Bisimulation Metric for Continuous MDPs

COMP 767

Pascale Gourdeau

March 24th, 2017

Overview

Definition of Continuous MDPs

Definition

A continuous MDP is a tuple (S, Σ, A, P, r) where

- (S, Σ) is a measurable space,

Definition of Continuous MDPs

Definition

A continuous MDP is a tuple (S, Σ, A, P, r) where

- (S, Σ) is a measurable space,
- A is a finite set of actions,

Definition of Continuous MDPs

Definition

A continuous MDP is a tuple (S, Σ, A, P, r) where

- (S, Σ) is a measurable space,
- A is a finite set of actions,
- $r : S \times A \to \mathbb{R}$ is a measurable reward function,

Definition of Continuous MDPs

Definition

A continuous MDP is a tuple (S, Σ, A, P, r) where

- (S, Σ) is a measurable space,
- A is a finite set of actions,
- $r : S \times A \to \mathbb{R}$ is a measurable reward function,
- $P : S \times A \times \Sigma \to [0, 1]$ is a labelled stochastic transition kernel:

Definition of Continuous MDPs

Definition

A continuous MDP is a tuple (S, Σ, A, P, r) where

- (S, Σ) is a measurable space,
- A is a finite set of actions,
- $r : S \times A \to \mathbb{R}$ is a measurable reward function,
- $P : S \times A \times \Sigma \to [0, 1]$ is a labelled stochastic transition kernel:
  - $\forall a \in A$, $\forall s \in S$, $P(s, a, \cdot) : \Sigma \to [0, 1]$ is a probability measure,

Definition of Continuous MDPs

Definition

A continuous MDP is a tuple (S, Σ, A, P, r) where

- (S, Σ) is a measurable space,
- A is a finite set of actions,
- $r : S \times A \to \mathbb{R}$ is a measurable reward function,
- $P : S \times A \times \Sigma \to [0, 1]$ is a labelled stochastic transition kernel:
  - $\forall a \in A$, $\forall s \in S$, $P(s, a, \cdot) : \Sigma \to [0, 1]$ is a probability measure,
  - $\forall a \in A$, $\forall X \in \Sigma$, $P(\cdot, a, X) : S \to [0, 1]$ is a measurable function.

Motivation

- Very large state space.

Motivation

- Very large state space.
- CMDPs $\rightarrow$ *continuous state space*.

Motivation

- Very large state space.
- CMDPs $\rightarrow$ *continuous state space*.
- First thought: discretization. Inconvenient because:

Motivation

- Very large state space.
- CMDPs $\rightarrow$ *continuous state space*.
- First thought: discretization. Inconvenient because:
  - We may aggregate states that are "physically" close together, but that have very different state values.

Motivation

- Very large state space.
- CMDPs $\rightarrow$ *continuous state space*.
- First thought: discretization. Inconvenient because:
  - We may aggregate states that are "physically" close together, but that have very different state values.
  - As we refine the space, we have an exponential blowup in the number of states.

Motivation

- Very large state space.
- CMDPs $\rightarrow$ *continuous state space*.
- First thought: discretization. Inconvenient because:
  - We may aggregate states that are "physically" close together, but that have very different state values.
  - As we refine the space, we have an exponential blowup in the number of states.
- Second thought: aggregate states with similar values.

Motivation

- Very large state space.
- CMDPs $\rightarrow$ *continuous state space*.
- First thought: discretization. Inconvenient because:
  - We may aggregate states that are "physically" close together, but that have very different state values.
  - As we refine the space, we have an exponential blowup in the number of states.
- Second thought: aggregate states with similar values.
  - Bisimulation!

Definitions

Definition

A *metric* on a set X is a map $d : X \times X \to [0, \infty)$ such that for all $x, y, z \in X$:

1. $x = y \iff d(x, y) = 0$
2. $d(x, y) = d(y, x)$
3. $d(x, y) \leq d(x, z) + d(z, y)$

Definition

We say that the tuple (X,d) where X is a set with a metric $d : X \times X \to [0, \infty)$ is a *metric space*.

Separable Metric Space

Definition

A metric space (X, d) is said to be *separable* if it has some countable dense subset.

Separable Metric Space

Definition
A metric space (X, d) is said to be *separable* if it has some countable dense subset.

Definition
A metric space (X, d) is said to be *complete* if every Cauchy sequence converges.

