

Probabilistic Aspects of Computer Science: TD6

Reachability Objectives in MDP & Probabilistic Automata

Emile Contal

<http://econtal.perso.math.cnrs.fr/teaching>

October 21, 2015

We are interested here in computing the minimum and maximum probabilities to reach a subset of states of a given MDP, and in describing policies achieving these optima.

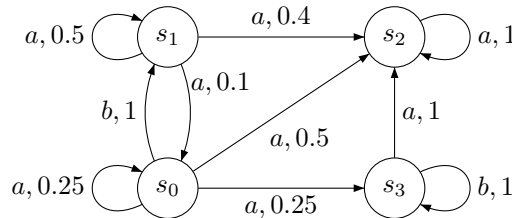
In the following, we consider an MDP $\mathcal{M} = (S, \{A_s\}_{s \in S}, p)$ with no reward functions. We recall that a *history* σ in \mathcal{M} is an infinite sequence $(s_0, a_0, s_1, a_1, s_2, \dots)$, such that for all $i \geq 0$, $s_i \in S$, $a_i \in A_{s_i}$ and $p(s_{i+1} | s_i, a_i) > 0$. Given an initial state $s \in S$, we denote $\text{Hist}(s)$ the set of histories starting in s . Then a policy $\pi \in \Pi^{HR}$ (history-dependent and randomized) permits to define discrete-time Markov chain \mathcal{M}^π with set of states being the finite prefixes of histories in $\text{Hist}(s)$.

Given a target subset T of S , we denote as $\text{Hist}(s, T)$ the set of histories, starting from state s , and reaching at some moment a state of T , i.e., such that there exists n with $s_n \in T$. As the reachability property only depends on a finite prefix of the history, $\text{Hist}(s, T)$ is a measurable subset of $\text{Hist}(s)$, hence, the DTMC \mathcal{M}^π defines its probability, denoted $\mathbf{Pr}^\pi(s, T)$. In the following, we study the two quantities

$$p_{\min}(s, T) = \inf_{\pi \in \Pi^{HR}} \mathbf{Pr}^\pi(s, T) \quad \text{and} \quad p_{\max}(s, T) = \sup_{\pi \in \Pi^{HR}} \mathbf{Pr}^\pi(s, T).$$

We will suppose in the following that states of T are absorbing, i.e., for all $t \in T$, $A_t = \{\alpha_t\}$ with $p(t | t, \alpha_t) = 1$.

You are invited to use the example depicted below throughout the rest, with $s = s_0$ and $T = \{s_2\}$.



Exercise 1 (Qualitative analysis). We start by considering the problem of determining states s for which $p_{\min}(s, T)$ or $p_{\max}(s, T)$ is zero or one: we denote these four possible sets of states as $S_T^{\min=0}$, $S_T^{\min=1}$, $S_T^{\max=0}$ and $S_T^{\max=1}$.

1. Find an iterative algorithm to compute the four sets.
2. What is the complexity of your algorithms?

Exercise 2 (Stationary deterministic policies are enough). We consider known, thanks to the previous exercise, the sets $S_T^{\min=0}$ and $S_T^{\min=1}$, and denote as $S^?$ the set $S \setminus (S_T^{\min=0} \cup S_T^{\min=1})$. We define (E) as the equation of the variable $\mathbf{x} \in \mathbb{R}^S$:

$$x_s = \begin{cases} 1 & \text{if } s \in S_T^{\min=1} \\ 0 & \text{if } s \in S_T^{\min=0} \\ \min_{a \in A_s} \sum_{s' \in S} p(s' | s, a) x_{s'} & \text{if } s \in S^?. \end{cases}$$

1. Show that the vector $(p_{\min}(s, T))_{s \in S}$ is a solution of (E) .

2. Consider a stationary deterministic policy d^∞ . Find a simple equation (E') having as unique solution the vector $(\mathbf{Pr}^{d^\infty}(s, T))_{s \in S^?}$. (Hint: to prove uniqueness, you may search for the classification of the states of the underlying DTMC with set of states S .)
3. Prove that (E) has then a unique solution and that $p_{\min}(s, T)$ is indeed a minimum computable by $\min_{\pi \in \Pi^{SD}} \mathbf{Pr}^\pi(s, T)$ (since Π^{SD} is a finite set). This shows the existence of an optimal strategy.

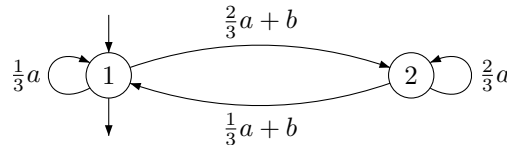
Exercise 3 (Computing $p_{\min}(s, T)$ and an optimal policy). We will study the three principal methods enabling the computation of $p_{\min}(s, T)$ and an optimal policy: value iteration, linear programming and policy iteration.

1. Write a value iteration algorithm to estimate the probability $p_{\min}(s, T)$ and an associated almost optimal policy. (Recall: value iteration is based on the iteration of the operator F suggested in equation (E) of the previous exercise, such that $p_{\min}(s, T)$ is its unique fixed point.)
2. Show that the vector $(p_{\min}(s, T))_{s \in S}$ is the unique solution of the following linear programming:

$$\text{Maximize } \sum_{s \in S} x_s \text{ subject to } \begin{cases} \forall s \in S_T^{\min=1} & x_s = 1 \\ \forall s \in S_T^{\min=0} & x_s = 0 \\ \forall s \in S^? \forall a \in A_s & x_s \leq \sum_{s' \in S} p(s' | s, a) x_{s'} \end{cases}$$

3. Write a policy iteration algorithm to find the exact value of $p_{\min}(s, T)$ and an associated optimal policy.

Exercise 4. We consider the probabilistic automaton \mathcal{A} over alphabet $A = \{a, b\}$ depicted below:



1. Describe the probability associated with word ab^n for every natural number n .
2. What are the possible probabilities that can be associated to a word by \mathcal{A} ?
3. Determine the (finite) class $\mathcal{L} = \{L_{>\theta}(\mathcal{A}) \mid \theta \in [0, 1]\}$.

Exercise 5. Prove or disprove that the following languages are stochastic languages.

1. $\{w \in \{a, b\}^* \mid |w|_a = |w|_b\}$, $\{w \in \{a, b, c\}^* \mid |w|_a = |w|_b = |w|_c\}$ and $\{a^n b^n c^n \mid n > 0\}$
2. $\{a^n b^m \mid n \neq m\}$
3. $\{a^m b a^{m_1} b a^{m_2} b \dots a^{m_k} b a^* \mid \exists j \leq k \quad m_1 + m_2 + \dots + m_j = m\}$
4. language of palindromes over $\{a, b\}$, i.e., words w such that $\bar{w} = w$
5. $\{a^n b^m a^n b^m \mid n, m \geq 1\}$