Metrics for probabilistic processes[a]

Radha Jagadeesan
DePaul University

Prakash Panangaden
McGill University

[a]Joint work with Desharnais and Ferns and Gupta and Precup

The Main Idea

- Equational reasoning for probabilistic processes is well studied

Markov chains	Lumpability
Labelled Markov processes	Bisimulation
Markov decision processes	Bisimulation
Labelled Concurrent Markov Chains with τ transitions	Weak Bisimulation

- In the contxt of probability is exact equivalence reasonable? We say "no". Instead one should have a (pseudo)metric for probabilistic processes.

Lumpability/Bisimulation

- Fix a Markov chain. Let s, t be states.
- R is a bisimulation if for any two states within a partition induced by R, their aggregated transition rates to any other partition are the same.
- 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

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

Bisimulation: Labelled version

- Let R be an equivalence relation. R is a bisimulation if: $s\ R\ t$ if $(\forall\ a)$:
$$(s \xrightarrow{a} P) \Rightarrow [t \xrightarrow{a} Q, P =_R Q]$$
$$(t \xrightarrow{a} Q) \Rightarrow [s \xrightarrow{a} P, P =_R Q]$$
- s, t are bisimular if there is a bisimulation relating them.
- There is a maximum bisimulation relation.

Back
Close

Bisimulation for MDPs

- Markov Decision processes with rewards:
  R, an equivalence relation is a bisimulation if: $s \; R \; t \Rightarrow$ if $(\forall \; a)$:
  - $r_s^a = r_t^a$, and
  - $$(s \xrightarrow{a} P) \Rightarrow [t \xrightarrow{a} Q, \quad P =_R Q]$$
    $$(t \xrightarrow{a} Q) \Rightarrow [s \xrightarrow{a} P, P =_R Q]$$

Properties of Bisimulation

- Establishing equality of states: Coinduction.
  Establish a bisimulation R that relates states s, t.

Properties of Bisimulation

- Establishing equality of states: Coinduction.
- Distinguishing states: Simple logic is complete for bisimulation.

$$\phi ::= \texttt{true} \mid \phi_1 \wedge \phi_2 \mid \langle a \rangle_{>q} \phi$$

- Bisimulation is sound for much richer logic pCTL*.

Properties of Bisimulation

- Establishing equality of states: Coinduction.
- Distinguishing states: Logical view.
- Equational and logical views coincide.
- Compositional reasoning: Bisimulation is a congruence for usual process operators.

Are exact equivalences reasonable??

- Exact reasoning is unstable under (small) perturbations of probability numbers.

Are exact equivalences reasonable??

- Exact reasoning is unstable under (small) perturbations of probability numbers.
- Probability numbers are to be viewed as coming with some error estimate: reasoning principles based on the exact value of numbers are of dubious value.
  - Physical systems: Probability arises in the modelling of kinetics in biochemical reactions or as stochastic noise used as an abstraction to specify incomplete knowledge.
  - Diagnosis: Models of failure are based on empirical and statistical evidence.

Are exact equivalences reasonable??

- Exact reasoning is unstable under (small) perturbations of probability numbers.
- Probability numbers are to be viewed as coming with some error estimate.
- Approximation of probability distributions:
  - Monte-Carlo methods to approximate probability distributions by a sample.
  - Approximating continuous distributions by discrete distributions

A metric-based approximate viewpoint

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

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

- Quantitative measurement of the distinction between processes.

Summary of results

- Results work for Markov chains, Labelled Markov processes, Markov decision processes and Labelled Concurrent Markov chains with τ-transitions.
- Establishing closeness of states: Coinduction.
- Distinguishing states: Real-valued modal logics.
- Equational and logical views coincide: Metrics yield same distances as real-valued modal logics.
- Compositional reasoning by *Non-Expansivity*. Process-combinators take closeby processes to closeby processes. eg.

$$\frac{d(s_1, t_1) < \epsilon_1, \qquad d(s_2, t_2) < \epsilon_2}{d(s_1 \mid\mid s_2, t_1 \mid\mid t_2) < \epsilon_1 + \epsilon_2}$$

Back
Close

Criterion 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*."

Lumpability/Bisimulation

- 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-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.

- $\mathcal{M}$: 1-bounded pseudometrics on states.
- $$d(\mu,\nu) = \sup_f \left| \int f d\mu - \int f d\nu \right|, f \text{ 1-Lipschitz}$$
- 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:
  $$\begin{array}{l} \forall i. 0 \leq a_i \leq 1 \\ \forall i,j.\ a_i - a_j \leq m(s_i, s_j). \end{array}$$

Duality yields splitting

- (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:

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

Some examples

- $m(P, P) = 0$.
- In dual, match each state with itself. So:

$$\min \sum_{i,j} l_{ij} m(s_i, s_j) + \sum_i x_i + \sum_j y_j$$

becomes 0.

Back
Close

Some examples

- Let $m(s,t) = r < 1$. Let $\delta_s(\delta_t)$ be the probability measure concentrated at s Then,
$$m(\delta_s, \delta_t) = r$$
- Upper bound from dual: Choose $l_{st} = 1$.
Lower bound from primal: Choose $f(s) = 0, f(t) = r$.

