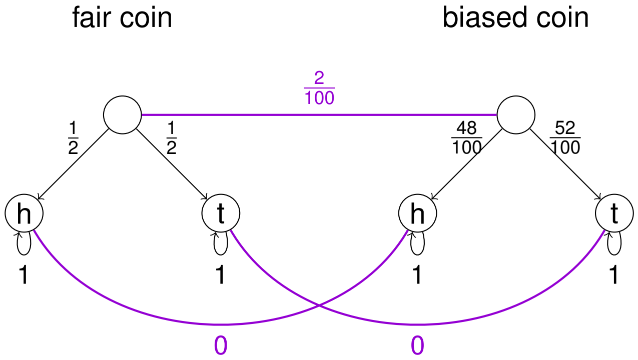


probabilistic bisimilarity

Probabilistic Bisimilarity is not Robust



Probabilistic Bisimilarity Distances



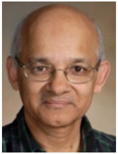
Each state has distance zero to itself. All other distances are one.

Probabilistic Bisimilarity Distances

Theorem

States are probabilistic bisimilar if and only if their probabilistic bisimilarity distance is zero.



Desharnais, Gupta, Jagadeesan and Panangaden.
CONCUR 1999.

Probabilistic Bisimilarity Distances

Franck van Breugel. Probabilistic bisimilarity distances. SIGLOG News, 4(4):33–51, October 2017.

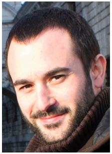
Algorithm to Compute Distances

1. Decide probabilistic bisimilarity in $O(m \lg n)$

Derisavi, Hermanns and Sanders. IPL 2003.

2. Policy iteration in $\Omega(2^n)$

Bacci, Bacci, Larsen and Mardare. TACAS 2013

Main Result

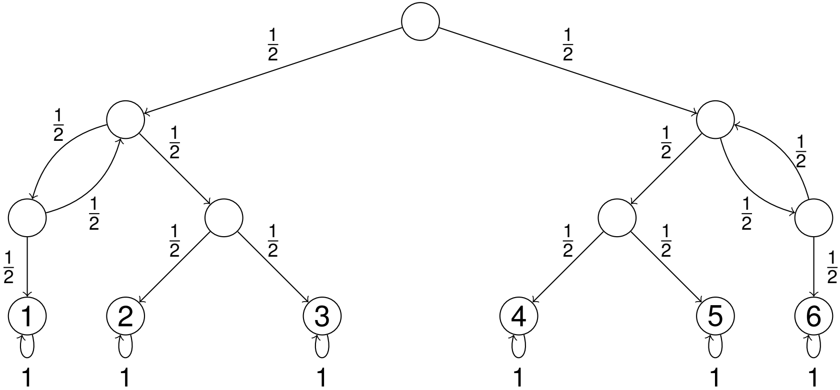
Theorem

Distance one can be decided in $O(n^2 + m^2)$.

