Probabilistic bisimulation and related metrics

Prakash Panangaden
School of Computer Science
McGill University
and
Montreal Institute of Learning Algorithms
and
Strategic Visitor University of Edinburgh

23 March 2022

Who am I?

- Professor at McGill University and Core Member of Montreal Institute of Learning Algorithms (MILA)

Who am I?

- Professor at McGill University and Core Member of Montreal Institute of Learning Algorithms (MILA)
- Research interests (1) reasoning about probabilistic systems: equivalences, logics, metrics, approximation, applications to verification and to machine learning.

Who am I?

- Professor at McGill University and Core Member of Montreal Institute of Learning Algorithms (MILA)
- Research interests (1) reasoning about probabilistic systems: equivalences, logics, metrics, approximation, applications to verification and to machine learning.
- (2) Logic: epistemic logic and distributed systems, logics of belief, quantitative equational logic, categorical logic, modal logics for concurrency

Who am I?

- Professor at McGill University and Core Member of Montreal Institute of Learning Algorithms (MILA)
- Research interests (1) reasoning about probabilistic systems: equivalences, logics, metrics, approximation, applications to verification and to machine learning.
- (2) Logic: epistemic logic and distributed systems, logics of belief, quantitative equational logic, categorical logic, modal logics for concurrency
- (3) Concurrency: dataflow languages, concurrent constraint programming languages, expressiveness, type systems for dataflow languages, monoidal categories for concurrent systems

Who am I?

- Professor at McGill University and Core Member of Montreal Institute of Learning Algorithms (MILA)
- Research interests (1) reasoning about probabilistic systems: equivalences, logics, metrics, approximation, applications to verification and to machine learning.
- (2) Logic: epistemic logic and distributed systems, logics of belief, quantitative equational logic, categorical logic, modal logics for concurrency
- (3) Concurrency: dataflow languages, concurrent constraint programming languages, expressiveness, type systems for dataflow languages, monoidal categories for concurrent systems
- (4) Quantum information theory, quantum computation, quantum mechanics and formerly quantum field theory in curved spacetime.

Who am I?

- Professor at McGill University and Core Member of Montreal Institute of Learning Algorithms (MILA)
- Research interests (1) reasoning about probabilistic systems: equivalences, logics, metrics, approximation, applications to verification and to machine learning.
- (2) Logic: epistemic logic and distributed systems, logics of belief, quantitative equational logic, categorical logic, modal logics for concurrency
- (3) Concurrency: dataflow languages, concurrent constraint programming languages, expressiveness, type systems for dataflow languages, monoidal categories for concurrent systems
- (4) Quantum information theory, quantum computation, quantum mechanics and formerly quantum field theory in curved spacetime.
- (5) Type theory, programming language semantics.

Who am I?

- Professor at McGill University and Core Member of Montreal Institute of Learning Algorithms (MILA)
- Research interests (1) reasoning about probabilistic systems: equivalences, logics, metrics, approximation, applications to verification and to machine learning.
- (2) Logic: epistemic logic and distributed systems, logics of belief, quantitative equational logic, categorical logic, modal logics for concurrency
- (3) Concurrency: dataflow languages, concurrent constraint programming languages, expressiveness, type systems for dataflow languages, monoidal categories for concurrent systems
- (4) Quantum information theory, quantum computation, quantum mechanics and formerly quantum field theory in curved spacetime.
- (5) Type theory, programming language semantics.
- (6) Occasional forays into physics (GR) and pure mathematics.

Main achievements outside research

- bowled my elder brother out for a duck with a vicious leg break,

Main achievements outside research

- bowled my elder brother out for a duck with a vicious leg break,
- was MWTC Men's B Division Consolation Round Runner-up.

Today's topic

Probabilistic bisimulation: originally invented with a view to verification but we have found it useful in reinforcement learning.

Behavioural equivalence is fundamental

- When do two states have exactly the same behaviour?

Behavioural equivalence is fundamental

- When do two states have exactly the same behaviour?
- What can one observe of the behaviour?

Behavioural equivalence is fundamental

- When do two states have exactly the same behaviour?
- What can one observe of the behaviour?
- What should be guaranteed?

Behavioural equivalence is fundamental

- When do two states have exactly the same behaviour?
- What can one observe of the behaviour?
- What should be guaranteed?
- (i) If two states are equivalent we should not be able to “see” any differences in observable behaviour.

Behavioural equivalence is fundamental

- When do two states have exactly the same behaviour?
- What can one observe of the behaviour?
- What should be guaranteed?
- (i) If two states are equivalent we should not be able to “see” any differences in observable behaviour.
- (ii) If two states are equivalent they should stay equivalent as they evolve.

A bit of history

- Cantor and the back-and-forth argument

A bit of history

- Cantor and the back-and-forth argument
- Lumpability in queueing theory 1960's

A bit of history

- Cantor and the back-and-forth argument
- Lumpability in queueing theory 1960's
- Bisimulation of nondeterministic automata 1970's and process algebras 1980's: Milner and Park

A bit of history

- Cantor and the back-and-forth argument
- Lumpability in queueing theory 1960's
- Bisimulation of nondeterministic automata 1970's and process algebras 1980's: Milner and Park
- Probabilistic bisimulation, discrete systems: Larsen and Skou 1989

A bit of history

- Cantor and the back-and-forth argument
- Lumpability in queueing theory 1960's
- Bisimulation of nondeterministic automata 1970's and process algebras 1980's: Milner and Park
- Probabilistic bisimulation, discrete systems: Larsen and Skou 1989
- Bisimulation of Markov processes on continuous state spaces: Desharnais, Edalat, P. 1997...

