Bisimulation, Logics and Metrics for Labelled Markov Processes

Prakash Panangaden
School of Computer Science
McGill University
and
Research Visitor: The Alan Turing Institute

Logic for Data Science Seminar
The Alan Turing Institute
8th June 2018

Summary of Results

- Probabilistic bisimulation can be defined for continuous state-space systems. [LICS 97]

Summary of Results

- Probabilistic bisimulation can be defined for continuous state-space systems. [LICS 97]
- Logical characterization. [LICS 98, Info and Comp 02]

Summary of Results

- Probabilistic bisimulation can be defined for continuous state-space systems. [LICS 97]
- Logical characterization. [LICS 98, Info and Comp 02]
- Metric analogue of bisimulation. [CONCUR 99, TCS 04]

Summary of Results

- Probabilistic bisimulation can be defined for continuous state-space systems. [LICS 97]
- Logical characterization. [LICS 98, Info and Comp 02]
- Metric analogue of bisimulation. [CONCUR 99, TCS 04]
- Approximation of LMPs. [LICS00, Info and Comp 03, CONCUR 05, ICALP 09]

Summary of Results

- Probabilistic bisimulation can be defined for continuous state-space systems. [LICS 97]
- Logical characterization. [LICS 98, Info and Comp 02]
- Metric analogue of bisimulation. [CONCUR 99, TCS 04]
- Approximation of LMPs. [LICS00, Info and Comp 03, CONCUR 05, ICALP 09]
- Weak bisimulation. [LICS 02, CONCUR 02]

Summary of Results

- Probabilistic bisimulation can be defined for continuous state-space systems. [LICS 97]
- Logical characterization. [LICS 98, Info and Comp 02]
- Metric analogue of bisimulation. [CONCUR 99, TCS 04]
- Approximation of LMPs. [LICS00, Info and Comp 03, CONCUR 05, ICALP 09]
- Weak bisimulation. [LICS 02, CONCUR 02]
- Real time. [QEST 04, JLAP 03,LMCS 06]

Summary of Results

- Probabilistic bisimulation can be defined for continuous state-space systems. [LICS 97]
- Logical characterization. [LICS 98, Info and Comp 02]
- Metric analogue of bisimulation. [CONCUR 99, TCS 04]
- Approximation of LMPs. [LICS00, Info and Comp 03, CONCUR 05, ICALP 09]
- Weak bisimulation. [LICS 02, CONCUR 02]
- Real time. [QEST 04, JLAP 03,LMCS 06]
- Metrics for MDPs [UAI 04, 05, QEST 12, AAAI 15, NIPS 15]

Summary of Results

- Probabilistic bisimulation can be defined for continuous state-space systems. [LICS 97]
- Logical characterization. [LICS 98, Info and Comp 02]
- Metric analogue of bisimulation. [CONCUR 99, TCS 04]
- Approximation of LMPs. [LICS00, Info and Comp 03, CONCUR 05, ICALP 09]
- Weak bisimulation. [LICS 02, CONCUR 02]
- Real time. [QEST 04, JLAP 03,LMCS 06]
- Metrics for MDPs [UAI 04, 05, QEST 12, AAAI 15, NIPS 15]
- Duality theory [LICS 13, JACM 14, LICS 17]

Summary of Results

- Probabilistic bisimulation can be defined for continuous state-space systems. [LICS 97]
- Logical characterization. [LICS 98, Info and Comp 02]
- Metric analogue of bisimulation. [CONCUR 99, TCS 04]
- Approximation of LMPs. [LICS00, Info and Comp 03, CONCUR 05, ICALP 09]
- Weak bisimulation. [LICS 02, CONCUR 02]
- Real time. [QEST 04, JLAP 03,LMCS 06]
- Metrics for MDPs [UAI 04, 05, QEST 12, AAAI 15, NIPS 15]
- Duality theory [LICS 13, JACM 14, LICS 17]
- Quantitative equational logic [LICS 16, 17, 18]

Collaborators

- Josée Desharnais

Collaborators

- Josée Desharnais
- Radha Jagadeesan and Vineet Gupta

Collaborators

- Josée Desharnais
- Radha Jagadeesan and Vineet Gupta
- Abbas Edalat, Vincent Danos, Radu Mardare, François Laviolette

Collaborators

- Josée Desharnais
- Radha Jagadeesan and Vineet Gupta
- Abbas Edalat, Vincent Danos, Radu Mardare, François Laviolette
- Philippe Chaput, Florence Clerc, Nathanaël Fijalkow, Bartek Klin

Collaborators

- Josée Desharnais
- Radha Jagadeesan and Vineet Gupta
- Abbas Edalat, Vincent Danos, Radu Mardare, François Laviolette
- Philippe Chaput, Florence Clerc, Nathanaël Fijalkow, Bartek Klin
- Doina Precup, Norm Ferns, Gheorghe Comanici, Giorgio Bacci

Collaborators

- Josée Desharnais
- Radha Jagadeesan and Vineet Gupta
- Abbas Edalat, Vincent Danos, Radu Mardare, François Laviolette
- Philippe Chaput, Florence Clerc, Nathanaël Fijalkow, Bartek Klin
- Doina Precup, Norm Ferns, Gheorghe Comanici, Giorgio Bacci
- Dexter Kozen, Kim Larsen, Gordon Plotkin

The definition

- A set of states S,

The definition

- A set of states S,
- a set of *labels* or *actions*, L or $\mathcal{A}$ and

The definition

- A set of states S,
- a set of *labels* or *actions*, L or $\mathcal{A}$ and
- a transition relation $\subseteq S \times \mathcal{A} \times S$, usually written

$$\rightarrow_a \subseteq S \times S.$$

The transitions could be indeterminate (nondeterministic).