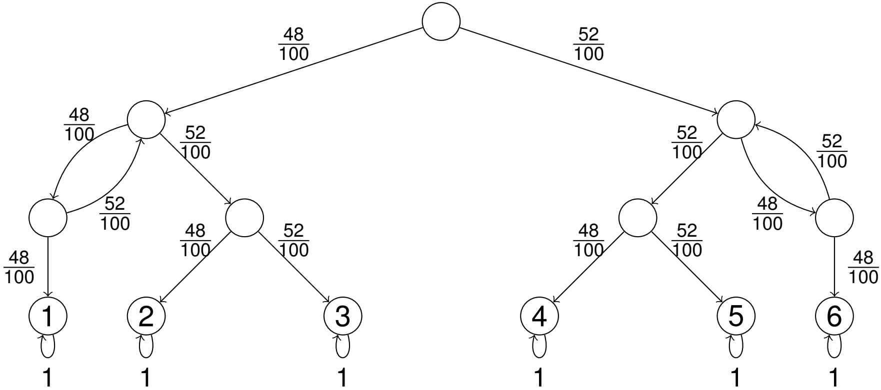
New Algorithm to Compute Distances

1. Decide distance zero in $O(m \lg n)$
2. Decide distance one in $O(n^2 + m^2)$
3. Policy iteration in $\Omega(2^n)$

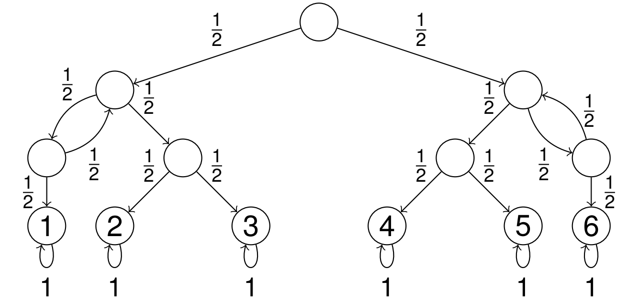
New Algorithm to Compute Distances

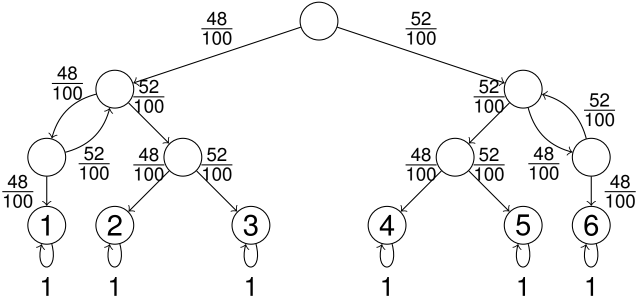


New Algorithm to Compute Distances



New Algorithm to Compute Distances





New Algorithm to Compute Distances

Donald Knuth and Andrew Yao. The Complexity of Nonuniform Random Number Generation. In Proceedings of a Symposium on New Directions and Recent Results in Algorithms and Complexity, pages 375–428, Pittsburgh, PA, USA, April 1976. Academic Press.

Labelled Markov chain with 26 states and 36 transitions

DHS + B^2LM algorithm: 4.753 seconds

New Algorithm to Compute Distances

Donald Knuth and Andrew Yao. The Complexity of Nonuniform Random Number Generation. In Proceedings of a Symposium on New Directions and Recent Results in Algorithms and Complexity, pages 375–428, Pittsburgh, PA, USA, April 1976. Academic Press.

Labelled Markov chain with 26 states and 36 transitions

DHS + B^2LM algorithm: 4.753 seconds
Our algorithm: 0.237 seconds

New Algorithm to Compute Distances

Alon Itai and Michael Rodeh. Symmetry Breaking in Distributed Networks. Information and Computation, 88(1):60–87, September 1990.

Labelled Markov chain with 147 states and 210 transitions

DHS + B^2LM algorithm: 49 hours

New Algorithm to Compute Distances

Alon Itai and Michael Rodeh. Symmetry Breaking in Distributed Networks. Information and Computation, 88(1):60–87, September 1990.

Labelled Markov chain with 147 states and 210 transitions

DHS + B^2LM algorithm: 49 hours
Our algorithm: 0.013 seconds

Any Non-Trivial Distances?

1. Decide distance zero in $O(m \lg n)$
2. Decide distance one in $O(n^2 + m^2)$

Any Non-Trivial Distances?

Alon Itai and Michael Rodeh. Symmetry Breaking in Distributed Networks. Information and Computation, 88(1):60–87, September 1990.

Labelled Markov chain with 12400 states and 16495 transitions

Any Non-Trivial Distances?

Alon Itai and Michael Rodeh. Symmetry Breaking in Distributed Networks. Information and Computation, 88(1):60–87, September 1990.