A bit of history

- Cantor and the back-and-forth argument
- Lumpability in queueing theory 1960's
- Bisimulation of nondeterministic automata 1970's and process algebras 1980's: Milner and Park
- Probabilistic bisimulation, discrete systems: Larsen and Skou 1989
- Bisimulation of Markov processes on continuous state spaces: Desharnais, Edalat, P. 1997...
- Bisimulation metrics for Markov processes Desharnais, Gupta, Jagadeesan, P. 1999

A bit of history

- Cantor and the back-and-forth argument
- Lumpability in queueing theory 1960's
- Bisimulation of nondeterministic automata 1970's and process algebras 1980's: Milner and Park
- Probabilistic bisimulation, discrete systems: Larsen and Skou 1989
- Bisimulation of Markov processes on continuous state spaces: Desharnais, Edalat, P. 1997...
- Bisimulation metrics for Markov processes Desharnais, Gupta, Jagadeesan, P. 1999
- Fixed-point version: van Breugel and Worrell 2001

A bit of history

- Cantor and the back-and-forth argument
- Lumpability in queueing theory 1960's
- Bisimulation of nondeterministic automata 1970's and process algebras 1980's: Milner and Park
- Probabilistic bisimulation, discrete systems: Larsen and Skou 1989
- Bisimulation of Markov processes on continuous state spaces: Desharnais, Edalat, P. 1997...
- Bisimulation metrics for Markov processes Desharnais, Gupta, Jagadeesan, P. 1999
- Fixed-point version: van Breugel and Worrell 2001
- Bisimulation for MDP's : Givan and Dean 2003

A bit of history

- Cantor and the back-and-forth argument
- Lumpability in queueing theory 1960's
- Bisimulation of nondeterministic automata 1970's and process algebras 1980's: Milner and Park
- Probabilistic bisimulation, discrete systems: Larsen and Skou 1989
- Bisimulation of Markov processes on continuous state spaces: Desharnais, Edalat, P. 1997...
- Bisimulation metrics for Markov processes Desharnais, Gupta, Jagadeesan, P. 1999
- Fixed-point version: van Breugel and Worrell 2001
- Bisimulation for MDP's : Givan and Dean 2003
- Bisimulation metrics for MDP's: Ferns, Precup, P. 2004, 2005 (UAI)

A bit of history

- Cantor and the back-and-forth argument
- Lumpability in queueing theory 1960's
- Bisimulation of nondeterministic automata 1970's and process algebras 1980's: Milner and Park
- Probabilistic bisimulation, discrete systems: Larsen and Skou 1989
- Bisimulation of Markov processes on continuous state spaces: Desharnais, Edalat, P. 1997...
- Bisimulation metrics for Markov processes Desharnais, Gupta, Jagadeesan, P. 1999
- Fixed-point version: van Breugel and Worrell 2001
- Bisimulation for MDP's : Givan and Dean 2003
- Bisimulation metrics for MDP's: Ferns, Precup, P. 2004, 2005 (UAI)
- Representation learning using "metrics": Castro, Kastner, P., Rowland 2021 (NeurIPS)

The definition of an LTS

- A set of states S,

The definition of an LTS

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

The definition of an LTS

- 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).

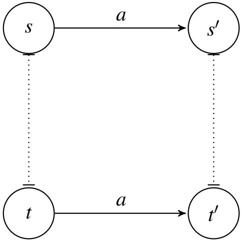
The definition of an LTS

- 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 $(s, s') \in \rightarrow_a$.

Formal definition



[Bisimulation definition]

If $s \sim t$ then

$$\forall s \in S, \forall a \in \mathcal{A}, s \xrightarrow{a} s' \Rightarrow \exists t', t \xrightarrow{a} t' \text{ with } s' \sim t'$$

and *vice versa* with s and t interchanged.

Discrete probabilistic transition systems

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

Discrete probabilistic transition systems

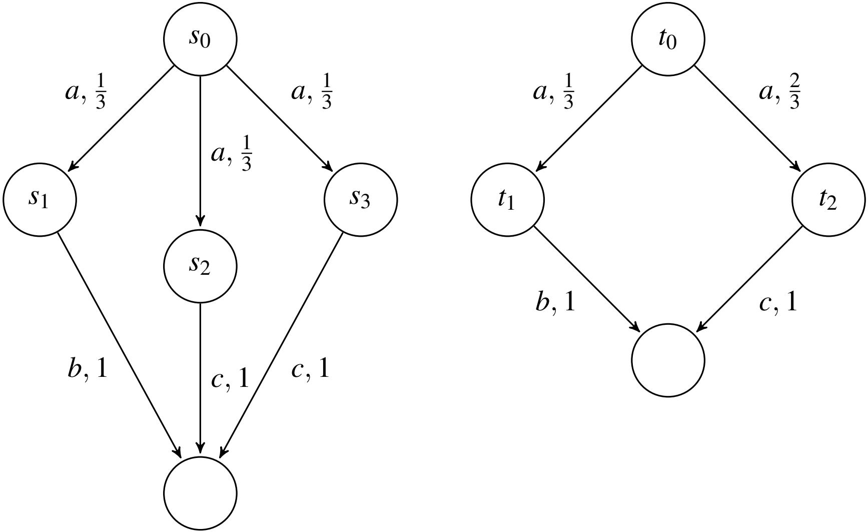
- Just like a labelled transition system with probabilities associated with the transitions.
-

$$(S, \mathcal{A}, \forall a \in \mathcal{A}\ T_a : S \times S \longrightarrow [0, 1])$$

Discrete probabilistic transition systems

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

Probabilistic bisimulation : Larsen and Skou



Are s_0 and t_0 bisimilar?

Yes, but one needs to add up the probabilities to s_2 and s_3.

Are s_0 and t_0 bisimilar?

Yes, but one needs to add up the probabilities to s_2 and s_3.

If s is a state, a an action and C a set of states, we write $T_a(s, C) = \sum_{s' \in S} T_a(s, s')$ for the probability of jumping on an a-action to one of the states in C.

Are s_0 and t_0 bisimilar?

Yes, but one needs to add up the probabilities to s_2 and s_3.

If s is a state, a an action and C a set of states, we write $T_a(s, C) = \sum_{s' \in S} T_a(s, s')$ for the probability of jumping on an a-action to one of the states in C.

Definition

R is a bisimulation relation if whenever sRt and C is an equivalence class of R then $T_a(s, C) = T_a(t, C)$.

Markov decision processes?

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

Markov decision processes?

- Markov decision processes are probabilistic versions of labelled transition systems. Labelled transition systems where the final state is governed by a probability distribution - no other indeterminacy.
- There is a *reward* associated with each transition.

Markov decision processes?

- Markov decision processes are probabilistic versions of labelled transition systems. Labelled transition systems where the final state is governed by a probability distribution - no other indeterminacy.
- There is a *reward* associated with each transition.
- We observe the interactions and the rewards - not the internal states.

Markov decision processes: formal definition

$$(S, \mathcal{A}, \forall a \in \mathcal{A}, P^a : S \longrightarrow \mathcal{D}(S), \mathcal{R} : \mathcal{A} \times S \longrightarrow \mathbf{R})$$

where

S : the state space, we will take it to be a finite set.