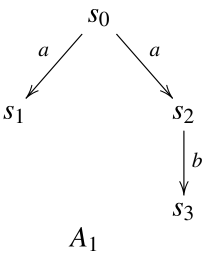
The definition

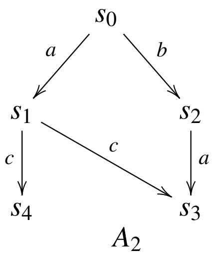
- A set of states S,
- a set of *labels* or *actions*, L or $\mathcal{A}$ and
- a transition relation $\subseteq S \times \mathcal{A} \times S$, usually written

$$\rightarrow_a \subseteq S \times S.$$

  The transitions could be indeterminate (nondeterministic).
- We write $s \xrightarrow{a} s'$ for $(s, s') \in \rightarrow_a$.

A simple example





Bisimulation

s and t are states of a labelled transition system. We say s is **bisimilar** to t – written $s \sim t$ – if

$$s \xrightarrow{a} s' \Rightarrow \exists t' \text{ such that } t \xrightarrow{a} t' \text{ and } s' \sim t'$$

and

$$t \xrightarrow{a} t' \Rightarrow \exists s' \text{ such that } s \xrightarrow{a} s' \text{ and } s' \sim t'.$$

Bisimulation relations

- Define *a* (note the indefinite article) bisimulation relation R to be an equivalence relation on S such that

Bisimulation relations

- Define *a* (note the indefinite article) bisimulation relation R to be an equivalence relation on S such that
-
$$sRt \text{ means } \forall a, s \xrightarrow{a} s' \Rightarrow \exists t', t \xrightarrow{a} t' \text{ with } s'Rt'$$

  and vice versa.

Bisimulation relations

- Define *a* (note the indefinite article) bisimulation relation R to be an equivalence relation on S such that
-
$$sRt \text{ means } \forall a, s \xrightarrow{a} s' \Rightarrow \exists t', t \xrightarrow{a} t' \text{ with } s'Rt'$$
  and vice versa.
- This is not circular; it is a condition on R.

Bisimulation relations

- Define *a* (note the indefinite article) bisimulation relation R to be an equivalence relation on S such that
- $$sRt \text{ means } \forall a, s \xrightarrow{a} s' \Rightarrow \exists t', t \xrightarrow{a} t' \text{ with } s'Rt'$$

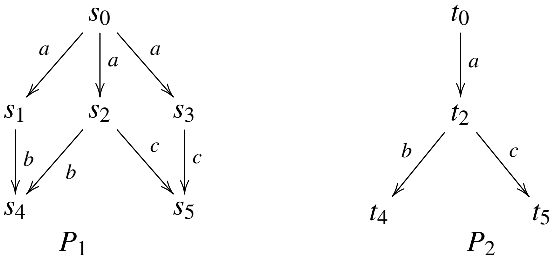
  and vice versa.
- This is not circular; it is a condition on R.
- We define $s \sim t$ if there is *some* bisimulation relation R with sRt.

Bisimulation relations

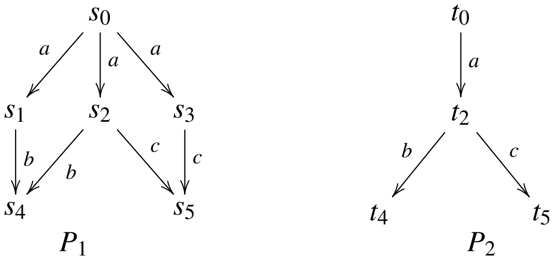
- Define *a* (note the indefinite article) bisimulation relation R to be an equivalence relation on S such that
- $$sRt \text{ means } \forall a, s \xrightarrow{a} s' \Rightarrow \exists t', t \xrightarrow{a} t' \text{ with } s'Rt'$$

  and vice versa.
- This is not circular; it is a condition on R.
- We define $s \sim t$ if there is *some* bisimulation relation R with sRt.
- This is the version that is used most often.

An example

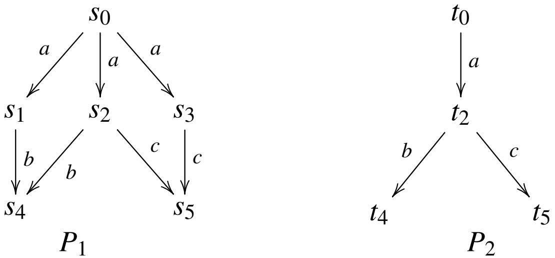
- 

An example

- 

- Here s_0 and t_0 are not bisimilar.