Some examples

- Let $P(s) = r$, $P(t) = 0$ if $s \neq t$.
  Let $Q(s) = r'$, $P(t) = 0$ if $s \neq t$.

$$d(P, Q) = |r - r'|$$

- Upper bound from dual: Match each state with itself.

$$\min \sum_{i,j} l_{ij} m(s_i, s_j) + \sum_i x_i + \sum_j y_j$$

  becomes $|r - r'|$.
  Lower bound from primal: yielded by constant function 1.

Return from Detour

- Summary of detour: Given a metric on states in a metric space, can lift to a metric on probability distributions on states.

Metric Bisimulation

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

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

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

- The required canonical metric on processes is the least such: ie. the distance numbers are the least possible.
- Thm: *Canonical least metric exists*

A metric for Markov decision processes

- Let $\alpha + \beta \leq 1$.
- m, a pseudo-metric is a (α, β)-metric bisimulation if: $m(s,t) < \epsilon \Rightarrow (\forall\ a)(\exists \epsilon_1, \epsilon_2).\epsilon_1 + \epsilon_2 \leq \epsilon$:
  - $$|r_s^a - r_t^a| < \frac{\epsilon_1}{\alpha}, \quad \text{and}$$
  - $$(s \xrightarrow{a} P) \Rightarrow [t \xrightarrow{a} Q, m(P,Q) < \frac{\epsilon_2}{\beta}]$$
    $$(t \xrightarrow{a} Q) \Rightarrow [s \xrightarrow{a} P, m(P,Q) < \frac{\epsilon_2}{\beta}]$$

Back
Close

A metric for Markov decision processes: Results

- (Versions of) earlier results go through: real-valued modal logic, completeness etc..
- (Similar story for labelled concurrent markov chains with τ-transitions)

Calculating quantitative observables

- Nearby states have nearby quantitative observations.
  - Expectation of uniformly continuous functions is a continuous function of the metric.
  - An example from information theory: channel capacity is a continuous function of the metric.
  - MDP's: Nearby states have nearby (optimal) value functions.

Calculating the metric

- LPs are of special form: a transportation problem known to be strongly polynomial.
- For discounted (future) form of metrics, this yields algorithms to calculate metric distances to within a known error.

Conclusion

- Towards continuous state spaces.
- Experiments in progress to evaluate aggregation using metrics.
- Recipe to turn exact reasoning into approximate reasoning.

Extra slides

Revisit bisimulation: maximum fixed point definition

- Let R be an equivalence relation. $F(R)(s,t)$ if:

$$s \longrightarrow P \Rightarrow \; t \longrightarrow Q, \quad P =_R Q$$

$$t \longrightarrow Q \Rightarrow \; s \longrightarrow P, \quad P =_R Q$$

- Bisimulation is maximum fixed point of F.

Metrics: technical development

- $\mathcal{M}$: 1-bounded pseudometrics on states with ordering

$$m_1 \preceq m_2 \text{ if } (\forall s,t)\ [m_1(s,t) \geq m_2(s,t)]$$

- $(\mathcal{M}, \preceq)$ is a complete lattice.

$$\begin{aligned}
\bot(s,t) &= \begin{cases} 0 \text{ if } s = t \\ 1 \text{ otherwise} \end{cases} \\
\top(s,t) &= 0, (\forall s,t) \\
(\sqcap\{m_i\})(s,t) &= \sup_i m_i(s,t)
\end{aligned}$$

A maximum fixed point definition

- Let $m \in \mathcal{M}$. $F(m)(s,t) < \epsilon$ if:

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

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

- F is monotone on $\mathcal{M}$, and metric-bisimulation is the greatest fixed point of F.

Real-valued Modal Logic

- $$f ::= \mathbf{1} \mid \max(f, f) \mid h \circ f \mid a.f$$

- $$\begin{array}{lcll} \mathbf{1}(s) & = & 1 & \text{True} \\ \max(f_1, f_2)(s) & = & \max(f_1(s), f_2(s)) & \text{Conjunction} \\ h \circ f(s) & = & h(f(s)) & \text{Conditionals} \\ & & & h \text{ 1-Lipschitz} : [0,1] \longrightarrow [0,1] \end{array}$$

- $d(s,t) = \sup_f |f(s) - f(t)|$
- Thm: d coincides with the canonical metric-bisimulation.

Back
Close

"Finitary" syntax for Real-valued modal logic

•

$$
\begin{array}{lcll}
\mathbf{1}(s) & = & 1 & \text{True} \\
\max(f_1, f_2)(s) & = & \max(f_1(s), f_2(s)) & \text{Conjunction} \\
(1-f)(s) & = & 1 - f(s) & \text{Negation} \\
\lfloor f_q(s) \rfloor & = & \begin{cases} q \quad , & f(s) \geq q \\ f(s) \quad , & f(s) < q \end{cases} & \text{Lipschitz conditionals}
\end{array}
$$

q is a rational.