Labelled Markov chain with 12400 states and 16495 transitions

Our algorithm: 2971.244 seconds

Compute Distances smaller than ϵ

1. Decide distance zero
2. Decide distance one
3. Compute $\Delta(d)$ where

$$d(s,t) = \begin{cases} 1 & \text{if distance of } s \text{ and } t \text{ is one} \\ 0 & \text{otherwise} \end{cases}$$

4. Partial policy iteration for

$$\{\, (s,t) \in S \times S \mid \Delta(d)(s,t) \leq \epsilon \,\}$$

Compute Distances smaller than ϵ

Donald Knuth and Andrew Yao. The Complexity of Nonuniform Random Number Generation. In Proceedings of a Symposium on New Directions and Recent Results in Algorithms and Complexity, pages 375–428, Pittsburgh, PA, USA, April 1976. Academic Press.

Labelled Markov chain with 26 states and 36 transitions

DHS + B^2LM algorithm: 4.753 seconds

Compute Distances smaller than ϵ

Donald Knuth and Andrew Yao. The Complexity of Nonuniform Random Number Generation. In Proceedings of a Symposium on New Directions and Recent Results in Algorithms and Complexity, pages 375–428, Pittsburgh, PA, USA, April 1976. Academic Press.

Labelled Markov chain with 26 states and 36 transitions

DHS + B^2LM algorithm: 4.753 seconds
Our algorithm: 0.237 seconds

Compute Distances smaller than ϵ

Donald Knuth and Andrew Yao. The Complexity of Nonuniform Random Number Generation. In Proceedings of a Symposium on New Directions and Recent Results in Algorithms and Complexity, pages 375–428, Pittsburgh, PA, USA, April 1976. Academic Press.

Labelled Markov chain with 26 states and 36 transitions

DHS + B^2LM algorithm: 4.753 seconds
Our algorithm: 0.237 seconds
Our algorithm with $\epsilon = 0.2$: 0.076 seconds

Table of Contents

Labelled Markov Chain

Definition

A labelled Markov chain is a tuple $\langle S, L, \tau, \ell \rangle$ consisting of

- a nonempty finite set S of states,
- a nonempty finite set of L of labels,
- a transition function $\tau : S \to Distr(S)$ and
- a labelling function $\ell : S \to L$.

The probability of transitioning from state s to state t is $\tau(s)(t)$.

Probabilistic Bisimilarity Distances

Definition

The function $\Delta : [0,1]^{S \times S} \to [0,1]^{S \times S}$ is defined as follows. Let $d : S \times S \to [0,1]$ and $s, t \in S$.

- If $\ell(s) \neq \ell(t)$ then
$$\Delta(d)(s,t) = 1.$$
- If $\ell(s) = \ell(t)$ then
$$\Delta(d)(s,t) = \min_{c \in \mathcal{C}(\tau(s),\tau(t))} \sum_{u,v \in S} c(u,v)\, d(u,v).$$

Probabilistic Bisimilarity Distances

Definition

The function $\Delta : [0,1]^{S \times S} \to [0,1]^{S \times S}$ is defined as follows.
Let $d : S \times S \to [0,1]$ and $s, t \in S$.

- If $\ell(s) \neq \ell(t)$ then
$$\Delta(d)(s,t) = 1.$$
- If $\ell(s) = \ell(t)$ then
$$\Delta(d)(s,t) = \min_{c \in \mathcal{C}(\tau(s),\tau(t))} \sum_{u,v \in S} c(u,v)\, d(u,v).$$

Proposition

Δ is a monotone function from the complete lattice $[0,1]^{S \times S}$ to itself.

Probabilistic Bisimilarity Distances

Definition

The function $\Delta : [0,1]^{S \times S} \to [0,1]^{S \times S}$ is defined as follows. Let $d : S \times S \to [0,1]$ and $s, t \in S$.

- If $\ell(s) \neq \ell(t)$ then
$$\Delta(d)(s,t) = 1.$$
- If $\ell(s) = \ell(t)$ then
$$\Delta(d)(s,t) = \min_{c \in \mathcal{C}(\tau(s),\tau(t))} \sum_{u,v \in S} c(u,v)\, d(u,v).$$

Proposition

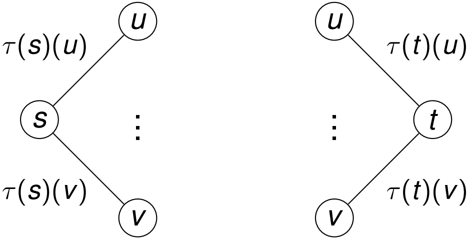
Δ is a monotone function from the complete lattice $[0,1]^{S \times S}$ to itself.

