Bisimulation and Simulation for Labelled Markov Processes

Prakash Panangaden\textsuperscript{1}

\textsuperscript{1}School of Computer Science
McGill University

IRIF Lecture Series Jan 2018
Overview

- Discrete probabilistic transition system.
- Labelled Markov processes: probabilistic transition systems with continuous state spaces.
- Bisimulation for LMPs.
- A game for bisimulation.
- Simulation
- Logical characterization.
Summary of Results

- Probabilistic bisimulation can be defined for continuous state-space systems. [LICS97]
- Logical characterization. [LICS98, Info and Comp 2002]
- Metric analogue of bisimulation. [CONCUR99, TCS2004]
- Approximation of LMPs. [LICS00, Info and Comp 2003]
- Weak bisimulation. [LICS02, CONCUR02]
- Real time. [QEST 2004, JLAP 2003, LMCS 2006]
- Event bisimulation [Info and Comp 2006]
- Approximation by averaging [ICALP 2009, JACM 2014]
- Duality [JACM 2014, LICS 2013, 2017]
- Quantitative equational logic [LICS 2016, 2017]
Collaborators

Labelled Transition System

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

$$\rightarrow a \subseteq S \times S.$$  

The transitions could be indeterminate (nondeterministic).
A *discrete-time* Markov chain is a finite set $S$ (the state space) together with a transition probability function $T : S \times S \rightarrow [0, 1]$. The key property is that the transition probability from $s$ to $s'$ only depends on $s$ and $s'$ and not on the past history of how it got there. This is what allows the probabilistic data to be given as a single matrix $T$. 
Discrete probabilistic transition systems

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

$$(S, L, \forall a \in L T_a : S \times S \rightarrow [0, 1])$$

- The model is *reactive*: All probabilistic data is *internal* - no probabilities associated with environment behaviour.
Examples of PTSs

\[
\begin{align*}
A_1 & \quad A_2 \\
S_0 & \\
| & \\
\downarrow a[\frac{1}{4}] & \downarrow a[\frac{3}{4}] \\
S_1 & \quad S_2 \\
| & \quad | \\
\downarrow a[1] & \downarrow c[\frac{1}{2}] \\
S_3 & \quad S_4 \\
| & \quad | \\
\downarrow c[\frac{1}{2}] & \downarrow a[1] \\
S_4 & \quad S_3
\end{align*}
\]
Consider

Should $s_0$ and $t_0$ be bisimilar?

Yes, but we need to add the probabilities.
Let $S = (S, L, T_a)$ 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 $a \in A$ 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.
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*. 
Basic fact: There are subsets of $\mathbb{R}$ for which no sensible notion of size can be defined.

More precisely, there is no non-trivial translation-invariant measure defined on all the subsets of the reals.
A stochastic kernel (Markov kernel) is a function \( h : S \times \Sigma \rightarrow [0, 1] \) with (a) \( h(s, \cdot) : \Sigma \rightarrow [0, 1] \) a (sub)probability measure and (b) \( h(\cdot, A) : X \rightarrow [0, 1] \) a measurable function.

Though apparently asymmetric, these are the stochastic analogues of binary relations

and the uncountable generalization of a matrix.

They are the Kleisli arrows of a monad: the Giry monad.
An LMP is a tuple $(S, \Sigma, L, \forall \alpha \in L. \tau_\alpha)$ where $\tau_\alpha : S \times \Sigma \rightarrow [0, 1]$ 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.
Let $S = (S, i, \Sigma, \tau)$ be a labelled Markov process. An equivalence relation $R$ on $S$ is a **bisimulation** if whenever $sR s'$, with $s, s' \in S$, we have that for all $a \in A$ 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.
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 $(x, y)$. 

Panangaden (McGill University)  
Lecture 1  
IRIF 16th January 2018  
16 / 33
Logical Characterization

\[ \mathcal{L} ::= T|\phi_1 \land \phi_2|\langle a \rangle_q \phi \]

We say \( s \models \langle a \rangle_q \phi \) iff

\[ \exists A \in \Sigma. (\forall s' \in A. s' \models \phi) \land (\tau_a(s, A) > q). \]

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!

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.
Proof idea

- Show that the relation “$s$ and $s'$ satisfy exactly the same formulas” is a bisimulation.
- Can easily show that $\tau_a(s, A) = \tau_a(s', A)$ for $A$ of the form $[\phi]$.
- Use Dynkin’s lemma to show that we get a well defined measure on the $\sigma$-algebra generated by such sets and the above equality holds.
- Use special properties of analytic spaces to show that this $\sigma$-algebra is the same as the original $\sigma$-algebra.
Let $S = (S, \Sigma, \tau)$ be a labelled Markov process. A preorder $R$ on $S$ is a **simulation** if whenever $sRs'$, we have that for all $a \in A$ and every $R$-closed measurable set $A \in \Sigma$, $\tau_a(s, A) \leq \tau_a(s', A)$. We say $s$ is simulated by $s'$ if $sRs'$ for some simulation relation $R$. 
The logic used in the characterization has no negation, not even a limited negative construct.

One can show that if $s$ simulates $s'$ then $s$ satisfies all the formulas of $\mathcal{L}$ that $s'$ satisfies.

What about the converse?
Counter example!

In the following picture, $t$ satisfies all formulas of $\mathcal{L}$ that $s$ satisfies but $t$ does not simulate $s$.