An example

- 

- Here s_0 and t_0 are not bisimilar.
- However s_0 and t_0 can *simulate each other*!

How do we know that two processes are **not** bisimilar?

- Define a logic as follows:

How do we know that two processes are **not** bisimilar?

- Define a logic as follows:
- $$\phi ::== \mathsf{T} | \neg\phi | \phi_1 \wedge \phi_2 | \langle a \rangle \phi$$

How do we know that two processes are **not** bisimilar?

- Define a logic as follows:
-
$$\phi ::== \mathsf{T} | \neg\phi | \phi_1 \wedge \phi_2 | \langle a \rangle \phi$$
- $s \models \langle a \rangle \phi$ means that $s \xrightarrow{a} s'$ and $t \models \phi$.

How do we know that two processes are **not** bisimilar?

- Define a logic as follows:
- $$\phi ::== \mathsf{T} | \neg\phi | \phi_1 \wedge \phi_2 | \langle a \rangle \phi$$
- $s \models \langle a \rangle \phi$ means that $s \xrightarrow{a} s'$ and $t \models \phi$.
- We can define a dual to $\langle\rangle$ (written []) by using negation.

How do we know that two processes are **not** bisimilar?

- Define a logic as follows:
- $$\phi ::== \mathsf{T} | \neg\phi | \phi_1 \wedge \phi_2 | \langle a \rangle \phi$$
- $s \models \langle a \rangle \phi$ means that $s \xrightarrow{a} s'$ and $t \models \phi$.
- We can define a dual to $\langle\rangle$ (written $[]$) by using negation.
- $s \models [a]\phi$ means that *if* s can do an a the resulting state must satisfy ϕ.

Examples of HM Logic

- T is satisfied by any process, F is not satisfied by any process.

Examples of HM Logic

- T is satisfied by any process, F is not satisfied by any process.
- $s \models \langle a \rangle T$ means s can do an a action.

Examples of HM Logic

- T is satisfied by any process, F is not satisfied by any process.
- $s \models \langle a \rangle T$ means s can do an a action.
- $s \models \neg\langle a \rangle T$ or $s \models [a]F$ means s cannot do an a action.

Examples of HM Logic

- T is satisfied by any process, F is not satisfied by any process.
- $s \models \langle a \rangle T$ means s can do an a action.
- $s \models \neg \langle a \rangle T$ or $s \models [a]F$ means s cannot do an a action.
- $s \models \langle a \rangle (\langle b \rangle T)$ means that s can do an a *and then* do a b.

The logical characterization theorem

- Two processes are bisimilar if and only if they satisfy the same formulas of HM logic.

The logical characterization theorem

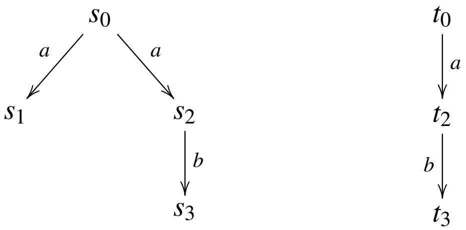
- Two processes are bisimilar if and only if they satisfy the same formulas of HM logic.
- Basic assumption: the processes are finitely-branching (otherwise you need infinitary conjunctions).

The logical characterization theorem

- Two processes are bisimilar if and only if they satisfy the same formulas of HM logic.
- Basic assumption: the processes are finitely-branching (otherwise you need infinitary conjunctions).
- To show that two processes are not bisimilar find a formula on which they disagree.

The role of negation

- Consider the processes below:



The role of negation

- Consider the processes below:

$$
\begin{tikzcd}
 & s_0 \arrow[dl, "a"'] \arrow[dr, "a"] & \\
s_1 & & s_2 \arrow[d, "b"] \\
 & & s_3
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
t_0 \arrow[d, "a"] \\
t_2 \arrow[d, "b"'] \\
t_3
\end{tikzcd}
$$

- $s_0 \models \langle a \rangle \neg \langle b \rangle T$ but t_0 does not.

The role of negation

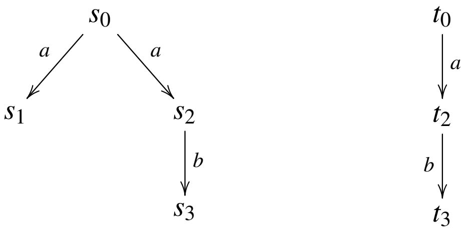
- Consider the processes below:

$$
\begin{tikzcd}
 & s_0 \arrow[dl, "a"'] \arrow[dr, "a"] & & & t_0 \arrow[d, "a"] \\
s_1 & & s_2 \arrow[d, "b"] & & t_2 \arrow[d, "b"'] \\
 & & s_3 & & t_3
\end{tikzcd}
$$

- $s_0 \models \langle a \rangle \neg \langle b \rangle T$ but t_0 does not.
- s_0 and t_0 agree on all formulas *without negation*.