Corollary

Δ has a least fixed point, denoted $\mathrm{lfp}(\Delta)$.

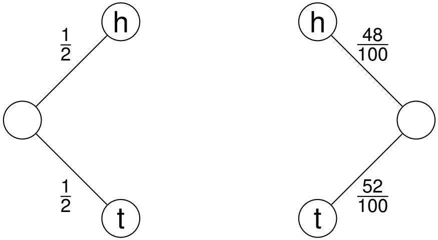
Probabilistic Bisimilarity Distances

$$\min_{c \in \mathcal{C}(\tau(s), \tau(t))} \sum_{u,v \in S} c(u,v)\, d(u,v)$$

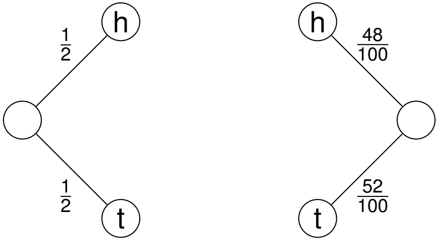


$d(u,v)$: cost to transport one unit between u and v
$c(u,v)$: amount transported between u and v

Probabilistic Bisimilarity Distances

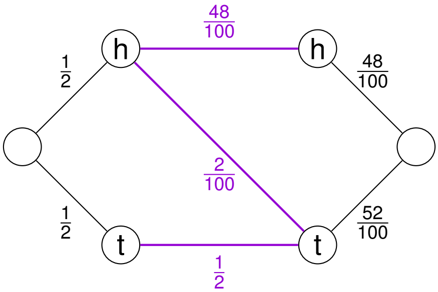


Probabilistic Bisimilarity Distances



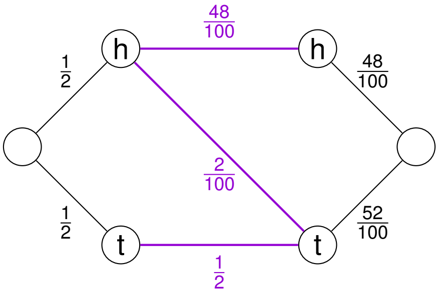
	h	t
h	0	1
t	1	0

Probabilistic Bisimilarity Distances



	h	t
h	0	1
t	1	0

Probabilistic Bisimilarity Distances



	h	t
h	0	1
t	1	0

$$\frac{48}{100} \times 0 + \frac{2}{100} \times 1 + \frac{1}{2} \times 1 = \frac{2}{100}$$

Distance Zero and One

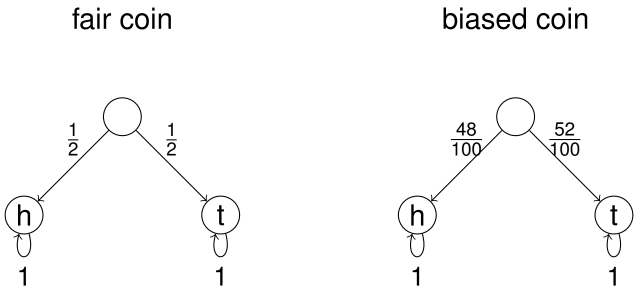
The set $S^2 = S \times S$ is partitioned:

$$\begin{aligned} S^2_0 &= \{ (s,t) \in S^2 \mid s \sim t \} \\ S^2_1 &= \{ (s,t) \in S^2 \mid \ell(s) \neq \ell(t) \} \\ S^2_? &= S^2 \setminus (S^2_0 \cup S^2_1) \end{aligned}$$

Distance Zero and One

The set $S^2 = S \times S$ is partitioned:

