A Survey of State Abstraction Techniques for Markov Decision Processes

Austin Mohr
Southern Illinois University at Carbondale
Carbondale, IL 62901
E-mail: austinmohr@gmail.com

Abstract

In many real-world applications of Markov Decision Processes (MPDs), the number of states is so large as to be infeasible for computation. State abstraction is a means by which similar states are aggregated, resulting in reduction of the state space size. Under certain conditions, it can be shown that the optimal policy derived from the "abstract state space" will also be optimal for the "ground state space". We consider the five abstraction schemes presented in [1] and the related theoretical results on optimality therein.

1 Markov Decision Processes

Li et al. present the following definition for a Markov Decision Process (MPD) in [1].

Definition 1.1. A Markov Decision Process (MPD) is a five-tuple $\langle S, A, P, R, \gamma \rangle$ where

- S is a finite set of states,
- A is a finite set of actions,
- P is the transition function with $P^a_{ss'}$ denoting the next-state distribution after taking action a in state s,
- R is a bounded reward function with R^a_s denoting the expected immediate reward gained by taking action a in state s, and

- $\gamma \in [0, 1]$ is a discount factor.

In solving an MPD, we seek a mapping $\pi : S \rightarrow A$ (called a policy) where, for each state $s \in S$, $\pi(s)$ gives the action $a \in A$ of maximum expected utility under transition function P and reward function R. We denote this optimal policy by π^*. In order to later discuss specific types of abstractions, we also require the following definition.

Definition 1.2. Given a policy π, we define the state-value function $V^\pi(s)$ as the expected cumulative reward received by executing π from state s. Similarly, the state-action value function $Q^\pi(s, a)$ is the expected cumulative reward received by taking action a in state s and following π thereafter.

As a matter of notation, we denote V^{π^*} by V^* and Q^{π^*} by Q^*.

2 State Abstraction

Loosely speaking, state abstraction is any technique which transforms a given state space S to another (hopefully smaller) abstract state space $\bar{S}$ while still preserving some properties of S. Generally, the property we are most concerned with preserving is the optimality of policies. That is, given a policy $\bar{\pi}^*$ which is optimal for the abstract state space, we want to know whether we can easily derive an optimal policy for the ground state space. If optimality is not feasible, then we may consider approximate techniques and ask the same question regarding policies within a certain bound of optimal. More formally, we consider the definition of state abstraction presented in [1].

Definition 2.1. Let $M = \langle S, A, P, R, \gamma \rangle$ be the ground MDP and its abstract version be $\bar{M} = \langle \bar{S}, A, \bar{P}, \bar{R}, \gamma \rangle$. Define the *abstraction function* as $\phi : S \rightarrow \bar{S}$ where $\phi(s)$ is the abstract state corresponding to ground state $s \in S$ and the inverse image $\phi^{-1}(\bar{s}) = \{s \in S | \phi(s) = \bar{s}\}$. Note that under these assumptions, $\{\phi^{-1}(\bar{s}) | \bar{s} \in \bar{S}\}$ partitions the ground state space S. To guarantee $\bar{P}$ and $\bar{R}$ are well-defined, a weighting function is needed: $w : S \rightarrow [0, 1]$ where, for each $\bar{s} \in \bar{S}$, $\sum_{s \in \phi^{-1}(\bar{s})} w(s) = 1$. We can now define the transition and reward functions for $\bar{M}$ as follows:

$$\bar{R}^a_{\bar{s}} = \sum_{s \in \phi^{-1}(\bar{s})} w(s) R^a_s,$$

$$\bar{P}^a_{\bar{s}\bar{s}'} = \sum_{s \in \phi^{-1}(\bar{s})} \sum_{s \in \phi^{-1}(\bar{s}')} w(s) P^a_{ss'}.$$

Under this definition, state aggregation is accomplished by the use of a function that maps several distinct states in the ground state space to a single state in the abstract state space. The weighting function can be viewed as a function which, given a ground state s, measures the extent to which s contributes to $\phi(s)$. With this understanding, the formulas for $\bar{R}$ and $\bar{P}$ follow directly.

3 Selected Forms of State Abstraction

Depending on which properties of the ground MPD we wish to preserve, we will necessarily arrive at different types of abstraction. The following five abstraction methods (and the properties each preserves) are presented in [1].

Definition 3.1. Given an MPD $M = \langle S, A, P, R, \gamma \rangle$ and any states $s_1, s_2 \in S$, we define five types of abstraction as below, with an arbitrary but fixed weighting function $w(s)$.

1. A *model-irrelevance abstraction* ϕ_{model} is such that for any action a and any abstract state $\bar{s}$, $\phi_{model}(s_1) = \phi_{model}(s_2)$ implies $R^a_{s_1} = R^a_{s_2}$ and $\sum_{s' \in \phi^{-1}_{model}(\bar{s})} P^a_{s_1 s'} = \sum_{s' \in \phi^{-1}_{model}(\bar{s})} P^a_{s_2 s'}$.