The role of negation

- Consider the processes below:



- $s_0 \models \langle a \rangle \neg \langle b \rangle T$ but t_0 does not.
- s_0 and t_0 agree on all formulas *without negation*.
- Note that $[a]$ has an implicit negation.

Discrete probabilistic transition systems

- Just like a labelled transition system with probabilities associated with the transitions.

Discrete probabilistic transition systems

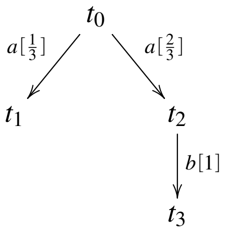
- Just like a labelled transition system with probabilities associated with the transitions.
- $$(S, \mathsf{L}, \forall a \in \mathsf{L}\ T_a : S \times S \longrightarrow [0, 1])$$

Discrete probabilistic transition systems

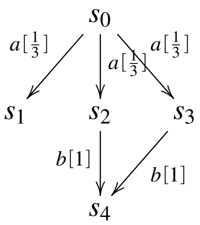
- Just like a labelled transition system with probabilities associated with the transitions.
- $$(S, \mathsf{L}, \forall a \in \mathsf{L}\ T_a : S \times S \longrightarrow [0,1])$$
- The model is *reactive*: All probabilistic data is *internal* - no probabilities associated with environment behaviour.

Bisimulation for PTS: Larsen and Skou

- Consider



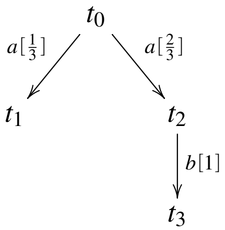
P_1



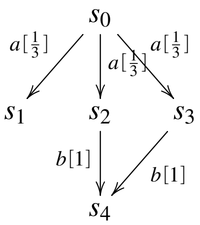
P_2

Bisimulation for PTS: Larsen and Skou

- Consider



P_1

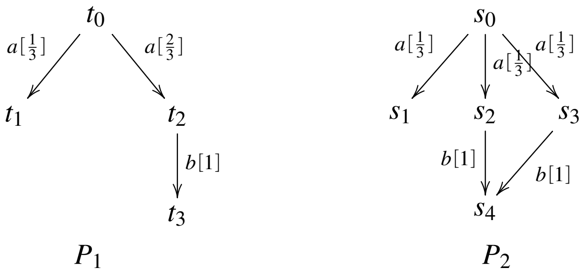


P_2

- Should s_0 and t_0 be bisimilar?

Bisimulation for PTS: Larsen and Skou

- Consider



- Should s_0 and t_0 be bisimilar?
- Yes, but we need to add the probabilities.

The Official Definition

- Let $\mathcal{S} = (S, \mathsf{L}, T_a)$ be a PTS. An equivalence relation R on S is a **bisimulation** if whenever sRs', with $s, s' \in S$, we have that for all $a \in \mathcal{A}$ and every R-equivalence class, A, $T_a(s, A) = T_a(s', A)$.

The Official Definition

- Let $\mathcal{S} = (S, \mathsf{L}, T_a)$ be a PTS. An equivalence relation R on S is a **bisimulation** if whenever sRs', with $s, s' \in S$, we have that for all $a \in \mathcal{A}$ and every R-equivalence class, A, $T_a(s, A) = T_a(s', A)$.
- The notation $T_a(s, A)$ means "the probability of starting from s and jumping to a state in the set A."

The Official Definition

- Let $\mathcal{S} = (S, \mathsf{L}, T_a)$ be a PTS. An equivalence relation R on S is a **bisimulation** if whenever sRs', with $s, s' \in S$, we have that for all $a \in \mathcal{A}$ and every R-equivalence class, A, $T_a(s, A) = T_a(s', A)$.
- The notation $T_a(s, A)$ means "the probability of starting from s and jumping to a state in the set A."
- Two states are bisimilar if there is some bisimulation relation R relating them.

What are labelled Markov processes?

- Labelled Markov processes are probabilistic versions of labelled transition systems. Labelled transition systems where the final state is governed by a probability distribution - no other indeterminacy.

What are labelled Markov processes?

- Labelled Markov processes are probabilistic versions of labelled transition systems. Labelled transition systems where the final state is governed by a probability distribution - no other indeterminacy.
- All probabilistic data is *internal* - no probabilities associated with environment behaviour.

What are labelled Markov processes?

- Labelled Markov processes are probabilistic versions of labelled transition systems. Labelled transition systems where the final state is governed by a probability distribution - no other indeterminacy.
- All probabilistic data is *internal* - no probabilities associated with environment behaviour.
- We observe the interactions - not the internal states.

What are labelled Markov processes?

- Labelled Markov processes are probabilistic versions of labelled transition systems. Labelled transition systems where the final state is governed by a probability distribution - no other indeterminacy.
- All probabilistic data is *internal* - no probabilities associated with environment behaviour.
- We observe the interactions - not the internal states.
- **In general, the state space of a labelled Markov process may be a *continuum*.**

Motivation

Model and reason about systems with *continuous* state spaces or continuous time evolution or both.

- hybrid control systems; e.g. flight management systems.

Motivation

Model and reason about systems with *continuous* state spaces or continuous time evolution or both.

