

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/2933305>

广义和游戏中的敌友 Q 学习

文章 · 2003 年 1 月

Source: CiteSeer

CITATIONS

463

READS

10,122

1 author:



Michael L. Littman
Brown University
301 PUBLICATIONS 37,807 CITATIONS

[SEE PROFILE](#)

¶

¶¶

¶¶¶¶¶

¶¶¶

$$B^i(u_1, \dots, u_m) \geq B^i(u_1, \dots, u_{i-1}, u_i^{-1}, u_{i+1}, \dots, u_m), \quad (I)$$

¶

$$\leq \lambda < \# \# \#$$

$$B^i(u_1, \dots, u_n) \leq B^i(u_1, \dots, u_{i-1}, u_i, u_{i+1}, \dots, u_n) \quad (5)$$

u = -

$$\begin{aligned} O^i(s^i \alpha_1 \cdots \alpha_n) &= B^i(s^i \alpha_1, \dots, \alpha_n) \\ &+ \lambda \sum_{s^j \in S} L(s^j \alpha_1, \dots, \alpha_n, s^j) O^i(s^j u_1, \dots, u_n) \end{aligned} \quad (2)$$

u u u =

$$B^i(u_1, \dots, u_n) = \max_{\alpha_1 \in V_1, \dots, \alpha_n \in V_n} B^i(\alpha_1, \dots, \alpha_n) \quad (3)$$

= \dots =

\# \# \#

$$\begin{bmatrix} & & \\ & & \\ & & \end{bmatrix} \quad \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix}$$

row: $B_I = \begin{bmatrix} -I & J \\ 0 & I \end{bmatrix}$, column: $B_J = \begin{bmatrix} I & -J \\ 0 & -I \end{bmatrix}$ (4)

$$u = ()u = () \# - \# \# \#$$

\# \# \#

$$b = ()b = () \#$$

$$\begin{aligned} O^i[s^i \alpha_1, \dots, \alpha_n] &:= (I - \alpha_i) O^i[s^i \alpha_1, \dots, \alpha_n] + \\ &\alpha_i (u_i + \lambda \# \# \# (s^i O^i[s^i \alpha_1, \dots, \alpha_n])) \end{aligned} \quad (6)$$

= u u u \# \# \#

$\underline{u} \underline{u} b \underline{b} \underline{H} \underline{H} \underline{b}$

$$\begin{aligned}& L^{\frac{1}{2}}(\underline{u}_I, \dots, \underline{u}_M) \\&\geq L^{\frac{1}{2}}(\underline{u}_I, \dots, \underline{u}_{i-1}, \underline{u}_i^1, \underline{u}_{i+1}, \dots, \underline{u}_M) \\&= L^{\frac{1}{2}}(b_I, \dots, b_{i-1}^1, b_i, b_{i+1}^1, \dots, b_M^1) \\&> L^{\frac{1}{2}}(b_I, \dots, b_M).\end{aligned}$$

$\underline{u} \underline{b} \underline{u} \underline{b}$

$$M\sigma\mu_I(z, \mathcal{O}_I, \mathcal{O}_S) = \sum_{\alpha_I \in \mathbb{V}_I, \alpha_S \in \mathbb{V}_S} \mathcal{O}_I[z, \alpha_I, \alpha_S] \quad (\Delta)$$

$$\begin{aligned} M\sigma\mu_I(z, \mathcal{O}_I, \mathcal{O}_S) &= \sum_{\underline{\alpha} \in \Pi(\mathbb{V}_I)} \sum_{\alpha_S \in \mathbb{V}_S} \sum_{\alpha_I \in \mathbb{V}_I} \underline{\alpha}(\alpha_I) \mathcal{O}_I[z, \alpha_I, \alpha_S] \\ &\quad (8) \end{aligned}$$

$$\begin{aligned} M\sigma\mu_I(z, \mathcal{O}_I, \dots, \mathcal{O}_v) &= \sum_{\substack{\underline{\alpha} \in \Pi(\mathbb{V}_I \times \dots \times \mathbb{V}_v) \\ \lambda_I, \dots, \lambda_v \in \Lambda_I \times \dots \times \Lambda_v \\ \alpha_I, \dots, \alpha_v \in \mathbb{V}_I \times \dots \times \mathbb{V}_v}} \\ &\quad \underline{\alpha}(x_I) \dots \underline{\alpha}(x_v) \mathcal{O}_I[z, x_I, \dots, x_v, \lambda_I, \dots, \lambda_v]. \\ &\quad \mathcal{B}_I(\underline{\alpha}_I, \dots, \underline{\alpha}_v) \geq \mathcal{B}_I(\underline{\alpha}'_I, \underline{\alpha}'_S, \dots, \underline{\alpha}'_v) \\ &= \mathcal{B}_I(b_I, b'_S, \dots, b'_v) \\ &\geq \mathcal{B}_I(b_I, \dots, b_v). \end{aligned}$$

$$\begin{aligned} &\mathcal{B}_I(b_I, \dots, b_v) \geq \mathcal{B}_I(b'_I, b'_S, \dots, b'_v) \\ &= \mathcal{B}_I(\underline{\alpha}_I, \underline{\alpha}'_S, \dots, \underline{\alpha}'_v) \\ &\geq \mathcal{B}_I(\underline{\alpha}_I, \dots, \underline{\alpha}_v). \end{aligned}$$

$$|M\sigma\mu_I(z, \mathcal{O}_I, \dots) - M\sigma\mu_I(z, \mathcal{O}'_I, \dots)| \leq$$

$$\sum_{\alpha_I, \alpha_S} |\mathcal{O}_I(z, \alpha_I, \alpha_S) - \mathcal{O}'_I(z, \alpha_I, \alpha_S)|.$$

¶

—

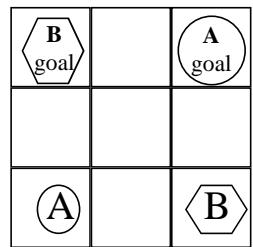
—

—

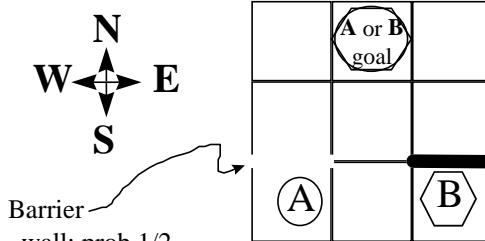
—

$$V^T = \begin{bmatrix} & & \\ C & B \\ \hline B & -100 & +100 \\ & +20 & +80 \end{bmatrix}$$

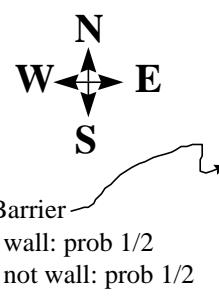
=



Grid Game 1



Grid Game 2



wall: prob 1/2

not wall: prob 1/2

¶ +

¶

¶ ª ª ª ——

¶

¶

+¶