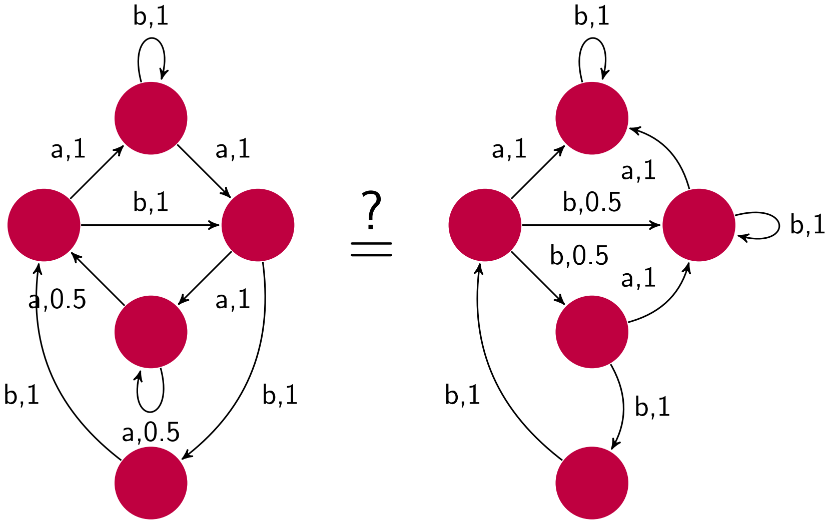
Separable Metric Space

Definition
A metric space (X, d) is said to be *separable* if it has some countable dense subset.

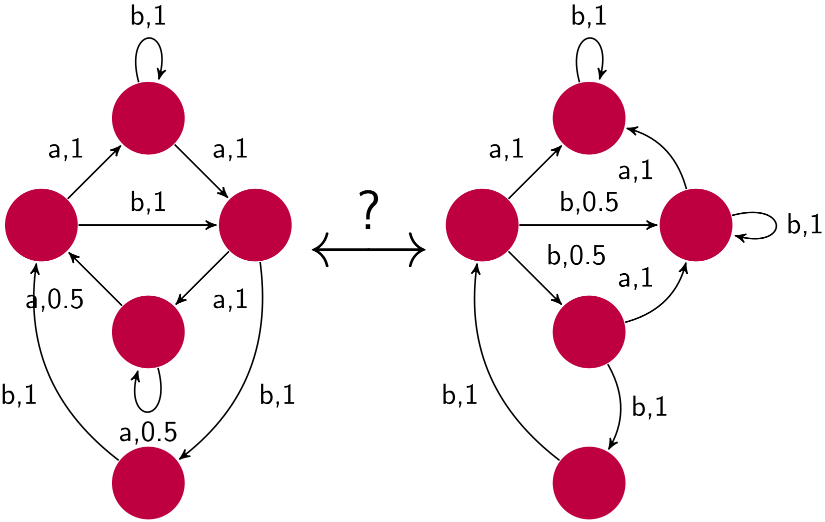
Definition
A metric space (X, d) is said to be *complete* if every Cauchy sequence converges.

Definition
A metric space (X, d) is said to be *Polish* if it is both separable and complete.

Motivation



Motivation



Bisimulation

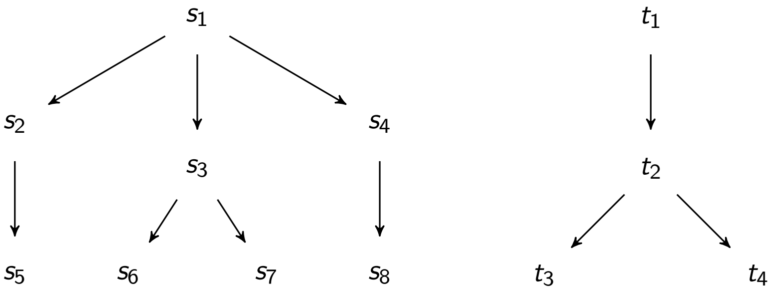
- Originally due to Park (1981) and extended to probabilistic systems by Larsen and Skou (1991).
- Abstract notion of *behavioural equivalence* between processes.
- More flexible and subtle than *isomorphism*.
- If I have two bisimilar systems, I can replace one by the other and no test (sequence of experiments) can distinguish them.