$\mathcal{A}$: the actions, a finite set

P^a : the transition function; $\mathcal{D}(S)$ denotes distributions over S

$\mathcal{R}$: the reward, could readily make it stochastic.

Will write $P^a(s, C)$ for $P^a(s)(C)$.

Policies

MDP

$$(S, \mathcal{A}, \forall a \in \mathcal{A}, P^a : S \longrightarrow \mathcal{D}(S), \mathcal{R} : \mathcal{A} \times S \longrightarrow \mathbf{R})$$

We control the choice of action; it is not some external scheduler.

Policies

MDP

$$(S, \mathcal{A}, \forall a \in \mathcal{A}, P^a : S \longrightarrow \mathcal{D}(S), \mathcal{R} : \mathcal{A} \times S \longrightarrow \mathbf{R})$$

We control the choice of action; it is not some external scheduler.

Policy

$$\pi : S \longrightarrow \mathcal{D}(\mathcal{A})$$

Policies

MDP

$$(S, \mathcal{A}, \forall a \in \mathcal{A}, P^a : S \longrightarrow \mathcal{D}(S), \mathcal{R} : \mathcal{A} \times S \longrightarrow \mathbf{R})$$

We control the choice of action; it is not some external scheduler.

Policy

$$\pi : S \longrightarrow \mathcal{D}(\mathcal{A})$$

Policies

MDP

$$(S, \mathcal{A}, \forall a \in \mathcal{A}, P^a : S \longrightarrow \mathcal{D}(S), \mathcal{R} : \mathcal{A} \times S \longrightarrow \mathbf{R})$$

We control the choice of action; it is not some external scheduler.

Policy

$$\pi : S \longrightarrow \mathcal{D}(\mathcal{A})$$

The goal is **choose** the best policy: numerous algorithms to find or approximate the optimal policy.

Bisimulation

- Let R be an equivalence relation. R is a bisimulation if: $s\ R\ t$ if $(\forall\ a)$ and all equivalence classes C of R:

Bisimulation

- Let R be an equivalence relation. R is a bisimulation if: $s\ R\ t$ if $(\forall\ a)$ and all equivalence classes C of R:
  (i) $\mathcal{R}(a, s) = \mathcal{R}(a, t)$

Bisimulation

- Let R be an equivalence relation. R is a bisimulation if: $s\ R\ t$ if $(\forall\ a)$ and all equivalence classes C of R:
  - (i) $\mathcal{R}(a, s) = \mathcal{R}(a, t)$
  - (ii) $P^a(s, C) = P^a(t, C)$

Bisimulation

- Let R be an equivalence relation. R is a bisimulation if: $s\ R\ t$ if $(\forall\ a)$ and all equivalence classes C of R:
  (i) $\mathcal{R}(a, s) = \mathcal{R}(a, t)$
  (ii) $P^a(s, C) = P^a(t, C)$
- s, t are bisimilar if there is a bisimulation relation R with sRt them.

Bisimulation

- Let R be an equivalence relation. R is a bisimulation if: $s\ R\ t$ if $(\forall\ a)$ and all equivalence classes C of R:
  (i) $\mathcal{R}(a,s) = \mathcal{R}(a,t)$
  (ii) $P^a(s,C) = P^a(t,C)$
- s,t are bisimilar if there is a bisimulation relation R with sRt them.
- Basic pattern: immediate rewards match (initiation), stay related after the transition (coinduction).

Bisimulation

- Let R be an equivalence relation. R is a bisimulation if: $s\ R\ t$ if $(\forall\ a)$ and all equivalence classes C of R:
  (i) $\mathcal{R}(a, s) = \mathcal{R}(a, t)$
  (ii) $P^a(s, C) = P^a(t, C)$
- s, t are bisimilar if there is a bisimulation relation R with sRt them.
- Basic pattern: immediate rewards match (initiation), stay related after the transition (coinduction).
- Bisimulation can be defined as the *greatest fixed point* of a relation transformer.

Continuous state spaces: why?

- Software controllers attached to physical devices or sensors - robots, controllers.

Continuous state spaces: why?

- Software controllers attached to physical devices or sensors - robots, controllers.
- Continuous state space but discrete time.

Continuous state spaces: why?

- Software controllers attached to physical devices or sensors - robots, controllers.
- Continuous state space but discrete time.
- Applications to control systems.

Continuous state spaces: why?

- Software controllers attached to physical devices or sensors - robots, controllers.
- Continuous state space but discrete time.
- Applications to control systems.
- Applications to probabilistic programming languages.

Some remarks on the use of continuous spaces

- Can be used for reasoning - but much better if we could have a finite-state version.

Some remarks on the use of continuous spaces

- Can be used for reasoning - but much better if we could have a finite-state version.
- Why not discretize right away and never worry about the continuous case?

Some remarks on the use of continuous spaces

- Can be used for reasoning - but much better if we could have a finite-state version.
- Why not discretize right away and never worry about the continuous case?
- How can we say that our discrete approximation is “accurate”?

Some remarks on the use of continuous spaces

- Can be used for reasoning - but much better if we could have a finite-state version.
- Why not discretize right away and never worry about the continuous case?
- How can we say that our discrete approximation is “accurate”?
- We lose the ability to *refine* the model later.

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.

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 apparantly 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 apparantly asymmetric, these are the stochastic analogues of binary relations
- and the uncountable generalization of a matrix.

Logical Characterization

- Very austere logic:

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

Logical Characterization

- Very austere logic:

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

- $s \models \langle a \rangle_q \phi$ means that if the system is in state s, then after the action a, with probability at least q the new state will satisfy the formula ϕ.

Logical Characterization

- Very austere logic:

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

- $s \models \langle a \rangle_q \phi$ means that if the system is in state s, then after the action a, with probability at least q the new state will satisfy the formula ϕ.
- Two systems are bisimilar iff they obey the same formulas of $\mathcal{L}$. [DEP 1998 LICS, I and C 2002]

Logical Characterization

- Very austere logic:

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

- $s \models \langle a \rangle_q \phi$ means that if the system is in state s, then after the action a, with probability at least q the new state will satisfy the formula ϕ.
- Two systems are bisimilar iff they obey the same formulas of $\mathcal{L}$. [DEP 1998 LICS, I and C 2002]
- No finite branching assumption.

