State Abstractions
Notations and Setup

- MDP $M = (S, A, P, R, \gamma)$
- Abstraction $\phi : S \rightarrow \phi(S)$
Notations and Setup

• MDP $M = (S, A, P, R, \gamma)$

• Abstraction $\phi : S \rightarrow \phi(S)$

• Surjection — aggregate states and treat as equivalent
Notations and Setup

- MDP $M = (S, A, P, R, \gamma)$
- Abstraction $\phi : S \rightarrow \phi(S)$
- Surjection — aggregate states and treat as equivalent
- Are the aggregated states really equivalent?
Notations and Setup

• MDP \( M = (S, A, P, R, \gamma) \)
• Abstraction \( \phi : S \rightarrow \phi(S) \)
• Surjection — aggregate states and treat as equivalent
• Are the aggregated states really equivalent?
• Do they have the same…
  • optimal action?
  • \( Q^* \) values?
  • dynamics and rewards?
Abstraction hierarchy

An abstraction $\phi$ is … if … $\forall \ s^{(1)}, s^{(2)}$ where $\phi(s^{(1)}) = \phi(s^{(2)})$
Abstraction hierarchy

An abstraction $\phi$ is … if … $\forall s^{(1)}, s^{(2)}$ where $\phi(s^{(1)}) = \phi(s^{(2)})$

• $\pi^*\text{-irrelevant: } \exists \pi_M^* \text{ s.t. } \pi_M^*(s^{(1)}) = \pi_M^*(s^{(2)})$
Abstraction hierarchy

An abstraction $\phi$ is ... if ... $\forall s^{(1)}, s^{(2)}$ where $\phi(s^{(1)}) = \phi(s^{(2)})$

- $\pi^*$-irrelevant: $\exists \pi^*_M$ s.t. $\pi^*_M(s^{(1)}) = \pi^*_M(s^{(2)})$
- $Q^*$-irrelevant: $\forall a, Q^*_M(s^{(1)}, a) = Q^*_M(s^{(2)}, a)$
Abstraction hierarchy

An abstraction $\phi$ is ... if ... $\forall s^{(1)}, s^{(2)}$ where $\phi(s^{(1)}) = \phi(s^{(2)})$

- $\pi^*$-irrelevant: $\exists \pi_M^*$ s.t. $\pi_M^*(s^{(1)}) = \pi_M^*(s^{(2)})$

- $Q^*$-irrelevant: $\forall a, Q_M^*(s^{(1)}, a) = Q_M^*(s^{(2)}, a)$

- Model-irrelevant: $\forall a \in A$
  
  - (bisimulation) $\forall a \in A, x' \in \phi(S)$, $R(s^{(1)}, a) = R(s^{(2)}, a)$
  
  $P(x' | s^{(1)}, a) = P(x' | s^{(2)}, a)$

$$\sum_{s' \in \phi^{-1}(x')} P(s' | s^{(1)}, a)$$
Abstraction hierarchy

An abstraction $\phi$ is … if … $\forall \; s^{(1)}, s^{(2)}$ where $\phi(s^{(1)}) = \phi(s^{(2)})$

- $\pi^*$-irrelevant: $\exists \; \pi^*_M$ s.t. $\pi^*_M(s^{(1)}) = \pi^*_M(s^{(2)})$

- $Q^*$-irrelevant: $\forall \; a \in A, \; Q^*_M(s^{(1)}, a) = Q^*_M(s^{(2)}, a)$

- Model-irrelevant: $\forall \; a \in A$, $x' \in \phi(S)$,
  (bisimulation) $\forall \; a \in A, \; x' \in \phi(S), \; P(x' \mid s^{(1)}, a) = P(x' \mid s^{(2)}, a)$

$$\sum_{s' \in \phi^{-1}(x')} P(s' \mid s^{(1)}, a)$$

Theorem: Model-irrelevance $\Rightarrow Q^*$-irrelevance $\Rightarrow \pi^*$-irrelevance
Why not $P(s' \mid s^{(1)}, a) = P(s' \mid s^{(2)}, a)$?
Why not $P(s' \mid s^{(1)}, a) = P(s' \mid s^{(2)}, a)$?

\[
P((x', z') \mid (x, z), a) = P_M(x' \mid x, a) \cdot P_C(z' \mid z)
\]
Why not $P(s' \mid s^{(1)}, a) = P(s' \mid s^{(2)}, a)$?

\[
P((x', z') \mid (x, z), a) = P_M(x' \mid x, a) \cdot P_C(z' \mid z)
\]

MDP $M$    Markov chain $C$

(x, z^{(1)}) and (x, z^{(2)}) cannot be aggregated under the $s'$-based condition
Why not $P(s' \mid s^{(1)}, a) = P(s' \mid s^{(2)}, a)$?