Bisimulation as a game

- Two-way simulation
- Bisimulation

Bisimulation as a game

- Two-way simulation
- Bisimulation



Bisimulation on CMDPs

MDP assumptions:

- S is a Polish space with its Borel σ-algebra Σ.

Bisimulation on CMDPs

MDP assumptions:

- S is a Polish space with its Borel σ-algebra Σ.
- $\mathrm{img}(r) \subseteq [0, 1]$

Bisimulation on CMDPs

MDP assumptions:

- S is a Polish space with its Borel σ-algebra Σ.
- $\text{img}(r) \subseteq [0, 1]$
- For each $a \in A$, $r(\cdot, a)$ is continuous on S.

Bisimulation on CMDPs

MDP assumptions:

- S is a Polish space with its Borel σ-algebra Σ.
- $\text{img}(r) \subseteq [0, 1]$
- For each $a \in A$, $r(\cdot, a)$ is continuous on S.
- For each $a \in A$, P_s^a is continuous as a function of s.

Bisimulation on CMDPs

MDP assumptions:

- S is a Polish space with its Borel σ-algebra Σ.
- $\text{img}(r) \subseteq [0, 1]$
- For each $a \in A$, $r(\cdot, a)$ is continuous on S.
- For each $a \in A$, P_s^a is continuous as a function of s.

Bisimulation on CMDPs

MDP assumptions:

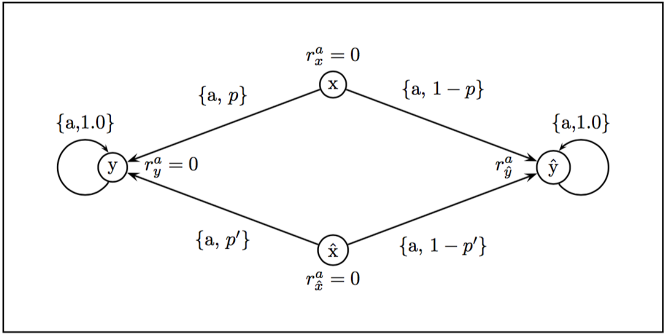
- S is a Polish space with its Borel σ-algebra Σ.
- $\text{img}(r) \subseteq [0, 1]$
- For each $a \in A$, $r(\cdot, a)$ is continuous on S.
- For each $a \in A$, P_s^a is continuous as a function of s.

Definition

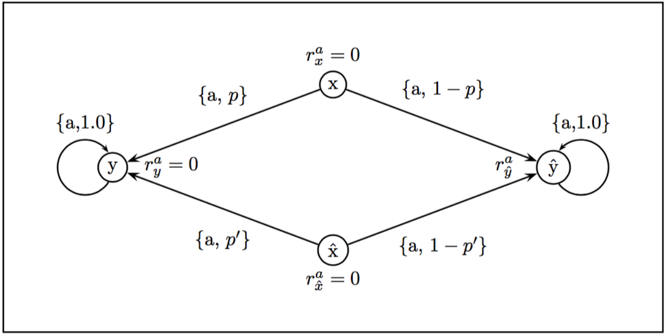
Let (S, Σ, A, P, r) be an MDP satisfying the above assumptions. An equivalence relation R on S is a *bisimulation relation* if and only if it satisfies

$$sRs' \iff \begin{array}{l} \text{for every } a \in A,\ r_s^a = r_{s'}^a \text{ and} \\ \text{for every } X \in \Sigma(R),\ P_s^a(X) = P_{s'}^a(X). \end{array}$$

Example of bisimulation on MDPs

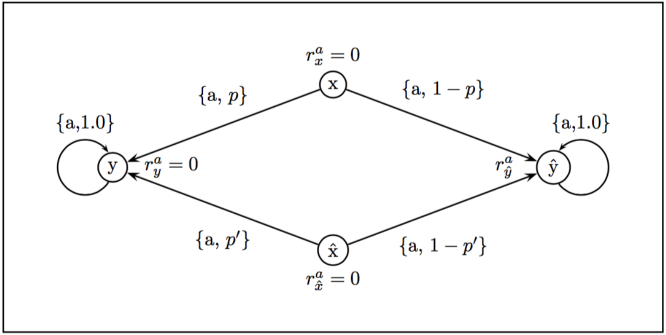


Example of bisimulation on MDPs



If $r_{\hat{y}} = 0$, all states are bisimilar.