Logical Characterization

- Very austere logic:

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

- $s \models \langle a \rangle_q \phi$ means that if the system is in state s, then after the action a, with probability at least q the new state will satisfy the formula ϕ.
- Two systems are bisimilar iff they obey the same formulas of $\mathcal{L}$. [DEP 1998 LICS, I and C 2002]
- No finite branching assumption.
- No negation in the logic,

Logical Characterization

- Very austere logic:

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

- $s \models \langle a \rangle_q \phi$ means that if the system is in state s, then after the action a, with probability at least q the new state will satisfy the formula ϕ.
- Two systems are bisimilar iff they obey the same formulas of $\mathcal{L}$. [DEP 1998 LICS, I and C 2002]
- No finite branching assumption.
- No negation in the logic,
- so one can obtain a logical characterization result for simulation

Logical Characterization

- Very austere logic:

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

- $s \models \langle a \rangle_q \phi$ means that if the system is in state s, then after the action a, with probability at least q the new state will satisfy the formula ϕ.
- Two systems are bisimilar iff they obey the same formulas of $\mathcal{L}$. [DEP 1998 LICS, I and C 2002]
- No finite branching assumption.
- No negation in the logic,
- so one can obtain a logical characterization result for simulation
- but it needs disjunction.

Logical Characterization

- Very austere logic:

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

- $s \models \langle a \rangle_q \phi$ means that if the system is in state s, then after the action a, with probability at least q the new state will satisfy the formula ϕ.
- Two systems are bisimilar iff they obey the same formulas of $\mathcal{L}$. [DEP 1998 LICS, I and C 2002]
- No finite branching assumption.
- No negation in the logic,
- so one can obtain a logical characterization result for simulation
- but it needs disjunction.
- The proof uses tools from descriptive set theory and measure theory.

But...

- In the context of probability is exact equivalence reasonable?

But...

- In the context of probability is exact equivalence reasonable?
- We say “no”. A small change in the probability distributions may result in bisimilar processes no longer being bisimilar though they may be very “close” in behaviour.

But...

- In the context of probability is exact equivalence reasonable?
- We say “no”. A small change in the probability distributions may result in bisimilar processes no longer being bisimilar though they may be very “close” in behaviour.
- Instead one should have a (pseudo)metric for probabilistic processes.

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).
- Quantitative measurement of the distinction between processes.

In lieu of several slides of greek letters and symbols

- If two states are **not** bisimilar there is a some observation on which they disagree.

In lieu of several slides of greek letters and symbols

- If two states are **not** bisimilar there is a some observation on which they disagree.
- They may diasagree on the reward or on the probability distribution that results from a transition.

In lieu of several slides of greek letters and symbols

- If two states are **not** bisimilar there is a some observation on which they disagree.
- They may diasagree on the reward or on the probability distribution that results from a transition.
- We need to measure the latter, we use the ~~Wasserstein~~ Kantorovich metric between probability distributions.

In lieu of several slides of greek letters and symbols

- If two states are **not** bisimilar there is a some observation on which they disagree.
- They may diasagree on the reward or on the probability distribution that results from a transition.
- We need to measure the latter, we use the ~~Wasserstein~~ Kantorovich metric between probability distributions.
- Intuitively, if the difference shows up only after a long and elaborate observation then we should make the states “nearby” in the bisimulation metric.

In lieu of several slides of greek letters and symbols

- If two states are **not** bisimilar there is a some observation on which they disagree.
- They may diasagree on the reward or on the probability distribution that results from a transition.
- We need to measure the latter, we use the ~~Wasserstein~~ Kantorovich metric between probability distributions.
- Intuitively, if the difference shows up only after a long and elaborate observation then we should make the states “nearby” in the bisimulation metric.
- All this can be formalized and was originally done by Desharnais et al. and later with a beautiful fixed-point construction by van Breugel and Worrell.

In lieu of several slides of greek letters and symbols

- If two states are **not** bisimilar there is a some observation on which they disagree.
- They may diasagree on the reward or on the probability distribution that results from a transition.
- We need to measure the latter, we use the ~~Wasserstein~~ Kantorovich metric between probability distributions.
- Intuitively, if the difference shows up only after a long and elaborate observation then we should make the states "nearby" in the bisimulation metric.
- All this can be formalized and was originally done by Desharnais et al. and later with a beautiful fixed-point construction by van Breugel and Worrell.
- Ferns et al. added rewards and showed that the bisimulation metric bounds the difference in optimal value functions.

The basic setting: metric spaces

- A *pseudometric* on a set X is a function $d : X \times X \longrightarrow \mathbf{R}^{\geq 0}$ such that

The basic setting: metric spaces

- A *pseudometric* on a set X is a function $d : X \times X \longrightarrow \mathbf{R}^{\geq 0}$ such that
  1. $\forall x \in X, d(x,x) = 0$

The basic setting: metric spaces

- A *pseudometric* on a set X is a function $d : X \times X \longrightarrow \mathbf{R}^{\geq 0}$ such that
  1. $\forall x \in X, d(x, x) = 0$
  2. $\forall x, y \in X, d(x, y) = d(y, x)$

The basic setting: metric spaces

- A *pseudometric* on a set X is a function $d : X \times X \longrightarrow \mathbf{R}^{\geq 0}$ such that
  1. $\forall x \in X, d(x, x) = 0$
  2. $\forall x, y \in X, d(x, y) = d(y, x)$
  3. $\forall x, y, z \in X, d(x, y) \leq d(x, z) + d(z, y)$

The basic setting: metric spaces

- A *pseudometric* on a set X is a function $d : X \times X \longrightarrow \mathbf{R}^{\geq 0}$ such that
  1. $\forall x \in X, d(x,x) = 0$
  2. $\forall x, y \in X, d(x,y) = d(y,x)$
  3. $\forall x, y, z \in X, d(x,y) \leq d(x,z) + d(z,y)$
  4. If $d(x,y) = 0$ implies $x = y$ we say that it is a *metric*

The basic setting: metric spaces

- A *pseudometric* on a set X is a function $d : X \times X \longrightarrow \mathbf{R}^{\geq 0}$ such that
  1. $\forall x \in X, d(x,x) = 0$
  2. $\forall x,y \in X, d(x,y) = d(y,x)$
  3. $\forall x,y,z \in X, d(x,y) \leq d(x,z) + d(z,y)$
  4. If $d(x,y) = 0$ implies $x = y$ we say that it is a *metric*

The setup

A set M equipped with a **metric** d obeying the above axioms (unlike, for example, KL-divergence which is **not** a metric). A metric space is **complete** if every Cauchy sequence has a limit point to which it converges.