- hybrid control systems; e.g. flight management systems.
- telecommunication systems with spatial variation; e.g. cell phones

Motivation

Model and reason about systems with *continuous* state spaces or continuous time evolution or both.

- hybrid control systems; e.g. flight management systems.
- telecommunication systems with spatial variation; e.g. cell phones
- performance modelling,

Motivation

Model and reason about systems with *continuous* state spaces or continuous time evolution or both.

- hybrid control systems; e.g. flight management systems.
- telecommunication systems with spatial variation; e.g. cell phones
- performance modelling,
- continuous time systems,

Motivation

Model and reason about systems with *continuous* state spaces or continuous time evolution or both.

- hybrid control systems; e.g. flight management systems.
- telecommunication systems with spatial variation; e.g. cell phones
- performance modelling,
- continuous time systems,
- probabilistic process algebra with recursion.

The Need for Measure Theory

- Basic fact: There are subsets of $\mathbf{R}$ for which no sensible notion of size can be defined.

The Need for Measure Theory

- Basic fact: There are subsets of $\mathbf{R}$ for which no sensible notion of size can be defined.
- More precisely, there is no translation-invariant measure defined on all the subsets of the reals.

The Need for Measure Theory

- Basic fact: There are subsets of $\mathbf{R}$ for which no sensible notion of size can be defined.
- More precisely, there is no translation-invariant measure defined on all the subsets of the reals.
- Actually there is if you only require finite additivity.

Stochastic Kernels

- A *stochastic kernel* (Markov kernel) is a function $h : S \times \Sigma \longrightarrow [0, 1]$ with (a) $h(s, \cdot) : \Sigma \longrightarrow [0, 1]$ a (sub)probability measure and (b) $h(\cdot, A) : X \longrightarrow [0, 1]$ a measurable function.

Stochastic Kernels

- A *stochastic kernel* (Markov kernel) is a function $h : S \times \Sigma \longrightarrow [0, 1]$ with (a) $h(s, \cdot) : \Sigma \longrightarrow [0, 1]$ a (sub)probability measure and (b) $h(\cdot, A) : X \longrightarrow [0, 1]$ a measurable function.
- Though apparently asymmetric, these are the stochastic analogues of binary relations

Stochastic Kernels

- A *stochastic kernel* (Markov kernel) is a function $h : S \times \Sigma \longrightarrow [0, 1]$ with (a) $h(s, \cdot) : \Sigma \longrightarrow [0, 1]$ a (sub)probability measure and (b) $h(\cdot, A) : X \longrightarrow [0, 1]$ a measurable function.
- Though apparently asymmetric, these are the stochastic analogues of binary relations
- and the uncountable generalization of a matrix.

Formal Definition of LMPs

- An LMP is a tuple $(S, \Sigma, \mathsf{L}, \forall\alpha \in \mathsf{L}.\tau_\alpha)$ where $\tau_\alpha : S \times \Sigma \longrightarrow [0, 1]$ is a *transition probability* function such that

Formal Definition of LMPs

- An LMP is a tuple $(S, \Sigma, \mathsf{L}, \forall\alpha \in \mathsf{L}.\tau_\alpha)$ where $\tau_\alpha : S \times \Sigma \longrightarrow [0, 1]$ is a *transition probability* function such that
- $\forall s : S.\lambda A : \Sigma.\tau_\alpha(s, A)$ is a subprobability measure
  and
  $\forall A : \Sigma.\lambda s : S.\tau_\alpha(s, A)$ is a measurable function.

Larsen-Skou Bisimulation

- Let $\mathcal{S} = (S, i, \Sigma, \tau)$ be a labelled Markov process. An equivalence relation R on S is a **bisimulation** if whenever sRs', with $s, s' \in S$, we have that for all $a \in \mathcal{A}$ and every R-closed measurable set $A \in \Sigma$, $\tau_a(s, A) = \tau_a(s', A)$.
  Two states are bisimilar if they are related by a bisimulation relation.

Larsen-Skou Bisimulation

- Let $\mathcal{S} = (S, i, \Sigma, \tau)$ be a labelled Markov process. An equivalence relation R on S is a **bisimulation** if whenever sRs', with $s, s' \in S$, we have that for all $a \in \mathcal{A}$ and every R-closed measurable set $A \in \Sigma$, $\tau_a(s, A) = \tau_a(s', A)$.
  Two states are bisimilar if they are related by a bisimulation relation.
- Can be extended to bisimulation between two different **LMP**s.

A game for bisimulation

- Two players: spoiler (S) and duplicator (D).

A game for bisimulation

- Two players: spoiler (S) and duplicator (D).
- Duplicator claims x, y are bisimilar.

A game for bisimulation

- Two players: spoiler (S) and duplicator (D).
- Duplicator claims x, y are bisimilar.
- Spoiler exhibits a set C and says C is bisimulation-closed and that $\tau(x, C) \neq \tau(y, C)$. Assume that the inequality holds; it is easy to check.