$P((x', z') \mid (x, z), a) = P_M(x' \mid x, a) \cdot P_C(z' \mid z)$

Integrated out by bisimulation

$(x, z^{(1)})$ and $(x, z^{(2)})$ cannot be aggregated under the $s'$-based condition
The abstract MDP implied by bisimulation

\[ \phi \text{ is bisimulation: } R(s^{(1)}, a) = R(s^{(2)}, a), \quad P(x' \mid s^{(1)}, a) = P(x' \mid s^{(2)}, a) \]
The abstract MDP implied by bisimulation

\( \phi \) is bisimulation: \( R(s^{(1)}, a) = R(s^{(2)}, a) , \ P(x' | s^{(1)}, a) = P(x' | s^{(2)}, a) \)

- MDP \( M_\phi = (\phi(S), A, P_\phi, R_\phi, \gamma') \)
The abstract MDP implied by bisimulation

\( \phi \) is bisimulation: \( R(s^{(1)}, a) = R(s^{(2)}, a) \), \( P(x' \mid s^{(1)}, a) = P(x' \mid s^{(2)}, a) \)

- MDP \( M_\phi = (\phi(S), A, P_\phi, R_\phi, \gamma) \)
- For any \( x \in \phi(S), a \in A, x' \in \phi(S) \)
  - \( R_\phi(x, a) = R(s, a) \) for any \( s \in \phi^{-1}(x) \)
  - \( P_\phi(x' \mid x, a) = P(x' \mid s, a) \) for any \( s \in \phi^{-1}(x) \)
The abstract MDP implied by bisimulation

φ is bisimulation: \( R(s^{(1)}, a) = R(s^{(2)}, a) \), \( P(x' | s^{(1)}, a) = P(x' | s^{(2)}, a) \)

- MDP \( M_\phi = (\phi(S), A, P_\phi, R_\phi, \gamma) \)
- For any \( x \in \phi(S), a \in A, x' \in \phi(S) \)
  - \( R_\phi(x, a) = R(s, a) \) for any \( s \in \phi^{-1}(x) \)
  - \( P_\phi(x' | x, a) = P(x' | s, a) \) for any \( s \in \phi^{-1}(x) \)
- No way to distinguish between the two routes:

\[ \begin{align*}
M & \quad \text{generate data} \rightarrow \{(s, a, r, s')\} \\
M_\phi & \quad \text{generate data} \rightarrow \{((\phi(s), a, r, \phi(s'))})
\end{align*} \]
Implications of bisimulation

• $Q^*$-irrelevance

• Plan in $M_\phi$ and get $Q^*_{M_\phi}$ (dimension: $\phi(S) \times A$)
Implications of bisimulation

• $Q^*$-irrelevance
  
  • Plan in $M_\phi$ and get $Q_{M_\phi}^*$ (dimension: $\phi(S) \times A$)
  
  • Lift $Q_{M_\phi}^*$ from $\phi(S)$ to $S$ (populate aggregated states with the same value)
Implications of bisimulation

- $Q^*$-irrelevance
  - Plan in $M_\phi$ and get $Q^*_{M_\phi}$ (dimension: $\phi(S) \times A$)
  - Lift $Q^*_{M_\phi}$ from $\phi(S)$ to $S$ (populate aggregated states with the same value)
  - Useful notation: $\Phi$ is a $|\phi(S)| \times |S|$ matrix, with
    \[
    \Phi(x, s) = \mathbb{1}[\phi(s) = x]
    \]
Implications of bisimulation

• $Q^*$-irrelevance

• Plan in $M_\phi$ and get $Q^*_{M_\phi}$ (dimension: $\phi(S) \times A$)

• *Lift* $Q^*_{M_\phi}$ from $\phi(S)$ to $S$ (populate aggregated states with the same value)

• Useful notation: $\Phi$ is a $|\phi(\mathcal{S})| \times |\mathcal{S}|$ matrix, with

\[
\Phi(x, s) = \mathbb{1}[\phi(s) = x]
\]

• lifting a state-value function: $[V^*_{M_\phi}]_M = \Phi^T V^*_{M_\phi}$
Implications of bisimulation

• $Q^*$-irrelevance

• Plan in $M_\phi$ and get $Q^*_{M_\phi}$ (dimension: $\phi(S) \times A$)

• Lift $Q^*_{M_\phi}$ from $\phi(S)$ to $S$ (populate aggregated states with the same value)

• Useful notation: $\Phi$ is a $|\phi(S)| \times |S|$ matrix, with

\[
\Phi(x, s) = \mathbb{I}[\phi(s) = x]
\]

