The Relationship between MDPSTs and MDPIPs

Jianmin Ji
Multi-Agent Systems Lab, Department of Computer Science
University of Science and Technology of China
Hefei, 230026, China
jizheng@mail.ustc.edu.cn

December 10, 2007

1 Preparation

Markov decision processes with imprecise probabilities (MDPIPs) [3] and *Markov decision processes with set-valued transitions* (MDPSTs) [2] are two import frameworks for imprecise MDPs.

An MDPIP is a Markov decision process where transitions are specified through sets of probability measures. An MDPIP is defined by a tuple $\langle \mathcal{S}, \mathcal{A}, \mathcal{K}, \mathcal{C} \rangle$, where

- $\mathcal{S}$ is a finite set of states of the system;
- $\mathcal{A} : \mathcal{S} \to 2^{\mathcal{S}}$ is the possible action function, where $\mathcal{A}(s)$ is a set of possible actions for the state s;
- $\mathcal{K}$ is a credal set over the state space, a nonempty credal set $\mathcal{K}_s(a)$ for all $s \in \mathcal{S}$ and $a \in \mathcal{A}(s)$, representing the set of probability distributions $P(s' \mid s, a)$ over successor states in $\mathcal{S}$;
- $\mathcal{C} : \mathcal{S} \times A \to \mathbb{R}$ is a real-valued, bounded reward function.

An MDPST is a Markov decision process where transitions move probabilistically to reachable sets, and the probability for a particular state is not resolved by the model. An MDPST is defined by a tuple $\langle \mathcal{S}, \mathcal{A}, m, \mathcal{C} \rangle$, where

- $\mathcal{S}$, $\mathcal{A}$, $\mathcal{C}$ are the same as those defined in MDPIPs;

- For any $s \in \mathcal{S}$, $a \in \mathcal{A}(s)$, and $k \in 2^{\mathcal{S}} \setminus \emptyset$, $m(k \mid s, a)$ stands for the probability of ending in the set of states k. $\mathbf{F}(s,a) = \{k \mid m(k \mid s,a) > 0\}$.

The relationship between MDPIPs and MDPSTs is discussed in the next section.

2 The main theorem

In this section we prove that any MDPST is expressible by a MDPIP. The proof is based on the Farkas Lemma [1]. We use A^T to denote the transpose of the matrix A.

Lemma 1 (Farkas Lemma)

$$(\exists x \geq 0, Ax = b) \Leftrightarrow (\forall y, (A^T y \geq 0 \Rightarrow b^T y \geq 0)).$$

With the help of the Farkas Lemma, we have the following fundamental relationship between MDPSTs and MDPIPs.

Theorem 1 *Any MDPST $q = \langle \mathcal{S}, \mathcal{A}, m, \mathcal{C} \rangle$ is expressible by an MDPIP $r = \langle \mathcal{S}, \mathcal{A}, \mathcal{K}, \mathcal{C} \rangle$, where for each $s \in \mathcal{S}$ and $a \in \mathcal{A}(s)$, any possible probability distribution $P(\cdot \mid s,a) \in \mathcal{K}_s(a)$ should satisfy the following condition: for any $k \subseteq \mathcal{S}$,*

$$\sum_{k' \in \mathbf{F}(s,a),\ s.t.\ k' \subseteq k} m(k' \mid s,a) \leq \sum_{s' \in k} P(s' \mid s,a). \tag{1}$$

Proof. For any $s \in \mathcal{S}$ and $a \in \mathcal{A}(s)$, a probability distribution $P(\cdot \mid s,a)$ is allowed by a MDPST $\langle \mathcal{S}, \mathcal{A}, m, \mathcal{C} \rangle$ iff the following linear equations have a non-negative solution.