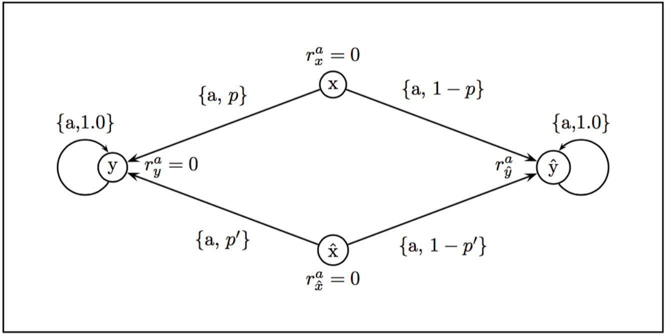
Example of bisimulation on MDPs



If $r_{\hat{y}} = 0$, all states are bisimilar.
If $r_{\hat{y}} > 0$,

- No state is bisimilar to $\hat{y}$ (other than itself),

Example of bisimulation on MDPs

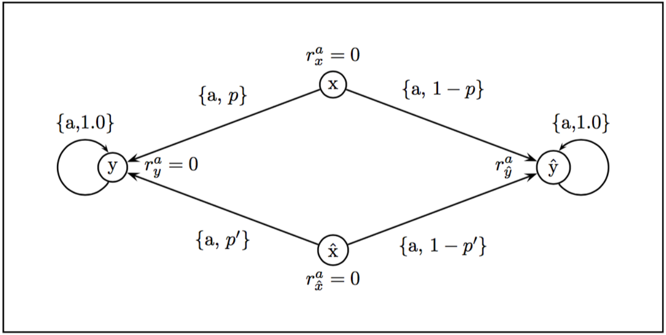


If $r_{\hat{y}} = 0$, all states are bisimilar.

If $r_{\hat{y}} > 0$,

- No state is bisimilar to $\hat{y}$ (other than itself),
- $x \sim \hat{x} \iff p = p'$,

Example of bisimulation on MDPs

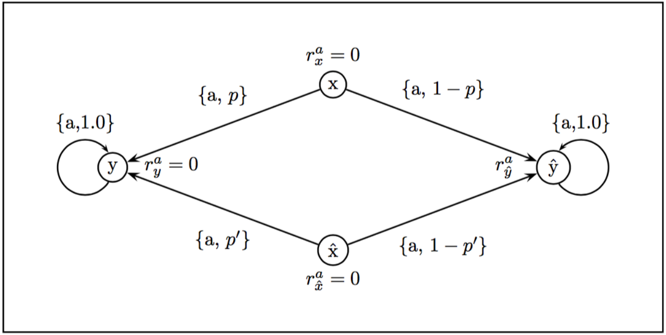


If $r_{\hat{y}} = 0$, all states are bisimilar.

If $r_{\hat{y}} > 0$,

- No state is bisimilar to $\hat{y}$ (other than itself),
- $x \sim \hat{x} \iff p = p'$,
- $x \sim y \iff p = 1$,

Example of bisimulation on MDPs



If $r_{\hat{y}} = 0$, all states are bisimilar.

If $r_{\hat{y}} > 0$,

- No state is bisimilar to $\hat{y}$ (other than itself),
- $x \sim \hat{x} \iff p = p'$,
- $x \sim y \iff p = 1$,
- $\hat{x} \sim y \iff p' = 1$.

