A strongly polynomial algorithm for payoff games

Bruno Scherrer

March 25, 2022

Abstract

...

Solve the n-horizon problem

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

Lemma 1. *For any policy μ, for any x, there exist y, $v \leq i < i+c \leq n$, such that*

$$v_\mu(x) - T^n v \leq \frac{\gamma^{i+c}}{1-\gamma^c} \mathbb{1}_y (T^{n-i} v - T^{n-i-c} v).$$

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

$$M_{x,y,i,c} = \{\mu \ ; \ \}$$

Lemma 2. *For any policy ν, for any x, there exist y, $v \leq i < i+c \leq n$, such that*

$$v_{\vec{\mu}_1^i (\vec{\mu}_{i+1}^{i+c})^\infty, \nu^\infty}(x) - T^n v \geq \frac{\gamma^i}{1-\gamma^c} \mathbb{1}_y (T^{n-i} v - T^{n-i-c} v).$$

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

Algorithm: compute a policy μ' that is better than $(\vec{\mu}_{i+1}^{i+c})^\infty$ for all i and c. From any x, μ' will end up cycling in y.

Warm-up: write the proof for the 1-player problem.