$$\begin{cases} \sum_{k \in F(s,a)} m(k \mid s,a) W(k, s_1) = P(s_1 \mid s,a), \\ \sum_{k \in F(s,a)} m(k \mid s,a) W(k, s_2) = P(s_2 \mid s,a), \\ \cdots \\ \sum_{k \in F(s,a)} m(k \mid s,a) W(k, s_n) = P(s_n \mid s,a), \\ \sum_{s' \in k_1} W(k_1, s') = 1, \\ \sum_{s' \in k_2} W(k_2, s') = 1, \\ \cdots \\ \sum_{s' \in k_l} W(k_l, s') = 1. \end{cases}, \tag{2}$$

where $n = |\mathcal{S}|$ and $l = |\mathbf{F}(s,a)|$. For each $k \in \mathbf{F}(s,a)$ and $s' \in \mathcal{S}$, $W(k,s')$ is the variable. Intuitively, $W(k,s')$ is the weight of s' in k. Clearly, there are $|\mathbf{F}(s,a)| \times |\mathcal{S}|$ variables and $|\mathbf{F}(s,a)| + |\mathcal{S}|$ equations.

From the Farkas Lemma, it is equal to prove that, for all possible y, $A^T y \geq 0$ implies $b^T y \geq 0$. Let $y = \{eq^{s_1}, \ldots, eq^{s_n}, eq^{k_1}, \ldots, eq^{k_l}\}^T$ and $A^T y \geq 0$ then for each $k \in \mathbf{F}(s,a)$, if $s' \in k$ then $eq^{s'} m(k \mid s,a) + eq^k \geq 0$, if $s' \notin k$ then $eq^{s'} m(k \mid s,a) \geq 0$. We need to prove that the inequality $P(s_1 \mid s,a) eq^{s_1} + \cdots + P(s_n \mid s,a) eq^{s_n} + eq^{k_1} + \cdots + eq^{k_l} \geq 0$ is valid. It is equal to prove that the following inequality is valid under the condition (1):

$$\begin{aligned} P(s_1 \mid s,a)\ eq^{s_1} + \cdots + P(s_n \mid s,a)\ eq^{s_n} - \min\{eq^{s'} \mid s' \in k_1\}\ m(k_1 \mid s,a) \\ - \cdots - \min\{eq^{s'} \mid s' \in k_l\}\ m(k_l \mid s,a) \geq 0. \end{aligned} \tag{3}$$

It is clear that if there exists some $k \in \mathbf{F}(s,a)$ s.t. $s' \notin k$, then $eq^{s'} \geq 0$. Let $t = \{s' \mid \forall k \in \mathbf{F}(s,a), s' \in k \text{ and } eq^{s'} \leq 0\}$. If $t \neq \emptyset$ then inequality (3) is valid if the following inequality is valid

$$\sum_{s' \in t} P(s' \mid s,a)\ eq^{s'} + \sum_{s' \notin t} P(s' \mid s,a)\ eq^{s'} - \min\{eq^{s'} \mid s' \in t\} \geq 0. \tag{4}$$

Clearly, inequality(4) is valid, now we only need to consider $t = \emptyset$.

If $t = \emptyset$ then for each $s' \in \mathcal{S}$, $eq^{s'} \geq 0$. Let $c_0 = \emptyset$, $\mathcal{S}_i = \mathcal{S} \setminus c_{i-1}$, and $c_i = c_{i-1} \cup \{s' \mid eq^{s'} = \max\{eq^{s''} \mid s'' \in \mathcal{S}_i\}\}$, $0 < i < \infty$. It is clear that $\mathcal{S} = \bigcup_{0 \leq i < \infty} c_i$ and from the condition (1) the following inequality is valid

$$\sum_{s' \in c_i} P(s' \mid s,a) - \sum_{k \in \mathbf{F}(s,a),\ s.t.\ k \subseteq c_i} m(k \mid s,a) \geq 0, \tag{5}$$

where $0 < i < \infty$.