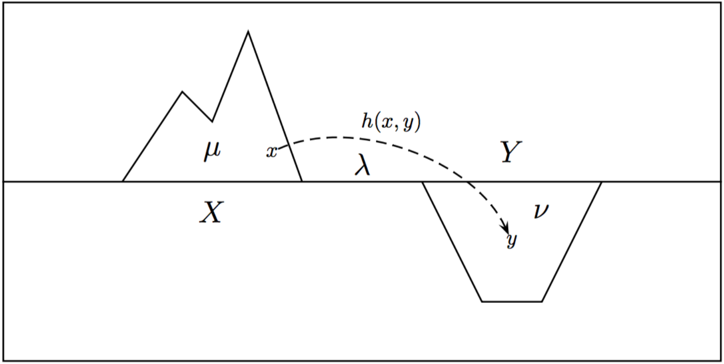
Kantorovich Metric

Definition

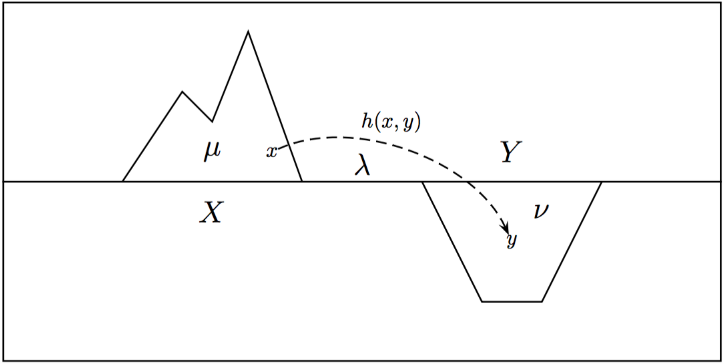
Let (S, d) be a Polish metric space, h a bounded pseudo-metric on S that is lower semi-continuous on $S \times S$ and $Lip(h)$ the set of all bounded functions $f : S \to \mathbb{R}$ that are measurable w.r.t. $\mathcal{B}(S)$ and satisfy the Lipschitz condition $f(x) - f(y) \leq h(x, y)$ for every $x, y \in S$. Given two probability measures P and Q, the *Kantorovich distance* $T_K(h)$ is defined by

$$T_K(h)(P, Q) = \sup_{f \in Lip(h)} (P(f) - Q(f)) = \sup_{f \in Lip(h)} \left(\int f dP - \int f dQ \right)$$

Link to Transportation Problem



Link to Transportation Problem



Goal: determine a plan for transferring all the mass from X to Y while minimizing the cost.

Kantorovich Metric

Theorem (Kantorovich-Rubinstein Duality Theorem)

$$T_K(h)(P,Q) = \sup_{f \in Lip(h)} (P(f) - Q(f)) = \inf_{\lambda \in \Lambda(P,Q)} h(\lambda)$$

Kantorovich Metric

Theorem (Kantorovich-Rubinstein Duality Theorem)

$$T_K(h)(P,Q) = \sup_{f \in Lip(h)} (P(f) - Q(f)) = \inf_{\lambda \in \Lambda(P,Q)} h(\lambda)$$

Lemma

Let $\mathfrak{lsc}_{\mathfrak{m}}$ be the set of bounded pseudometrics on S which are lower semi-continuous on $S \times S$, $h \in \mathfrak{lsc}_{\mathfrak{m}}$ and $Rel(h)$ be the kernel of h. Then

$$T_K(h)(P,Q) = 0 \iff P(X) = Q(X)\ \forall X \in \Sigma(Rel(h)) \ .$$

What is a Bisimulation Metric?

Definition

A pseudometric $\rho : S \times S \to [0, +\infty)$ on the states of an MDP is a *bisimulation metric* if it satisfies

$$\rho(s, s') = 0 \iff s \sim s'.$$

A Map on Pseudometrics

Let (S, Σ, A, P, r) be an MDP satisfying the conditions in the previous slide, $c \in (0, 1)$ be a discount factor and $\mathfrak{lsc}_{\mathfrak{m}}$ be the set of bounded pseudometrics on S which are lower semi-continuous on $S \times S$. Define the map $F : \mathfrak{lsc}_{\mathfrak{m}} \to \mathfrak{lsc}_{\mathfrak{m}}$:

$$F(h)(s, s') = \max_{a \in A} [(1 - c)|r_s^a - r_{s'}^a| + cT_K(h)(P_s^a, P_{s'}^a)]$$

Bisimulation Metric through Fixed Point

The map

$$F(h)(s, s') = \max_{a \in A} [(1-c)|r_s^a - r_{s'}^a| + cT_K(h)(P_s^a, P_{s'}^a)]$$

has a unique fixed point $\rho^* : S \times S \to [0, 1]$.

Bisimulation Metric through Fixed Point

The map

$$F(h)(s, s') = \max_{a \in A} [(1 - c)|r_s^a - r_{s'}^a| + cT_K(h)(P_s^a, P_{s'}^a)]$$

has a unique fixed point $\rho^* : S \times S \to [0, 1]$. This ρ^*

- is a bisimulation metric;
- scales with rewards.