The setup

- We will assume that we have an underlying metric space—the state space—and we are looking at probability distributions on top of this space.

The setup

- We will assume that we have an underlying metric space—the state space—and we are looking at probability distributions on top of this space.
- We will then look at ways to define a metric on the space of probability distributions.

The setup

- We will assume that we have an underlying metric space—the state space—and we are looking at probability distributions on top of this space.
- We will then look at ways to define a metric on the space of probability distributions.
- It should be, somehow, related to the metric of the underlying space.

The ~~Wasserstein~~ Kantorovich metric

- What is the observable aspect of a probability distribution?

The ~~Wasserstein~~ Kantorovich metric

- What is the observable aspect of a probability distribution?
- Expectation values.

The ~~Wasserstein~~ Kantorovich metric

- What is the observable aspect of a probability distribution?
- Expectation values.
- $\kappa(P, Q) = \sup_{f \in ??} | \int f \mathrm{d}P - \int f \mathrm{d}Q|$

The ~~Wasserstein~~ Kantorovich metric

- What is the observable aspect of a probability distribution?
- Expectation values.
- $\kappa(P, Q) = \sup_{f \in ??} | \int f \mathrm{d}P - \int f \mathrm{d}Q|$
- But what kind of functions should we allow? Not just continuous ones.

The ~~Wasserstein~~ Kantorovitch metric

- What is the observable aspect of a probability distribution?
- Expectation values.
- $\kappa(P,Q) = \sup_{f\in ??} |\int f \mathrm{d}P - \int f \mathrm{d}Q|$
- But what kind of functions should we allow? Not just continuous ones.
- Nonexpansive or Lipschitz-1 functions: $d(f(x), f(y)) \leq d(x,y)$.

The ~~Wasserstein~~ Kantorovich metric

- What is the observable aspect of a probability distribution?
- Expectation values.
- $\kappa(P,Q) = \sup_{f \in ??} |\int f \mathrm{d}P - \int f \mathrm{d}Q|$
- But what kind of functions should we allow? Not just continuous ones.
- Nonexpansive or Lipschitz-1 functions: $d(f(x), f(y)) \leq d(x,y)$.
- Such functions are always continuous but, clearly, continuous functions are not necessarily Lipschitz-1.

The ~~Wasserstein~~ Kantorovich metric

- What is the observable aspect of a probability distribution?
- Expectation values.
- $\kappa(P, Q) = \sup_{f \in ??} |\int f \mathrm{d}P - \int f \mathrm{d}Q|$
- But what kind of functions should we allow? Not just continuous ones.
- Nonexpansive or Lipschitz-1 functions: $d(f(x), f(y)) \leq d(x, y)$.
- Such functions are always continuous but, clearly, continuous functions are not necessarily Lipschitz-1.
- $\kappa(P, Q) = \sup_{f \in \mathrm{Lip}_1} |\int f \mathrm{d}P - \int f \mathrm{d}Q|$

The ~~Wasserstein~~ Kantorovich metric

- What is the observable aspect of a probability distribution?
- Expectation values.
- $\kappa(P, Q) = \sup_{f \in ??} |\int f \mathrm{d}P - \int f \mathrm{d}Q|$
- But what kind of functions should we allow? Not just continuous ones.
- Nonexpansive or Lipschitz-1 functions: $d(f(x), f(y)) \leq d(x, y)$.
- Such functions are always continuous but, clearly, continuous functions are not necessarily Lipschitz-1.
- $\kappa(P, Q) = \sup_{f \in \mathrm{Lip}_1} |\int f \mathrm{d}P - \int f \mathrm{d}Q|$
- It is easy to verify all the metric conditions.

The ~~Wasserstein~~ Kantorovich metric

- What is the observable aspect of a probability distribution?
- Expectation values.
- $\kappa(P, Q) = \sup_{f \in ??} |\int f \mathrm{d}P - \int f \mathrm{d}Q|$
- But what kind of functions should we allow? Not just continuous ones.
- Nonexpansive or Lipschitz-1 functions: $d(f(x), f(y)) \leq d(x, y)$.
- Such functions are always continuous but, clearly, continuous functions are not necessarily Lipschitz-1.
- $\kappa(P, Q) = \sup_{f \in \mathrm{Lip}_1} |\int f \mathrm{d}P - \int f \mathrm{d}Q|$
- It is easy to verify all the metric conditions.
- But this definition is only half the story.

Couplings

- How to relate two distributions? Think of a distribution as a pile of sand.

Couplings

- How to relate two distributions? Think of a distribution as a pile of sand.
- We need to move some sand around to make the pile P look like Q.

Couplings

- How to relate two distributions? Think of a distribution as a pile of sand.
- We need to move some sand around to make the pile P look like Q.
- There are many different ways to do it. Each way is a “transport plan.”

Couplings

- How to relate two distributions? Think of a distribution as a pile of sand.
- We need to move some sand around to make the pile P look like Q.
- There are many different ways to do it. Each way is a “transport plan.”
- A **coupling** of two distributions P, Q defined on X is a *joint* distribution γ on $X \times X$ such that the *marginals* of γ are P and Q.

Couplings

- How to relate two distributions? Think of a distribution as a pile of sand.
- We need to move some sand around to make the pile P look like Q.
- There are many different ways to do it. Each way is a "transport plan."
- A **coupling** of two distributions P, Q defined on X is a *joint* distribution γ on $X \times X$ such that the *marginals* of γ are P and Q.
- There is always the independent coupling: $\gamma(A \times B) = P(A)Q(B)$.

Couplings

- How to relate two distributions? Think of a distribution as a pile of sand.
- We need to move some sand around to make the pile P look like Q.
- There are many different ways to do it. Each way is a "transport plan."
- A **coupling** of two distributions P, Q defined on X is a *joint* distribution γ on $X \times X$ such that the *marginals* of γ are P and Q.
- There is always the independent coupling: $\gamma(A \times B) = P(A)Q(B)$.
- But there are many others: the convex combinations of couplings are couplings.

Couplings

- How to relate two distributions? Think of a distribution as a pile of sand.
- We need to move some sand around to make the pile P look like Q.
- There are many different ways to do it. Each way is a "transport plan."
- A **coupling** of two distributions P, Q defined on X is a *joint* distribution γ on $X \times X$ such that the *marginals* of γ are P and Q.
- There is always the independent coupling: $\gamma(A \times B) = P(A)Q(B)$.
- But there are many others: the convex combinations of couplings are couplings.
- We write $\mathcal{C}(P, Q)$ for the set of couplings of P and Q.