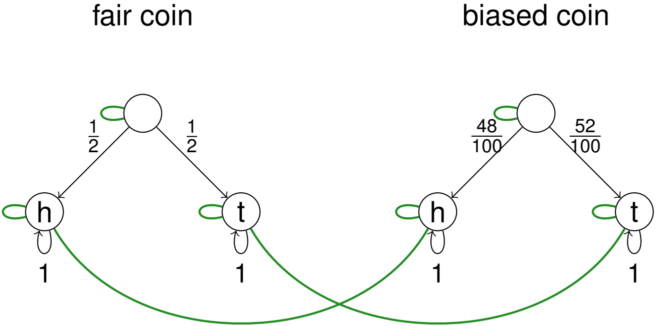
$$\begin{aligned} S^2_0 &= \{\, (s,t) \in S^2 \mid s \sim t \,\} \\ S^2_1 &= \{\, (s,t) \in S^2 \mid \ell(s) \neq \ell(t) \,\} \\ S^2_? &= S^2 \setminus (S^2_0 \cup S^2_1) \end{aligned}$$



Distance Zero and One

The set $S^2 = S \times S$ is partitioned:

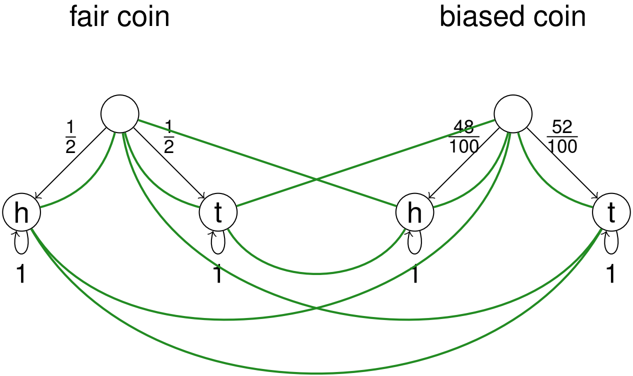
$$
\begin{aligned}
S^2_0 &= \{ (s,t) \in S^2 \mid s \sim t \} \\
S^2_1 &= \{ (s,t) \in S^2 \mid \ell(s) \neq \ell(t) \} \\
S^2_? &= S^2 \setminus (S^2_0 \cup S^2_1)
\end{aligned}
$$



Distance Zero and One

The set $S^2 = S \times S$ is partitioned:

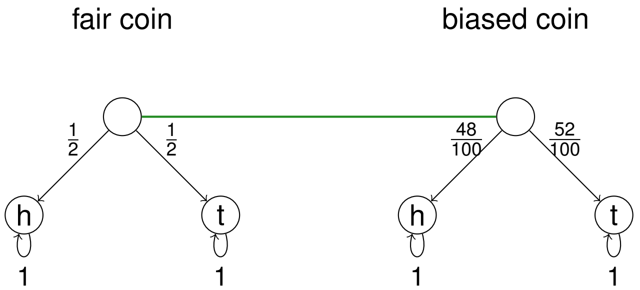
$$
\begin{aligned}
S^2_0 &= \{ (s,t) \in S^2 \mid s \sim t \} \\
S^2_1 &= \{ (s,t) \in S^2 \mid \ell(s) \neq \ell(t) \} \\
S^2_? &= S^2 \setminus (S^2_0 \cup S^2_1)
\end{aligned}
$$



Distance Zero and One

The set $S^2 = S \times S$ is partitioned:

$$
\begin{aligned}
S_0^2 &= \{ (s,t) \in S^2 \mid s \sim t \} \\
S_1^2 &= \{ (s,t) \in S^2 \mid \ell(s) \neq \ell(t) \} \\
S_?^2 &= S^2 \setminus (S_0^2 \cup S_1^2)
\end{aligned}
$$



Distance Zero and One

The set $S^2 = S \times S$ is partitioned:

$$\begin{array}{rcl} S^2_0 & = & \{\, (s,t) \in S^2 \mid s \sim t \,\} \\ S^2_1 & = & \{\, (s,t) \in S^2 \mid \ell(s) \neq \ell(t) \,\} \\ S^2_? & = & S^2 \setminus (S^2_0 \cup S^2_1) \end{array}$$

Theorem (DGJP 1999)

$S^2_0 = D_0 = \{\, (s,t) \in S^2 \mid \mathrm{lfp}(\Delta)(s,t) = 0 \,\}$.

