Temporally Extended Metrics for Markov Decision Processes

Philip Amortila†  Marc G. Bellemare◦*  Prakash Panangaden†  Doina Precup†
†McGill University  °Google Brain  *CIFAR Fellow

SafeAI at AAAI 2019
Jan 27 2019
Bisimulation is a canonical example of a **safe state abstraction**
- States with different risk properties will never get clustered
Motivation – Safe state abstraction for MDPs

- Bisimulation is a canonical example of a **safe state abstraction**
  - States with different risk properties will never get clustered
- Bisimulation is too stringent – any $\varepsilon$ perturbation to the model can cause states to lose their bisimilarity
Motivation – Safe state abstraction for MDPs

- Bisimulation is a canonical example of a **safe state abstraction**
  - States with different risk properties will never get clustered
- Bisimulation is too stringent – any $\varepsilon$ perturbation to the model can cause states to lose their bisimilarity
- Quantitative analogues (the *bisimulation metrics*) are expensive to compute and require the true model
Bisimulation is a canonical example of a **safe state abstraction**
- States with different risk properties will never get clustered

Bisimulation is too stringent – any $\varepsilon$ perturbation to the model can cause states to lose their bisimilarity

Quantitative analogues (the *bisimulation metrics*) are expensive to compute and require the true model

Motivation: investigate alternative metrics for behavioural equivalence
Contributions

- We provide an alternative characterization of bisimulation via couplings using this, we generalize bisimulation by making it depend on arbitrary comparisons between states instead of strict reward matching. We develop temporally extended metrics, which reflect the extent to which the difference between states is preserved through the course of transitions. We provide formal safety bounds and compare with bisimulation metrics by examining the dynamics computed by the two metrics.
Contributions

- We provide an alternative characterization of bisimulation via couplings.
- Using this, we generalize bisimulation by making it depend on arbitrary comparisons between states instead of strict reward matching.

Philip Amortila
Temporally Extended Metrics
Jan 27 2019
Contributions

- We provide an alternative characterization of bisimulation via couplings.
- Using this, we generalize bisimulation by making it depend on arbitrary comparisons between states instead of strict reward matching.
- Develop *temporally extended metrics*, which reflect the extent to which the difference between states is preserved through the course of transitions.
Contributions

- We provide an alternative characterization of bisimulation via couplings
- Using this, we generalize bisimulation by making it depend on arbitrary comparisons between states instead of strict reward matching
- Develop *temporally extended metrics*, which reflect the extent to which the difference between states is preserved through the course of transitions.
  - We provide formal safety bounds and compare with bisimulation metrics by examining the dynamics computed by the two metrics