A strongly polynomial algorithm for payoff games

Bruno Scherrer

March 25, 2022

Abstract

...

Let

$$T_{\vec{\mu},\vec{\nu}}v = T^n$$

Lemma 1. *For any policy μ, for any x, there exists $0 \leq i < i+c \leq n$,*

$$v_*(x) \leq \mathbb{1}_x \gamma^i P_{\mu_*^i,\vec{\nu}_1^i}(I - \gamma^c P_{\mu_*^c,\vec{\nu}_{i+1}^{i+c}})^{-1}(T^{n-i}0 - T^{n-i-c}0).$$

Proof. μ_* play against $\vec{\nu}$. Once a cycle is found, one loops. □

Lemma 2. *For any policy ν, for any x, there exists $0 \leq i < i+c \leq n$,*

$$v_{\vec{\mu}_1^i(\vec{\mu}_{i+1}^{i+c})^\infty,\nu^\infty}(x) \geq \mathbb{1}_x \gamma^i P(\vec{\mu}_1^i,\nu)(I - \gamma^c P_{\vec{\mu}_{i+1}^{i+c},\nu^c})^{-1}(T^{n-i}0 - T^{n-i-c}0).$$

Proof. $\vec{\mu}$ plays against ν. Once a cycle is found, one loops. □

Algorithm: compute a policy that is better than $\vec{\mu}_1^i(\vec{\mu}_{i+1}^{i+c})^\infty$ for all i and c.