A game for bisimulation

- Two players: spoiler (S) and duplicator (D).
- Duplicator claims x, y are bisimilar.
- Spoiler exhibits a set C and says C is bisimulation-closed and that $\tau(x, C) \neq \tau(y, C)$. Assume that the inequality holds; it is easy to check.
- Duplicator responds by saying that C is not bisimulation-closed and that exhibits $x' \in C$ and $y' \notin C$ and claims that x', y' are bisimilar.

A game for bisimulation

- Two players: spoiler (S) and duplicator (D).
- Duplicator claims x, y are bisimilar.
- Spoiler exhibits a set C and says C is bisimulation-closed and that $\tau(x, C) \neq \tau(y, C)$. Assume that the inequality holds; it is easy to check.
- Duplicator responds by saying that C is not bisimulation-closed and that exhibits $x' \in C$ and $y' \notin C$ and claims that x', y' are bisimilar.
- A player loses when he or she cannot make a move. Note that if C is all of the state space, duplicator loses. Duplicator wins if she can play forever.

A game for bisimulation

- Two players: spoiler (S) and duplicator (D).
- Duplicator claims x, y are bisimilar.
- Spoiler exhibits a set C and says C is bisimulation-closed and that $\tau(x, C) \neq \tau(y, C)$. Assume that the inequality holds; it is easy to check.
- Duplicator responds by saying that C is not bisimulation-closed and that exhibits $x' \in C$ and $y' \notin C$ and claims that x', y' are bisimilar.
- A player loses when he or she cannot make a move. Note that if C is all of the state space, duplicator loses. Duplicator wins if she can play forever.
- We prove that x is bisimilar to y iff Duplicator has a winning strategy starting from (x, y).

Logical Characterization

- $$\mathcal{L} ::== \mathsf{T} | \phi_1 \wedge \phi_2 | \langle a \rangle_q \phi$$

Logical Characterization

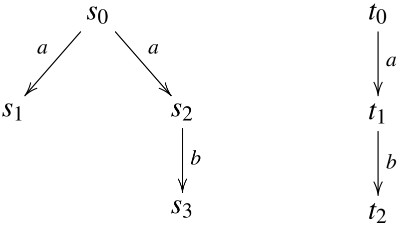
- $$\mathcal{L} ::== \mathsf{T} | \phi_1 \wedge \phi_2 | \langle a \rangle_q \phi$$
- We say $s \models \langle a \rangle_q \phi$ iff

$$\exists A \in \Sigma.(\forall s' \in A. s' \models \phi) \wedge (\tau_a(s, A) > q).$$

Logical Characterization

- $$\mathcal{L} ::== \mathsf{T} | \phi_1 \wedge \phi_2 | \langle a \rangle_q \phi$$
- We say $s \models \langle a \rangle_q \phi$ iff
$$\exists A \in \Sigma.(\forall s' \in A.s' \models \phi) \wedge (\tau_a(s, A) > q).$$
- Two systems are bisimilar iff they obey the same formulas of $\mathcal{L}$. [DEP 1998 LICS, I and C 2002]

That cannot be right?



Two processes that cannot be distinguished without negation.
The formula that distinguishes them is $\langle a \rangle (\neg \langle b \rangle \top)$.

But it is!

- $$\begin{tikzcd}
 & s_0 \arrow[dl, "{a[p]}"'] \arrow[dr, "{a[q]}"] & & & t_0 \arrow[d, "{a[r]}"] \\
s_1 & & s_2 \arrow[d, "b"] & & t_1 \arrow[d, "b"] \\
 & & s_3 & & t_2
\end{tikzcd}$$

  We add probabilities to the transitions.

But it is!

- $$\begin{tikzcd} & s_0 \arrow[dl, "{a[p]}"'] \arrow[dr, "{a[q]}"] & \\ s_1 & & s_2 \arrow[d, "b"] \\ & & s_3 \end{tikzcd} \qquad \begin{tikzcd} t_0 \arrow[d, "{a[r]}"] \\ t_1 \arrow[d, "b"] \\ t_2 \end{tikzcd}$$

  We add probabilities to the transitions.
- If $p + q < r$ or $p + q > r$ we can easily distinguish them.

But it is!

- $$\begin{tikzcd} & s_0 \arrow[dl, "{a[p]}"'] \arrow[dr, "{a[q]}"] & & & t_0 \arrow[d, "{a[r]}"] \\ s_1 & & s_2 \arrow[d, "b"] & & t_1 \arrow[d, "b"] \\ & & s_3 & & t_2 \end{tikzcd}$$

  We add probabilities to the transitions.
- If $p+q<r$ or $p+q>r$ we can easily distinguish them.
- If $p+q=r$ and $p>0$ then $q<r$ so $\langle a\rangle_r\langle b\rangle_1\top$ distinguishes them.

A metric-based approximate viewpoint

- Move from equality between processes to distances between processes (Jou and Smolka 1990).

A metric-based approximate viewpoint

- Move from equality between processes to distances between processes (Jou and Smolka 1990).
- Formalize distance as a metric:

$$d(s,s) = 0, d(s,t) = d(t,s), d(s,u) \leqslant d(s,t) + d(t,u).$$

Quantitative analogue of an equivalence relation.