• lifting a state-value function: $[V^*_{M_\phi}]_M = \Phi^T V^*_{M_\phi}$

• collapsing the transition distribution: $\Phi P(s, a)$
Implications of bisimulation

• $Q^*$-irrelevance

• Plan in $M_\phi$ and get $Q^*_{M_\phi}$ (dimension: $\phi(S) \times A$)

• Lift $Q^*_{M_\phi}$ from $\phi(S)$ to $S$ (populate aggregated states with the same value)

• Useful notation: $\Phi$ is a $|\phi(\mathcal{S})| \times |\mathcal{S}|$ matrix, with

$$\Phi(x, s) = \mathbb{I}[\phi(s) = x]$$

• lifting a state-value function: $[V^*_{M_\phi}]_M = \Phi^T V^*_{M_\phi}$

• collapsing the transition distribution: $\Phi P(s, a)$

• Claim: $\left[Q^*_{M_\phi}\right]_M = Q^*_M$ (proof on board)
Implications of bisimulation

• $Q^*$-irrelevance
Implications of bisimulation

• $Q^*$-irrelevance

• $Q_M^{\pi}$ is preserved for any $\pi$ lifted from an abstract policy
Implications of bisimulation

• $Q^*$-irrelevance

• $Q_M^\pi$ is preserved for any $\pi$ lifted from an abstract policy

• Given any lifted $\pi$, distribution over reward seq. is preserved (assuming reward is deterministic function of $s, a$) (Is this sufficient?)

  • Can be extended to features of state to define a notion of saliency (think: what happens when the reward criterion is missing?)

  • For deeper thoughts along these lines, read Erik Talvitie’s thesis
Abstraction induces an equivalence relation

- Reflexivity, symmetry, transitivity
Abstraction induces an equivalence relation

- Reflexivity, symmetry, transitivity
- Equivalence notion is a canonical representation of abstraction (i.e., what symbol you associate with each abstract state doesn’t matter; what matters is which states are aggregated together)
Abstraction induces an **equivalence relation**

- Reflexivity, symmetry, transitivity

- Equivalence notion is a canonical representation of abstraction (i.e., what symbol you associate with each abstract state doesn’t matter; what matters is which states are aggregated together)

- Partition the state space into **equivalence classes**
Abstraction induces an equivalence relation

- Reflexivity, symmetry, transitivity
- Equivalence notion is a canonical representation of abstraction (i.e., what symbol you associate with each abstract state doesn’t matter; what matters is which states are aggregated together)
- Partition the state space into equivalence classes
- Coarsest bisimulation is unique (proof), but is NP-hard to find
  - Is the hardness interesting from a learning perspective?
Extension to handle action aggregation

Figure from: Ravindran & Barto. Approximate Homomorphisms: A framework for non-exact minimization in Markov Decision Processes. 2004.
Approximate abstractions

1. $\phi$ is an $\epsilon_{\pi^*}$-approximate $\pi^*$-irrelevant abstraction, if there exists an abstract policy $\pi : \phi(S) \rightarrow A$, such that $\|V_M^* - V_M^{[\pi]} \|_{\infty} \leq \epsilon_{\pi^*}$.

2. $\phi$ is an $\epsilon_Q^*$-approximate $Q^*$-irrelevant abstraction if there exists an abstract $Q$-value function $f : \phi(S) \times A \rightarrow \mathbb{R}$, such that $\|[f]_M - Q_M^* \|_{\infty} \leq \epsilon_{Q^*}$.

3. $\phi$ is an $(\epsilon_R, \epsilon_P)$-approximate model-irrelevant abstraction if for any $s^{(1)}$ and $s^{(2)}$ where $\phi(s^{(1)}) = \phi(s^{(2)})$, $\forall a \in A$,

$$|R(s^{(1)}, a) - R(s^{(2)}, a)| \leq \epsilon_R, \quad \|\Phi P(s^{(1)}, a) - \Phi P(s^{(2)}, a)\|_1 \leq \epsilon_P. \quad (3)$$

**Theorem 2.** (1) If $\phi$ is an $(\epsilon_R, \epsilon_P)$-approximate model-irrelevant abstraction, then $\phi$ is also an approximate $Q^*$-irrelevant abstraction with approximation error $\epsilon_{Q^*} = \frac{\epsilon_R}{1-\gamma} + \frac{\gamma \epsilon_P R_{\text{max}}}{2(1-\gamma)^2}$.

(2) If $\phi$ is an $\epsilon_{Q^*}$-approximate $Q^*$-irrelevant abstraction, then $\phi$ is also an approximate $\pi^*$-irrelevant abstraction with approximation error $\epsilon_{\pi^*} = 2\epsilon_{Q^*} / (1 - \gamma)$. 