At last, let $EQ(c_i) = \min\{eq^{s'} \mid s' \in c_i\}$, then for each $k \in \mathbf{F}(s,a)$, $\min\{eq^{s'} \mid s' \in k\} = EQ(c_j)$ and $j = \min\{i \mid k \subseteq c_i\}$. So

$$\begin{aligned} P(s_1 \mid s,a)\ eq^{s_1} + \cdots + P(s_n \mid s,a)\ eq^{s_n} - \min\{eq^{s'} \mid s' \in k_1\}\ m(k_1 \mid s,a) \\ - \cdots - \min\{eq^{s'} \mid s' \in k_l\}\ m(k_l \mid s,a) \quad = \\ \sum_{0<i<\infty} [EQ(c_i) - EQ(c_{i+1})] \cdot [\sum_{s' \in c_i} P(s' \mid s,a) - \sum_{k \in \mathbf{F}(s,a),\ s.t.\ k \subseteq c_i} m(k \mid s,a)] \\ \geq 0. \end{aligned} \tag{6}$$

So under the condition (1), inequality (3) is valid and the probability distribution $P(\cdot \mid s,a)$ is allowed. ■

In fact the condition (1) is also a necessary condition.

Proposition 1 *For each $s \in \mathcal{S}$ and $a \in \mathcal{A}(s)$, a probability distribution $P(\cdot \mid s,a)$ is allowed by a MDPST $\langle \mathcal{S}, \mathcal{A}, m, \mathcal{C} \rangle$, then it satisfies the following condition: for any $k \subseteq \mathcal{S}$,*

$$\sum_{k' \in \mathbf{F}(s,a),\ s.t.\ k' \subseteq k} m(k' \mid s,a) \leq \sum_{s' \in k} P(s' \mid s,a). \tag{7}$$

Proof. Clearly, in order to have non-negative solutions for linear equations (2), the condition should be satisfied. ■

There is another condition which is equal to the condition (1).

Proposition 2 *For any MDPST $\langle \mathcal{S}, \mathcal{A}, m, \mathcal{C} \rangle$, $s \in \mathcal{S}$ and $a \in \mathcal{A}(s)$, $P(\cdot \mid s, a)$ is a probability distribution allowed by the MDPST.*

$$\text{For all } k \subseteq \mathcal{S},\ \sum_{s' \in k} P(s' \mid s, a) \leq \sum_{k' \in \mathbf{F}(s,a),\ s.t.\ k \cap k' \neq \emptyset} m(k' \mid s, a) \tag{8}$$

iff

$$\text{for all } k \subseteq \mathcal{S},\ \sum_{k' \in \mathbf{F}(s,a),\ s.t.\ k' \subseteq k} m(k' \mid s, a) \leq \sum_{s' \in k} P(s' \mid s, a). \tag{9}$$

Proof. Assume inequality (8) is true, then

$$\text{for all } k \subseteq \mathcal{S},\ 1 - \sum_{s' \in k} P(s' \mid s, a) \geq 1 - \sum_{k' \in \mathbf{F}(s,a),\ s.t.\ k \cap k' \neq \emptyset} m(k' \mid s, a). \tag{10}$$

Let $\bar{k} = \mathcal{S} \setminus k$, then

$$1 - \sum_{s' \in k} P(s' \mid s, a) = \sum_{s' \in \bar{k}} P(s' \mid s, a)$$

and

$$1 - \sum_{k' \in \mathbf{F}(s,a),\ s.t.\ k \cap k' \neq \emptyset} m(k' \mid s, a) = \sum_{k' \in \mathbf{F}(s,a),\ s.t.\ k' \subseteq \bar{k}} m(k' \mid s, a).$$

So from (10) we get

$$\text{for all } \bar{k} \subseteq \mathcal{S},\ \sum_{s' \in \bar{k}} P(s' \mid s, a) \geq \sum_{k' \in \mathbf{F}(s,a),\ s.t.\ k' \subseteq \bar{k}} m(k' \mid s, a). \tag{11}$$

So we get (9) from (8), and (8) can be driven from (9) similarly and dually. ■

References

[1] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

[2] F. W. Trevizan, F. G. Cozman, and L. N. de Barros. Planning under risk and knightian uncertainty. In *Proc. of the 20th IJCAI*, pages 2023–2028, 2007.

[3] C. C. White III and H. K. Eldeib. Markov Decision Processes with Imprecise Transition Probabilities. *Operations Research*, 42(4):739–749, 1994.