Couplings

- How to relate two distributions? Think of a distribution as a pile of sand.
- We need to move some sand around to make the pile P look like Q.
- There are many different ways to do it. Each way is a "transport plan."
- A **coupling** of two distributions P, Q defined on X is a *joint* distribution γ on $X \times X$ such that the *marginals* of γ are P and Q.
- There is always the independent coupling: $\gamma(A \times B) = P(A)Q(B)$.
- But there are many others: the convex combinations of couplings are couplings.
- We write $\mathcal{C}(P, Q)$ for the set of couplings of P and Q.
- We can also define a coupling to be a pair of random variables R, S with distributions P, Q respectively.

Couplings

- How to relate two distributions? Think of a distribution as a pile of sand.
- We need to move some sand around to make the pile P look like Q.
- There are many different ways to do it. Each way is a "transport plan."
- A **coupling** of two distributions P, Q defined on X is a *joint* distribution γ on $X \times X$ such that the *marginals* of γ are P and Q.
- There is always the independent coupling: $\gamma(A \times B) = P(A)Q(B)$.
- But there are many others: the convex combinations of couplings are couplings.
- We write $\mathcal{C}(P, Q)$ for the set of couplings of P and Q.
- We can also define a coupling to be a pair of random variables R, S with distributions P, Q respectively.
- We can also define couplings easily between two different underlying spaces X and Y.

The W metrics

- A coupling γ defines a transport plan, how much does it cost?

The W metrics

- A coupling γ defines a transport plan, how much does it cost?
- If we measure the cost by a metric d we get

The W metrics

- A coupling γ defines a transport plan, how much does it cost?
- If we measure the cost by a metric d we get
- cost $= \int_{X\times X} d(x,y)\mathrm{d}\gamma$

The W metrics

- A coupling γ defines a transport plan, how much does it cost?
- If we measure the cost by a metric d we get
- cost $= \int_{X\times X} d(x,y)\mathrm{d}\gamma$
- We define a metric: $W_1(P,Q) = \inf_{\gamma\in\mathcal{C}(P,Q)} \int_{X\times X} d(x,y)\mathrm{d}\gamma$.

The W metrics

- A coupling γ defines a transport plan, how much does it cost?
- If we measure the cost by a metric d we get
- cost $= \int_{X\times X} d(x,y)\mathrm{d}\gamma$
- We define a metric: $W_1(P,Q) = \inf_{\gamma\in\mathcal{C}(P,Q)} \int_{X\times X} d(x,y)\mathrm{d}\gamma$.
- Kantorovich-Rubinstein duality: $\kappa = W_1$.

The W metrics

- A coupling γ defines a transport plan, how much does it cost?
- If we measure the cost by a metric d we get
- cost $= \int_{X\times X} d(x,y)\mathrm{d}\gamma$
- We define a metric: $W_1(P,Q) = \inf_{\gamma\in\mathcal{C}(P,Q)} \int_{X\times X} d(x,y)\mathrm{d}\gamma$.
- Kantorovich-Rubinstein duality: $\kappa = W_1$.
- $W_p(P,Q) = \inf_{\gamma\in\mathcal{C}(P,Q)}[\int_{X\times X}[d(x,y)]^p\mathrm{d}\gamma]^{\frac{1}{p}}$.

The W metrics

- A coupling γ defines a transport plan, how much does it cost?
- If we measure the cost by a metric d we get
- cost $= \int_{X\times X} d(x,y)\mathrm{d}\gamma$
- We define a metric: $W_1(P,Q) = \inf_{\gamma\in\mathcal{C}(P,Q)} \int_{X\times X} d(x,y)\mathrm{d}\gamma$.
- Kantorovich-Rubinstein duality: $\kappa = W_1$.
- $W_p(P,Q) = \inf_{\gamma\in\mathcal{C}(P,Q)}[\int_{X\times X}[d(x,y)]^p\mathrm{d}\gamma]^{\frac{1}{p}}$.
- Crucial point: if I find *any* coupling it gives an *upper bound* on W_1.

The W metrics

- A coupling γ defines a transport plan, how much does it cost?
- If we measure the cost by a metric d we get
- cost $= \int_{X\times X} d(x,y)\mathrm{d}\gamma$
- We define a metric: $W_1(P,Q) = \inf_{\gamma\in\mathcal{C}(P,Q)} \int_{X\times X} d(x,y)\mathrm{d}\gamma$.
- Kantorovich-Rubinstein duality: $\kappa = W_1$.
- $W_p(P,Q) = \inf_{\gamma\in\mathcal{C}(P,Q)}[\int_{X\times X}[d(x,y)]^p\mathrm{d}\gamma]^{\frac{1}{p}}$.
- Crucial point: if I find *any* coupling it gives an *upper bound* on W_1.
- We can define a map from a metric space (M,d) to the space $(\mathcal{P}(M),W_1)$ by $x\mapsto\delta_x$. This map is an *isometry*.

Bisimulation via couplings

- Recall MDP's

$$(S, \mathcal{A}, \forall a \in \mathcal{A}, P^a : S \longrightarrow \mathcal{D}(S), \mathcal{R} : \mathcal{A} \times S \longrightarrow \mathbf{R})$$

Bisimulation via couplings

- Recall MDP's

$$(S, \mathcal{A}, \forall a \in \mathcal{A}, P^a : S \longrightarrow \mathcal{D}(S), \mathcal{R} : \mathcal{A} \times S \longrightarrow \mathbf{R})$$

- An equivalence relation R on S is a **bisimulation** if sRt implies that $\forall a \in \mathcal{A}$ there is a *coupling* ω of $P^a(s)$ and $P^a(t)$ such that the *support* of ω is contained in R.

Computing the bisimulation metric

- Let $\mathcal{M}$ be the space of 1-bounded pseudometrics over S, ordered by $d_1 \leq d_2$ if $\forall x, y; d_2(x, y) \leq d_1(x, y)$.

Computing the bisimulation metric

- Let $\mathcal{M}$ be the space of 1-bounded pseudometrics over S, ordered by $d_1 \leq d_2$ if $\forall x, y; d_2(x, y) \leq d_1(x, y)$.
- This is a complete lattice.