A metric-based approximate viewpoint

- Move from equality between processes to distances between processes (Jou and Smolka 1990).
- Formalize distance as a metric:

$$d(s, s) = 0, d(s, t) = d(t, s), d(s, u) \leqslant d(s, t) + d(t, u).$$

  Quantitative analogue of an equivalence relation.
- Quantitative measurement of the distinction between processes.

Criteria on Metrics

- Soundness:

$$d(s,t) = 0 \Leftrightarrow \; s,t \text{ are bisimilar}$$

Criteria on Metrics

- Soundness:
$$d(s,t) = 0 \Leftrightarrow \; s,t \text{ are bisimilar}$$
- Stability of distance under temporal evolution:“Nearby states stay close *forever*.”

Criteria on Metrics

- Soundness:
$$d(s,t) = 0 \Leftrightarrow \ s,t \text{ are bisimilar}$$
- Stability of distance under temporal evolution:“Nearby states stay close *forever*.”
- Metrics should be computable (efficiently?).

Bisimulation Recalled

Let R be an equivalence relation. R is a bisimulation if: $s\ R\ t$ if:

$$(s \longrightarrow P) \Rightarrow [t \longrightarrow Q, P =_R Q]$$

$$(t \longrightarrow Q) \Rightarrow [s \longrightarrow P, P =_R Q]$$

where $P =_R Q$ if

$$(\forall R-\text{closed } E)\ P(E) = Q(E)$$

A putative definition of a metric analogue of bisimulation

- m is a metric-bisimulation if: $m(s, t) < \epsilon \Rightarrow$:

$$s \longrightarrow P \Rightarrow \ t \longrightarrow Q, \quad m(P, Q) < \epsilon$$

$$t \longrightarrow Q \Rightarrow \ s \longrightarrow P, \ \ m(P, Q) < \epsilon$$

A putative definition of a metric analogue of bisimulation

- m is a metric-bisimulation if: $m(s, t) < \epsilon \Rightarrow$:

$$s \longrightarrow P \Rightarrow \; t \longrightarrow Q, \quad m(P, Q) < \epsilon$$

$$t \longrightarrow Q \Rightarrow \; s \longrightarrow P, \;\; m(P, Q) < \epsilon$$

- Problem: what is $m(P, Q)$? — Type mismatch!!

A putative definition of a metric analogue of bisimulation

- m is a metric-bisimulation if: $m(s,t) < \epsilon \Rightarrow$:

$$s \longrightarrow P \Rightarrow \; t \longrightarrow Q, \quad m(P,Q) < \epsilon$$

$$t \longrightarrow Q \Rightarrow \; s \longrightarrow P, \quad m(P,Q) < \epsilon$$

- Problem: what is $m(P,Q)$? — Type mismatch!!
- Need a way to lift distances from states to a distances on distributions of states.

A detour: Kantorovich metric

- Metrics on probability measures on metric spaces.

A detour: Kantorovich metric

- Metrics on probability measures on metric spaces.
- $\mathcal{M}$: 1-bounded pseudometrics on states.

A detour: Kantorovich metric

- Metrics on probability measures on metric spaces.
- $\mathcal{M}$: 1-bounded pseudometrics on states.
- $$d(\mu, \nu) = \sup_f |\int f d\mu - \int f d\nu|, f \text{ 1-Lipschitz}$$

A detour: Kantorovich metric

- Metrics on probability measures on metric spaces.
- $\mathcal{M}$: 1-bounded pseudometrics on states.
- $$d(\mu, \nu) = \sup_f |\int f d\mu - \int f d\nu|, f \text{ 1-Lipschitz}$$
- Arises in the solution of an LP problem: *transshipment*.

An LP version for Finite-State Spaces

When state space is finite: Let P, Q be probability distributions. Then:

$$m(P, Q) = \max \sum_i (P(s_i) - Q(s_i))a_i$$

subject to:

$$\forall i. 0 \leqslant a_i \leqslant 1$$
$$\forall i, j.\ a_i - a_j \leqslant m(s_i, s_j).$$

The Dual Form

- Dual form from Worrell and van Breugel:

The Dual Form

- Dual form from Worrell and van Breugel:
- $$\min \sum_{i,j} l_{ij} m(s_i, s_j) + \sum_i x_i + \sum_j y_j$$

  subject to:

  $$\begin{array}{l} \forall i.\ \sum_j l_{ij} + x_i = P(s_i) \\ \forall j.\ \sum_i l_{ij} + y_j = Q(s_j) \\ \forall i,j.\ l_{ij}, x_i, y_j \geqslant 0. \end{array}$$

The Dual Form

- Dual form from Worrell and van Breugel:
- $$\min \sum_{i,j} l_{ij} m(s_i, s_j) + \sum_i x_i + \sum_j y_j$$

  subject to:

  $$\begin{array}{l} \forall i.\ \sum_j l_{ij} + x_i = P(s_i) \\ \forall j.\ \sum_i l_{ij} + y_j = Q(s_j) \\ \forall i,j.\ l_{ij}, x_i, y_j \geqslant 0. \end{array}$$

- We prove many equations by using the primal form to show one direction and the dual to show the other.

