广义和游戏中的敌友 Q 学习

文章·2003 年 1 月
Source: CiteSeer

Michael L. Littman
Brown University

$$R_i(\pi_1, \ldots, \pi_n) \leq R_i(\pi_1, \ldots, \pi_{i-1}, \pi_i', \pi_{i+1}, \ldots, \pi_n), \quad (1)$$

fi

$$R_i(\pi_1,\ldots,\pi_n) \leq R_i(\pi_1',\ldots,\pi_{i-1}',\pi_i,\pi_{i+1}',\ldots,\pi_n'), \tag{2}$$

$\pi\pi = -$

$$R_i(\pi_1,\ldots,\pi_n) = \max_{a_1\in A_1,\ldots,a_n\in A_n} R_i(a_1,\ldots,a_n) \tag{3}$$

$= = \cdots =$

$$\text{row:}R_1 = \begin{bmatrix} -1 & 2 \\ 0 & 1 \end{bmatrix}, \text{column:}R_2 = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}. \tag{4}$$

$\pi = ()\pi = ()$ fi− fi− fififi

$b = ()b = ()$ fi

fi==

$\leq \gamma <$ fi$\pi\pi$ fi

$$Q_i(s,a_1,\ldots,a_n) = R_i(s,a_1,\ldots,a_n) + \gamma\sum_{s'\in S} T(s,a_1,\ldots,a_n,s')Q_i(s',\pi_1,\ldots,\pi_n), \tag{5}$$

$\pi\pi\pi$

fi

fi

$$Q_i[s,a_1,\ldots,a_n] := (1-\alpha_t)Q_i[s,a_1,\ldots,a_n] + \alpha_t\left(r_i + \gamma \mathrm{Nash}_i(s,Q_1,\ldots,Q_n)\right) \tag{6}$$

$=\pi\pi\pi\ \alpha$

$$
\begin{aligned}
& R_i(\pi_1, \ldots, \pi_n) \\
\leq \quad & R_i(\pi_1, \ldots, \pi_{i-1}, \pi_i', \pi_{i+1}, \ldots \pi_n) \\
= \quad & R_i(\rho_1', \ldots, \rho_{i-1}', \rho_i, \rho_{i+1}', \ldots, \rho_n') \\
\leq \quad & R_i(\rho_1, \ldots, \rho_n).
\end{aligned}
$$

$$\mathrm{Nash}_1(s, Q_1, Q_2) = \max_{a_1 \in A_1, a_2 \in A_2} Q_1[s, a_1, a_2] \tag{7}$$

$$\mathrm{Nash}_1(s, Q_1, Q_2) = \max_{\pi \in \Pi(A_1)} \min_{a_2 \in A_2} \sum_{a_1 \in A_1} \pi(a_1) Q_1[s, a_1, a_2] \tag{8}$$

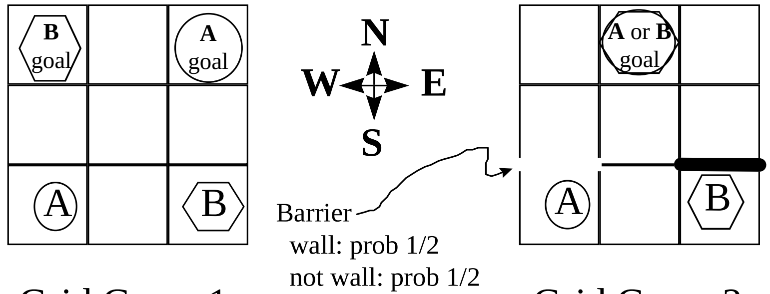
$$\mathrm{Nash}_i(s, Q_1, \ldots, Q_n) = \max_{\pi \in \Pi(X_1 \times \cdots \times X_k)} \min_{y_1, \ldots, y_l \in Y_1 \times \cdots \times Y_l} \sum_{x_1, \ldots, x_k \in X_1 \times \cdots \times X_k} \pi(x_1) \cdots \pi(x_k) Q_i[s, x_1, \ldots, x_k, y_1, \ldots, y_l].$$

$$|\mathrm{Nash}_1(s, Q_1, \ldots) - \mathrm{Nash}_1(s, Q_1', \ldots)| \leq \max_{a_1, a_2} |Q_1(s, a_1, a_2) - Q_1'(s, a_1, a_2)|.$$

$$\begin{aligned} R_1(\pi_1, \ldots, \pi_n) &\geq R_1(\pi_1', \pi_2, \ldots, \pi_n) \\ &= R_1(\rho_1, \rho_2', \ldots, \rho_n') \\ &\geq R_1(\rho_1, \ldots, \rho_n). \end{aligned}$$

$$\begin{aligned} R_1(\rho_1, \ldots, \rho_n) &\geq R_1(\rho_1', \rho_2, \ldots, \rho_n) \\ &= R_1(\pi_1, \pi_2', \ldots, \pi_n') \\ &\geq R_1(\pi_1, \ldots, \pi_n). \end{aligned}$$

$$
R_1 = \begin{array}{c c}
 & \begin{array}{cc} \mathbf{C} & \mathbf{B} \end{array} \\
\begin{array}{c} \mathbf{C} \\ \mathbf{B} \end{array} & \begin{bmatrix} -100 & +100 \\ +50 & +66 \end{bmatrix}
\end{array}
$$



Grid Game 1

Grid Game 2