Distance Zero and One

The set $S^2 = S \times S$ is partitioned:

$$
\begin{array}{rcl}
S^2_0 & = & \{\, (s,t) \in S^2 \mid s \sim t \,\} \\
S^2_1 & = & \{\, (s,t) \in S^2 \mid \ell(s) \neq \ell(t) \,\} \\
S^2_? & = & S^2 \setminus (S^2_0 \cup S^2_1)
\end{array}
$$

Theorem (DGJP 1999)

$S^2_0 = D_0 = \{\, (s,t) \in S^2 \mid \mathrm{lfp}(\Delta)(s,t) = 0 \,\}$.

Proposition

$S^2_1 \subseteq D_1 = \{\, (s,t) \in S^2 \mid \mathrm{lfp}(\Delta)(s,t) = 1 \,\}$.

Definition

The function $\Gamma : 2^{S \times S} \to 2^{S \times S}$ is defined by

$$\Gamma(X) = S^2_1 \cup \{ (s, t) \in S^2_? \mid \forall c \in \mathcal{C}(\tau(s), \tau(t)) : \text{support}(c) \subseteq X \}.$$

Distance One

Definition

The function $\Gamma : 2^{S\times S} \to 2^{S\times S}$ is defined by

$$\Gamma(X) = S_1^2 \cup \{\, (s,t) \in S_?^2 \mid \forall c \in \mathcal{C}(\tau(s),\tau(t)) : \text{support}(c) \subseteq X \,\}.$$

Proposition

Γ is a monotone function from the complete lattice $2^{S\times S}$ to itself.

Distance One

Definition

The function $\Gamma : 2^{S \times S} \to 2^{S \times S}$ is defined by

$$\Gamma(X) = S_1^2 \cup \{\, (s,t) \in S_?^2 \mid \forall c \in \mathcal{C}(\tau(s), \tau(t)) : \text{support}(c) \subseteq X \,\}.$$

Proposition

Γ is a monotone function from the complete lattice $2^{S \times S}$ to itself.

Corollary

Γ has a greatest fixed point, denoted gfp(Γ).

Distance One

Definition

The function $\Gamma : 2^{S \times S} \to 2^{S \times S}$ is defined by

$$\Gamma(X) = S_1^2 \cup \{ (s,t) \in S_?^2 \mid \forall c \in \mathcal{C}(\tau(s), \tau(t)) : \text{support}(c) \subseteq X \}.$$

Proposition

Γ is a monotone function from the complete lattice $2^{S \times S}$ to itself.

Corollary

Γ has a greatest fixed point, denoted $\text{gfp}(\Gamma)$.

Theorem

$D_1 = \text{gfp}(\Gamma)$.

Distance smaller than One

Definition

The function $\mathsf{L} : 2^{S \times S} \to 2^{S \times S}$ is defined by

$$\begin{aligned}
&\mathsf{L}(X) \\
&= S^2 \setminus \Gamma(S^2 \setminus X) \\
&= S^2_0 \cup \{\, (s,t) \in S^2_? \mid \exists c \in \mathcal{C}(\tau(s), \tau(t)) : \mathrm{support}(c) \not\subseteq S^2 \setminus X \,\} \\
&= S^2_0 \cup \{\, (s,t) \in S^2_? \mid \exists c \in \mathcal{C}(\tau(s), \tau(t)) : \mathrm{support}(c) \cap X \neq \emptyset \,\}.
\end{aligned}$$