Bisimulation Metrics are Optimal Value Functions

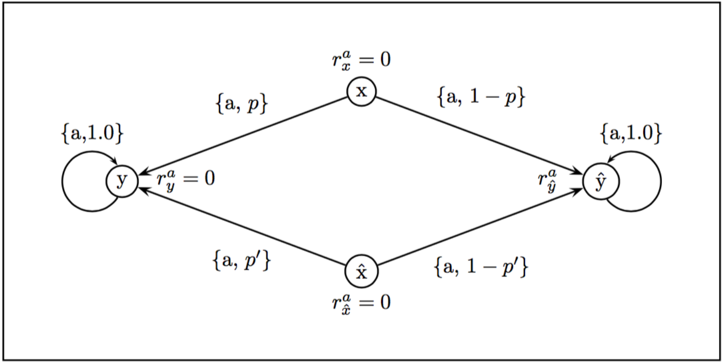
$$|V^*(s) - V^*(s')| \leq \frac{1}{1-c}\rho_c^*(s, s')$$

Bisimulation Metrics are Optimal Value Functions

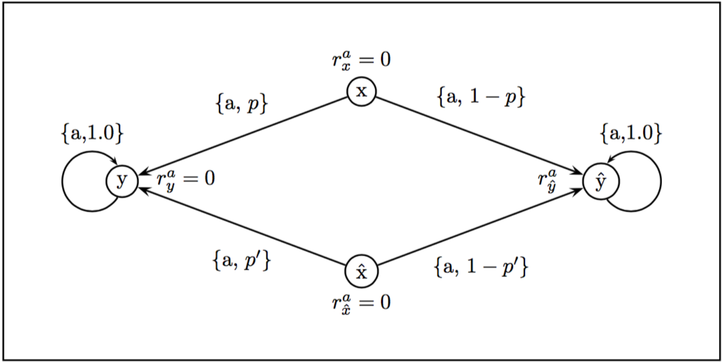
$$|V^*(s) - V^*(s')| \leq \frac{1}{1-c}\rho_c^*(s, s')$$

- The closer the distance (relative to bisimilarity) the more likely they will share optimal value functions (and hence policies).
- Aggregating states that are close in behaviour (w.r.t. bisimilarity) implies aggregating states with similar value functions.

Example of bisimulation metric on MDPs



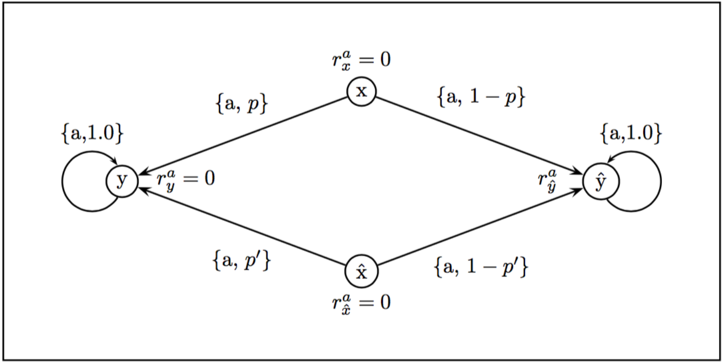
Example of bisimulation metric on MDPs



Note that:

1. There is only one action,

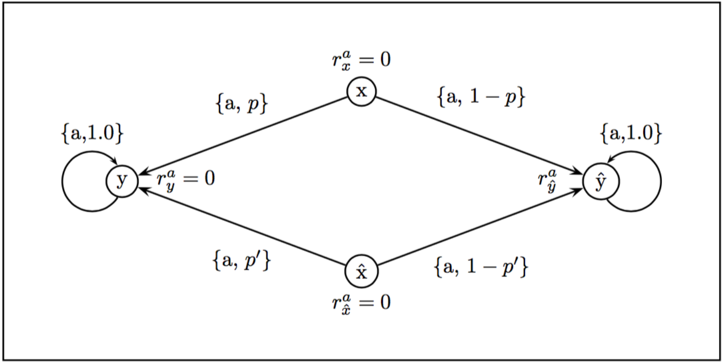
Example of bisimulation metric on MDPs



Note that:

1. There is only one action,
2. $T_K(\rho^*)(\delta_x, \delta_y) = \rho^*(x, y)$,

Example of bisimulation metric on MDPs



Note that:

1. There is only one action,
2. $T_K(\rho^*)(\delta_x, \delta_y) = \rho^*(x, y)$,
3. $F(\rho^*)(s, s') = (\rho^*)(s, s')$ and ρ^* is unique.