Computing the bisimulation metric

- Let $\mathcal{M}$ be the space of 1-bounded pseudometrics over S, ordered by $d_1 \leq d_2$ if $\forall x, y; d_2(x, y) \leq d_1(x, y)$.
- This is a complete lattice.
- We define $T_K : \mathcal{M} \longrightarrow \mathcal{M}$ by

Computing the bisimulation metric

- Let $\mathcal{M}$ be the space of 1-bounded pseudometrics over S, ordered by $d_1 \leq d_2$ if $\forall x, y; d_2(x, y) \leq d_1(x, y)$.
- This is a complete lattice.
- We define $T_K : \mathcal{M} \longrightarrow \mathcal{M}$ by
- $T_K(d)(x, y) = \max_a[|\mathcal{R}(x, a)\mathcal{R}(y, a)| + \gamma W_d(P^a(x), P^a(y))]$

Computing the bisimulation metric

- Let $\mathcal{M}$ be the space of 1-bounded pseudometrics over S, ordered by $d_1 \leq d_2$ if $\forall x, y; d_2(x, y) \leq d_1(x, y)$.
- This is a complete lattice.
- We define $T_K : \mathcal{M} \longrightarrow \mathcal{M}$ by
- $T_K(d)(x, y) = \max_a [|\mathcal{R}(x, a)\mathcal{R}(y, a)| + \gamma W_d(P^a(x), P^a(y))]$
- This is a monotone function on $\mathcal{M}$.

Computing the bisimulation metric

- Let $\mathcal{M}$ be the space of 1-bounded pseudometrics over S, ordered by $d_1 \leq d_2$ if $\forall x, y; d_2(x, y) \leq d_1(x, y)$.
- This is a complete lattice.
- We define $T_K : \mathcal{M} \longrightarrow \mathcal{M}$ by
- $T_K(d)(x, y) = \max_a[|\mathcal{R}(x, a)\mathcal{R}(y, a)| + \gamma W_d(P^a(x), P^a(y))]$
- This is a monotone function on $\mathcal{M}$.
- We can find the bisimulation as the fixed point of T_K by iteration: $d^{\sim}$.

Computing the bisimulation metric

- Let $\mathcal{M}$ be the space of 1-bounded pseudometrics over S, ordered by $d_1 \leq d_2$ if $\forall x, y; d_2(x, y) \leq d_1(x, y)$.
- This is a complete lattice.
- We define $T_K : \mathcal{M} \longrightarrow \mathcal{M}$ by
- $T_K(d)(x, y) = \max_a[|\mathcal{R}(x, a)\mathcal{R}(y, a)| + \gamma W_d(P^a(x), P^a(y))]$
- This is a monotone function on $\mathcal{M}$.
- We can find the bisimulation as the fixed point of T_K by iteration: $d^{\sim}$.
- An important bound proved by Ferns et al. $|V^*(x) - V^*(y)| \leq d^{\sim}(x, y)$.

Basic goals in RL

- We are often dealing with *large* or *infinite* transition systems whose behaviour is probabilistic.

Basic goals in RL

- We are often dealing with *large* or *infinite* transition systems whose behaviour is probabilistic.
- The system responds to stimuli (actions) and moves to a new state probabilistically and outputs a (possibly) random reward.

Basic goals in RL

- We are often dealing with *large* or *infinite* transition systems whose behaviour is probabilistic.
- The system responds to stimuli (actions) and moves to a new state probabilistically and outputs a (possibly) random reward.
- We seek optimal policies for extracting the largest possible reward in expectation.

Basic goals in RL

- We are often dealing with *large* or *infinite* transition systems whose behaviour is probabilistic.
- The system responds to stimuli (actions) and moves to a new state probabilistically and outputs a (possibly) random reward.
- We seek optimal policies for extracting the largest possible reward in expectation.
- A plethora of algorithms and techniques, but the cost depends on the size of the state space.

Basic goals in RL

- We are often dealing with *large* or *infinite* transition systems whose behaviour is probabilistic.
- The system responds to stimuli (actions) and moves to a new state probabilistically and outputs a (possibly) random reward.
- We seek optimal policies for extracting the largest possible reward in expectation.
- A plethora of algorithms and techniques, but the cost depends on the size of the state space.
- Can we *learn* representations of the state space that accelerate the learning process?

Representation learning

- For large state spaces, learning value functions $S \times \mathcal{A} \longrightarrow \mathbf{R}$ is not feasible.

Representation learning

- For large state spaces, learning value functions $S \times \mathcal{A} \longrightarrow \mathbf{R}$ is not feasible.
- Instead we define a new space of *features* M and try to come up with an embedding $\phi : S \longrightarrow \mathbf{R}^M$.

Representation learning

- For large state spaces, learning value functions $S \times \mathcal{A} \longrightarrow \mathbf{R}$ is not feasible.
- Instead we define a new space of *features* M and try to come up with an embedding $\phi : S \longrightarrow \mathbf{R}^M$.
- Then we can try to use this to predict values associated with state,action pairs.

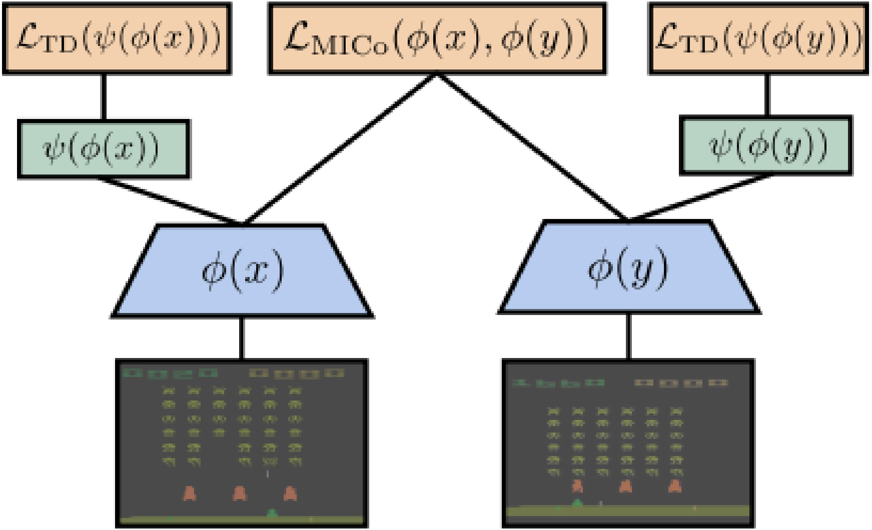
Representation learning

- For large state spaces, learning value functions $S \times \mathcal{A} \longrightarrow \mathbf{R}$ is not feasible.
- Instead we define a new space of *features* M and try to come up with an embedding $\phi : S \longrightarrow \mathbf{R}^M$.
- Then we can try to use this to predict values associated with state,action pairs.
- Representation learning means learning such a ϕ.