All transitions from $s$ and $t$ are labelled by $a$. 
Counter example (contd.)

- A formula of $\mathcal{L}$ that is satisfied by $t$ but not by $s$.
  \[ \langle a \rangle_0 \left( \langle a \rangle_0 T \land \langle b \rangle_0 T \right). \]

- A formula with disjunction that is satisfied by $s$ but not by $t$:
  \[ \langle a \rangle_{\frac{3}{4}} \left( \langle a \rangle_0 T \lor \langle b \rangle_0 T \right). \]
The logic $\mathcal{L}$ does not characterize simulation. One needs disjunction.

$$\mathcal{L}_\lor := \mathcal{L} \mid \phi_1 \lor \phi_2.$$ 

With this logic we have:

An LMP $s_1$ simulates $s_2$ if and only if for every formula $\phi$ of $\mathcal{L}_\lor$ we have

$$s_1 \models \phi \Rightarrow s_2 \models \phi.$$ 

The original proof uses domain theory and approximation.

New development (2017 ICALP) we can prove logical characterization for simulation and bisimulation in almost the same way.
An analytic set $A$ is the image of a Polish space $X$ (or a Borel subset of $X$) under a continuous (or measurable) function $f : X \to Y$, where $Y$ is Polish. If $(S, \Sigma)$ is a measurable space where $S$ is an analytic set in some ambient topological space and $\Sigma$ is the Borel $\sigma$-algebra on $S$.

Analytic sets do not form a $\sigma$-algebra but they are in the completion of the Borel algebra under any measure. [Universally measurable.]
Given a an analytic space and ∼ an equivalence relation such that there is a countable family of real-valued measurable functions $f_i : S \to \mathbb{R}$ such that

$$\forall s, s' \in S. s \sim s' \iff \forall i. f_i(s) = f_i(s')$$

then the quotient space $(Q, \Omega)$ - where $Q = S/\sim$ and $\Omega$ is the finest $\sigma$-algebra making the canonical surjection $q : S \to Q$ measurable - is also analytic.

If an analytic space $(S, \Sigma)$ has a sub-$\sigma$-algebra $\Sigma_0$ of $\Sigma$ which separates points and is countably generated then $\Sigma_0$ is $\Sigma$! The Unique Structure Theorem (UST).
A \( \pi \)-system is a family of sets closed under finite intersections.

A \( \lambda \)-system is a family of sets closed under complements and countable disjoint unions.

\( \lambda - \pi \) theorem: If \( \Pi \) is a \( \pi \)-system and \( \Lambda \) is a \( \lambda \)-system and \( \Pi \subset \Lambda \) then \( \sigma(\Pi) \subset \Lambda \).

Corollary: If two measures agree on the sets of a \( \pi \)-system then they agree on the generated \( \sigma \)-algebra.
Given \((S, \Sigma, \tau_a)\) an LMP, we define \(x \simeq y\) if \(x\) and \(y\) obey exactly the same formulas of \(L_0\).

We claim that \(\simeq\) is a bisimulation relation.

Suppose that \(x, y \in S\) and for some \(a\) and some \(\simeq\)-closed set \(C\), \(\tau_a(x, C) \neq \tau_a(y, C)\).

We need to show there is a formula on which \(x, y\) disagree.

Let \(\delta = \tau_a(x, \cdot)\) and \(\gamma = \tau_a(y, \cdot)\).

If \(\delta(S) > \gamma(S)\) then choose rational \(q\) such that \(\delta(S) > q > \gamma(S)\).

Now \(x \models \langle a \rangle_q \top\) and \(y \not\models \langle a \rangle_q \top\).
If $\delta(S) = \gamma(S)$ then pick an $\simeq$-closed set $C \in \Sigma$ with $\delta(C) \neq \gamma(C)$.

Define $\Pi = \{[\phi] | \phi \in L_0\}$ and $\Lambda = \{Y \in \Sigma | \delta(Y) = \gamma(Y)\}$. These are a $\pi$-system and a $\lambda$-system respectively.

By unique structure theorem $C \in \sigma(\Pi)$ but, by assumption $C \not\in \Lambda$ so $\Pi \not\subset \Lambda$ so there is a formula $\phi$ such that $\delta([\phi]) \neq \gamma([\phi])$.

Suppose $\delta([\phi]) > \gamma([\phi])$ choose $q$ rational in between and we have

- $x \models \langle a \rangle q \phi$ and $y \not\models \langle a \rangle q \phi$. 
How can we do this for simulation?

- Simulation is a preorder $\preceq$ rather than an equivalence relation.
- Simulation game can be defined similarly: Duplicator starts by claiming $x \preceq y$.
- Spoiler chooses $C$ which he claims is $\preceq$-closed and that $\tau(x, C) > \tau(y, C)$.
- Duplicator chooses $x' \in C$ and $y' \notin C$ and claims that $x' \preceq y'$.
- $x \preceq y$ iff Duplicator has a winning strategy starting from $x, y$. 
Positive theorems

- We had to come up with positive versions of the unique structure theorem and the monotone class theorem. With help from experts in descriptive theory.
- With these in place the proof of the logical characterization of simulation follows the same pattern.
The logical characterization theorem is *false* if you allow uncountably many labels. [Fijalkow]

However, if you require the transition functions to be continuous instead of measurable then logical characterization is restored.

For simulation as well as bisimulation.

We heavily use topological ideas in this proof.