Distance smaller than One

Definition

The function $\mathsf{L} : 2^{S \times S} \to 2^{S \times S}$ is defined by

$$\begin{aligned}
\mathsf{L}(X) \\
&= S^2 \setminus \Gamma(S^2 \setminus X) \\
&= S_0^2 \cup \{\, (s,t) \in S_?^2 \mid \exists c \in \mathcal{C}(\tau(s), \tau(t)) : \mathrm{support}(c) \not\subseteq S^2 \setminus X \,\} \\
&= S_0^2 \cup \{\, (s,t) \in S_?^2 \mid \exists c \in \mathcal{C}(\tau(s), \tau(t)) : \mathrm{support}(c) \cap X \neq \emptyset \,\}.
\end{aligned}$$

Proposition

$\mathrm{gfp}(\Gamma) = S^2 \setminus \mathrm{lfp}(\mathsf{L})$.

Distance smaller than One

Definition

The function $\mathrm{L} : 2^{S \times S} \to 2^{S \times S}$ is defined by

$$\begin{aligned}
\mathrm{L}(X) & \\
&= S^2 \setminus \Gamma(S^2 \setminus X) \\
&= S_0^2 \cup \{\, (s,t) \in S_?^2 \mid \exists c \in \mathcal{C}(\tau(s), \tau(t)) : \mathrm{support}(c) \not\subseteq S^2 \setminus X \,\} \\
&= S_0^2 \cup \{\, (s,t) \in S_?^2 \mid \exists c \in \mathcal{C}(\tau(s), \tau(t)) : \mathrm{support}(c) \cap X \neq \emptyset \,\}.
\end{aligned}$$

Proposition

$\mathrm{gfp}(\Gamma) = S^2 \setminus \mathrm{lfp}(\mathrm{L})$.

Proposition

$\exists c \in \mathcal{C}(\tau(s), \tau(t)) : \mathrm{support}(c) \cap X \neq \emptyset$
iff
$\exists (u,v) \in X : \tau(s)(u) > 0 \wedge \tau(t)(v) > 0$.

Distance smaller than One

Definition

The directed graph $G = \langle V, E \rangle$ is defined by

- $V = S^2$ and
- $E = \{ \langle (s, t), (u, v) \rangle \mid \tau(s)(u) > 0 \wedge \tau(t)(v) > 0 \}$.

Distance smaller than One

Definition

The directed graph $G = \langle V, E \rangle$ is defined by

- $V = S^2$ and
- $E = \{ \langle (s, t), (u, v) \rangle \mid \tau(s)(u) > 0 \wedge \tau(t)(v) > 0 \}$.

Proposition

$\text{lfp}(L) = \{ (u, v) \mid (u, v)$ is reachable from (s, t) with $s \sim t$ in $G \}$.

Distance smaller than One

Definition

The directed graph $G = \langle V, E \rangle$ is defined by

- $V = S^2$ and
- $E = \{ \langle (s,t), (u,v) \rangle \mid \tau(s)(u) > 0 \wedge \tau(t)(v) > 0 \}$.

Proposition

$\text{lfp}(L) = \{ (u,v) \mid (u,v)$ is reachable from (s,t) with $s \sim t$ in $G \}$.

Proposition

lfp(L) can be computed in $O(n^2 + m^2)$.

Distance smaller than One

Definition

The directed graph $G = \langle V, E \rangle$ is defined by

- $V = S^2$ and
- $E = \{ \langle (s,t), (u,v) \rangle \mid \tau(s)(u) > 0 \wedge \tau(t)(v) > 0 \}$.

Proposition

$\text{lfp}(L) = \{ (u,v) \mid (u,v)$ is reachable from (s,t) with $s \sim t$ in $G \}$.

Proposition

lfp(L) can be computed in $O(n^2 + m^2)$.

Proof

G has n^2 vertices and m^2 edges. Breadth first search, with the queue initially containing S_0^2, traverses all vertices in lfp(L) and takes $O(n^2 + m^2)$.

Conclusion

Distance one can be decided in $O(n^2 + m^2)$.

- New algorithm to compute distances.
- New polynomial time algorithm to decide if there are any non-trivial distances.
- New algorithm to compute all distances smaller than a given ϵ.