Example

- Let $m(s, t) = r < 1$. Let $\delta_s(\delta_t)$ be the probability measure concentrated at $s(t)$. Then,

$$m(\delta_s, \delta_t) = r$$

Example

- Let $m(s,t) = r < 1$. Let $\delta_s(\delta_t)$ be the probability measure concentrated at $s(t)$. Then,

$$m(\delta_s, \delta_t) = r$$

- Upper bound from dual: Choose $l_{st} = 1$ all other $l_{ij} = 0$. Then

$$\sum_{ij} l_{ij} m(s_i, s_j) = m(s,t) = r.$$

Example

- Let $m(s,t) = r < 1$. Let $\delta_s(\delta_t)$ be the probability measure concentrated at $s(t)$. Then,

$$m(\delta_s, \delta_t) = r$$

- Upper bound from dual: Choose $l_{st} = 1$ all other $l_{ij} = 0$. Then

$$\sum_{ij} l_{ij} m(s_i, s_j) = m(s,t) = r.$$

- Lower bound from primal: Choose $a_s = 0, a_t = r$, all others to match the constraints. Then

$$\sum_{i} (\delta_t(s_i) - \delta_s(s_i)) a_i = r.$$

The Importance of the Example

We can *isometrically* embed the original space in the metric space of distributions.

Return from Detour

Summary of detour: Given a metric on states in a metric space, it can be lifted to a metric on probability distributions on states.

Metric "Bisimulation"

- m is a metric-bisimulation if: $m(s,t) < \epsilon \Rightarrow$:

$$s \longrightarrow P \Rightarrow \; t \longrightarrow Q, \quad m(P,Q) < \epsilon$$

$$t \longrightarrow Q \Rightarrow \; s \longrightarrow P, \quad m(P,Q) < \epsilon$$

Metric “Bisimulation”

- m is a metric-bisimulation if: $m(s,t) < \epsilon \Rightarrow$:

$$s \longrightarrow P \Rightarrow \; t \longrightarrow Q, \quad m(P,Q) < \epsilon$$

$$t \longrightarrow Q \Rightarrow \; s \longrightarrow P, \;\; m(P,Q) < \epsilon$$

- The required canonical metric on processes is the least such: ie. the distances are the least possible.

Metric “Bisimulation”

- m is a metric-bisimulation if: $m(s,t) < \epsilon \Rightarrow$:

$$s \longrightarrow P \Rightarrow \; t \longrightarrow Q, \quad m(P,Q) < \epsilon$$

$$t \longrightarrow Q \Rightarrow \; s \longrightarrow P, \; m(P,Q) < \epsilon$$

- The required canonical metric on processes is the least such: ie. the distances are the least possible.
- Thm: *Canonical least metric exists*. Usual fixed-point theory arguments.

What about Continuous-State Systems?

- Develop a real-valued “modal logic” based on the analogy:

Program Logic	Probabilistic Logic
State s	Distribution μ
Formula ϕ	Random Variable f
Satisfaction $s \models \phi$	$\int f \mathrm{d}\mu$

What about Continuous-State Systems?

- Develop a real-valued “modal logic” based on the analogy:

Program Logic	Probabilistic Logic
State s	Distribution μ
Formula ϕ	Random Variable f
Satisfaction $s \models \phi$	$\int f \mathrm{d}\mu$

- Define a metric based on how closely the random variables agree.

What about Continuous-State Systems?

- Develop a real-valued "modal logic" based on the analogy:

Program Logic	Probabilistic Logic
State s	Distribution μ
Formula ϕ	Random Variable f
Satisfaction $s \models \phi$	$\int f \mathrm{d}\mu$

- Define a metric based on how closely the random variables agree.
- This metric coincides with the metric based on the Kantorovich metric.

Logic and metric

- Metrics seem like a piece of real analysis, where is the logic?

Logic and metric

- Metrics seem like a piece of real analysis, where is the logic?
- Quantitative equational logic: $=_\epsilon$, approximate equality

Logic and metric

- Metrics seem like a piece of real analysis, where is the logic?
- Quantitative equational logic: $=_\epsilon$, approximate equality
- Algebras are naturally equipped with a metric

Logic and metric

- Metrics seem like a piece of real analysis, where is the logic?
- Quantitative equational logic: $=_\epsilon$, approximate equality
- Algebras are naturally equipped with a metric
- Simple equational axioms capture Kantorovich and total variation metrics.

Conclusions

- Modal logics are closely related to behavioural equivalence.

Conclusions

- Modal logics are closely related to behavioural equivalence.
- Metrics arise as a natural quantitative relaxation of logic.

Conclusions

- Modal logics are closely related to behavioural equivalence.
- Metrics arise as a natural quantitative relaxation of logic.
- Quantitative equational logic captures metrics.

Conclusions

- Modal logics are closely related to behavioural equivalence.
- Metrics arise as a natural quantitative relaxation of logic.
- Quantitative equational logic captures metrics.
- Metrics and logics: exploring links with neural nets.