2. A *Q^π-irrelevance abstraction* ϕ_{Q^π} is such that for any policy π and any action a, $\phi_{Q^\pi}(s_1) = \phi_{Q^\pi}(s_2)$ implies $Q^\pi(s_1, a) = Q^\pi(s_2, a)$.

3. A *Q^*-irrelevance abstraction* ϕ_{Q^*} is such that for any action a, $\phi_{Q^*}(s_1) = \phi_{Q^*}(s_2)$ implies $Q^*(s_1, a) = Q^*(s_2, a)$.

4. An *a^*-irrelevance abstraction* ϕ_{a^*} is such that every abstract class has an action a^* that is optimal for all the states in that class, and $\phi_{a^*}(s_1) = \phi_{a^*}(s_2)$ implies that $Q^*(s_1, a^*) = max_a Q^*(s_1, a) = max_a Q^*(s_2, a) = Q^*(s_2, a^*)$.

5. A *π^*-irrelevance abstraction* ϕ_{π^*} is such that every abstract class has an action a^* that is optimal for all the states in that class, that is $\phi_{\pi^*}(s_1) = \phi_{\pi^*}(s_2)$ implies that $Q^*(s_1, a^*) = max_a Q^*(s_1, a)$ and $Q^*(s_2, a^*) = max_a Q^*(s_2, a)$.

In other words:

1. A model-irrelevance abstraction is such that two states are related only if the their reward and transition functions are identical.

2. A Q^π-irrelevance abstraction is such that two states are related only if their state-action value functions are identical under all possible policies.

3. A Q^*-irrelevance abstraction is such that two states are related only if their state-action value functions are identical under the optimal policy.

4. An a^*-irrelevance abstraction is such that all states of a given abstract class have the same optimal action a^* and, furthermore, two states are related only if a) a^* yields the maximum return from the state-action value function under the optimal policy while in either of the related states and b) the state-action value functions for the related states are identical under the optimal policy.

5. A π^*-irrelevance abstraction is such that all states of a given abstract class have the same optimal action a^* and, furthermore, two particular states are related only if a^* yields the maximum return from state-action value function under the optimal policy while in either of the related states.

4 Properties of Highlighted Abstractions

In order to gain some perspective on the relationships among these five abstractions, we define the following relation.

Definition 4.1. Given abstractions ϕ_1 and ϕ_2, we say ϕ_1 is *finer* than ϕ_2, denoted $\phi_1 \succeq \phi_2$, if for any states $s_1, s_2 \in S$, $\phi_1(s_1) = \phi_1(s_2)$ implies $\phi_2(s_1) = \phi_2(s_2)$.

Intuitively, one abstraction is finer than another if the former retains at least as much information about the ground state space as the latter. The following theorem follows readily from our definitions in 3.1. Note that ϕ_0 below refers to the identity abstraction (i.e. $\phi_0(s) = s$ for all $s \in S$).

Theorem 4.2. *For any MDP, we have* $\phi_0 \succeq \phi_{model} \succeq \phi_{Q^\pi} \succeq \phi_{Q^*} \succeq \phi_{a^*} \succeq \phi_{\pi^*}$.

The above theorem demonstrates the balance between the information retained by an abstraction and the size of the abstract state space. The finest abstraction ϕ_0 retains all information about the ground state space, but offers no reduction in size. At the other extreme, a π^*-irrelevance abstraction may offer great reduction in the size of the state space, but at the cost of much information loss. Precisely what balance is required for a given problem is a matter of empirical investigation.

Given an optimal policy for a state-abstracted MDP, a natural question is to ask how such a policy will perform on the ground MDP. The following theorem summarizes the performance of the abstractions in 3.1.

Theorem 4.3. *With abstractions ϕ_{model}, ϕ_{Q^π}, ϕ_{Q^*}, and ϕ_{a^*}, the optimal policy $\bar{\pi}^*$ is optimal in the ground MDP. However, there exist examples where the optimal policy with abstraction ϕ_{π^*} is suboptimal in the ground MDP.*

Of the abstractions discussed, only π^*-irrelevance fails to preserve policy optimality in all problems. In other words, the information loss resulting from applying a π^*-irrelevance abstraction can be so great that optimality is lost in the ground MDP.

5 Case Study: TAXI Problem

We conclude with some empirical results regarding the size of abstract state spaces for the TAXI problem. In this problem, an agent is to navigate a taxi in a 5x5 grid world, pick up a passenger, and deliver him/her to the destination. The ground state space contains 500 states and 6 actions. Each action performed costs -1, and the agent receives a reward of +20 after successfully delivering the passenger to the destination.

Abstraction	ϕ_0	ϕ_{model}	ϕ_{Q^*}	ϕ_{a^*}	ϕ_{π^*}
# of States	500	500	489	381	6

It should be noted that finding Q^π-irrelevance abstractions efficiently is still an open problem, but we know from 4.2 that the size of the abstract state space will be between that of ϕ_{model} and ϕ_{Q^*}.

References

[1] Lihong Li, Thomas J. Walsh, and Michael L. Littman. Toward a Unified Theory of Abstraction for MDPs. *Artificial Intelligence and Mathematics*, 2006.