Representation learning

- For large state spaces, learning value functions $S \times \mathcal{A} \longrightarrow \mathbf{R}$ is not feasible.
- Instead we define a new space of *features* M and try to come up with an embedding $\phi : S \longrightarrow \mathbf{R}^M$.
- Then we can try to use this to predict values associated with state,action pairs.
- Representation learning means learning such a ϕ.
- The elements of M are the "features" that are chosen. They can be based on any kind of knowledge or experience about the task at hand.

Experimental setup



Experiments

- Added the MICo loss term to a variety of existing agents: all those available in the Dopamine Library; 5 in all.

Experiments

- Added the MICo loss term to a variety of existing agents: all those available in the Dopamine Library; 5 in all.
- Ran each game 5 times with new seeds so 300 runs for each agent.

Experiments

- Added the MICo loss term to a variety of existing agents: all those available in the Dopamine Library; 5 in all.
- Ran each game 5 times with new seeds so 300 runs for each agent.
- Each game is run for 200 million environment interactions.

Experiments

- Added the MICo loss term to a variety of existing agents: all those available in the Dopamine Library; 5 in all.
- Ran each game 5 times with new seeds so 300 runs for each agent.
- Each game is run for 200 million environment interactions.
- We look at final scores and learning curve.

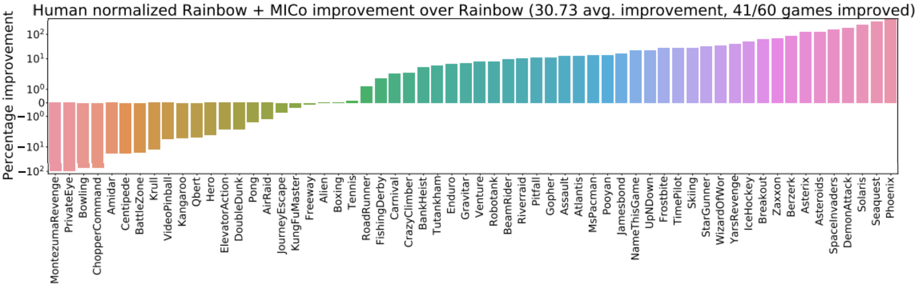
Experiments

- Added the MICo loss term to a variety of existing agents: all those available in the Dopamine Library; 5 in all.
- Ran each game 5 times with new seeds so 300 runs for each agent.
- Each game is run for 200 million environment interactions.
- We look at final scores and learning curve.
- We tried each agent with and without the MICo loss term on 60 different Atari games.

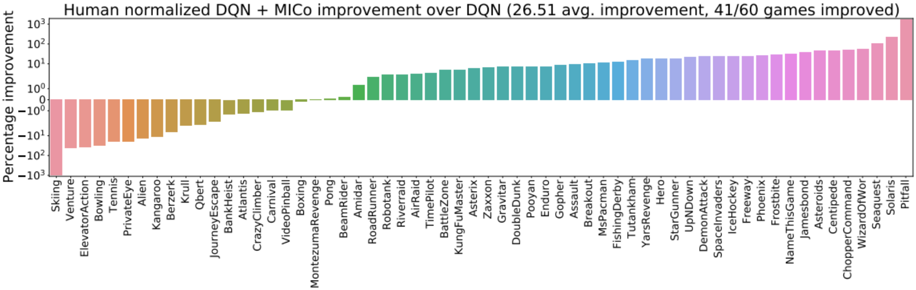
Experiments

- Added the MICo loss term to a variety of existing agents: all those available in the Dopamine Library; 5 in all.
- Ran each game 5 times with new seeds so 300 runs for each agent.
- Each game is run for 200 million environment interactions.
- We look at final scores and learning curve.
- We tried each agent with and without the MICo loss term on 60 different Atari games.
- *Every* agent performed better on about $\frac{2}{3}$ of the games.

Results for Rainbow



Results for DQN



Conclusions

- Bisimulation has a rich and venerable history.

Conclusions

- Bisimulation has a rich and venerable history.
- The metric analogue holds promise for quantitative reasoning and approximation as well as representation learning.

Conclusions

- Bisimulation has a rich and venerable history.
- The metric analogue holds promise for quantitative reasoning and approximation as well as representation learning.
- Research is alive and well and there are new areas where bisimulation is being “discovered”.

Conclusions

- Bisimulation has a rich and venerable history.
- The metric analogue holds promise for quantitative reasoning and approximation as well as representation learning.
- Research is alive and well and there are new areas where bisimulation is being “discovered”.
- Come talk to me any time about any of my research interests, or tennis or cricket or poetry or pure mathematics.

Conclusions

- Bisimulation has a rich and venerable history.
- The metric analogue holds promise for quantitative reasoning and approximation as well as representation learning.
- Research is alive and well and there are new areas where bisimulation is being “discovered”.
- Come talk to me any time about any of my research interests, or tennis or cricket or poetry or pure mathematics.
- Thanks to the Huawei Strategic Talent Scheme and to my host Chris Heunen.

Conclusions

- Bisimulation has a rich and venerable history.
- The metric analogue holds promise for quantitative reasoning and approximation as well as representation learning.
- Research is alive and well and there are new areas where bisimulation is being “discovered”.
- Come talk to me any time about any of my research interests, or tennis or cricket or poetry or pure mathematics.
- Thanks to the Huawei Strategic Talent Scheme and to my host Chris Heunen.
- Thanks to Ajitha Rajan and Gordon Winton for organizing this talk.

Conclusions

- Bisimulation has a rich and venerable history.
- The metric analogue holds promise for quantitative reasoning and approximation as well as representation learning.
- Research is alive and well and there are new areas where bisimulation is being “discovered”.
- Come talk to me any time about any of my research interests, or tennis or cricket or poetry or pure mathematics.
- Thanks to the Huawei Strategic Talent Scheme and to my host Chris Heunen.
- Thanks to Ajitha Rajan and Gordon Winton for organizing this talk.
- Special thanks to School of Informatics for providing such a wonderful welcoming environment.