Value Iteration with Guessing

K. Chatterjee, M. JafariRaviz, R. Saona, J. Svoboda
Institute of Science and Technology Austria

Objective

Compute the value of Markov Decision Processes,
using only local updates,
for reachability (and shortest stochastic path).

$$v \in \mathbb{R}^S \mapsto \mathrm{Update}(v)(s) \coloneqq \max_{a \in A} \left\{ \sum_{s' \in S} p(s' \mid s, a)\ v(s') \right\}$$

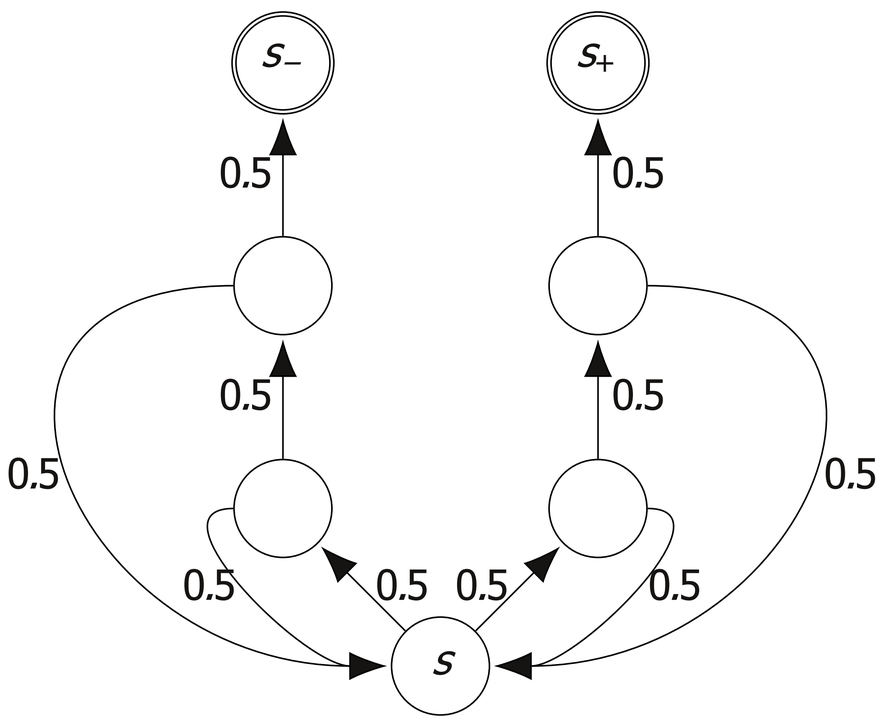
Value Iteration

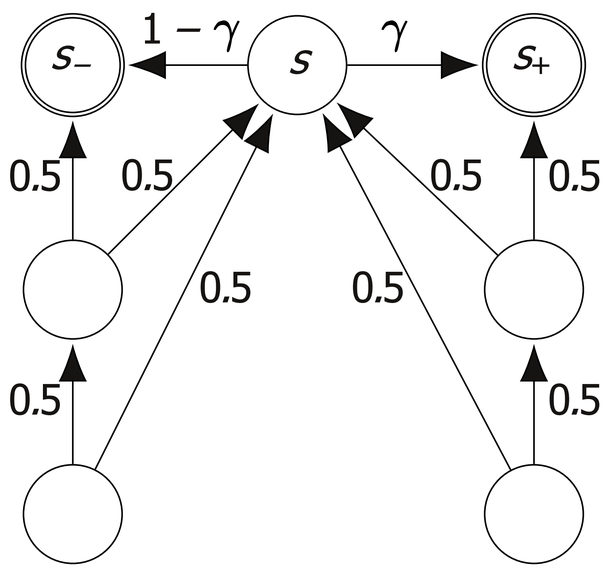
$$v_0(s) \coloneqq 1[s \text{ is target}]$$
$$v_{i+1} \coloneqq \mathrm{Update}(v_i)$$

$$\| v_{|S|i} - \mathrm{val} \|_\infty \leq \left(1 - p_{\min}^{|S|}\right)^i \| v_0 - \mathrm{val} \|_\infty$$

Number of updates required for approximation: $(-\log(\varepsilon)|S|/p_{\min})^{\Omega(|S|)}$

Guessing to simplify structure





Inequalities

Gueessed values can be verified with the recursive solution.

$$\mathrm{Update}(\gamma, \mathrm{val}[s \coloneqq \gamma])(s) > \gamma \ \Leftrightarrow\ \mathrm{val}(s) > \gamma$$

Approximate verification is possible.

$$\mathrm{Update}(\gamma, \mathrm{val}[s \coloneqq \gamma])(s) > \gamma - \varepsilon p_{\min}^{|S|} \ \Rightarrow\ \mathrm{val}(s) > \gamma - \varepsilon$$

Markov Chains

We can precompute which updates to ask for using linear space and almost linear time
to obtain an approximation of the value using few updates.

Precomputation

$$\mathcal{O}((|S| + |E|) \log |S|)$$

Numer of updates

$$(-\log(\varepsilon)|S|/p_{\min})^{\mathcal{O}(\sqrt{|S|})}$$

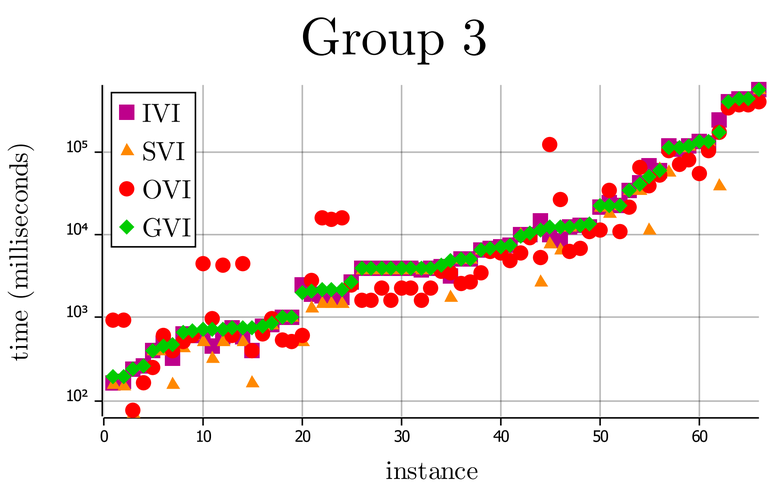
Markov Decision Processes

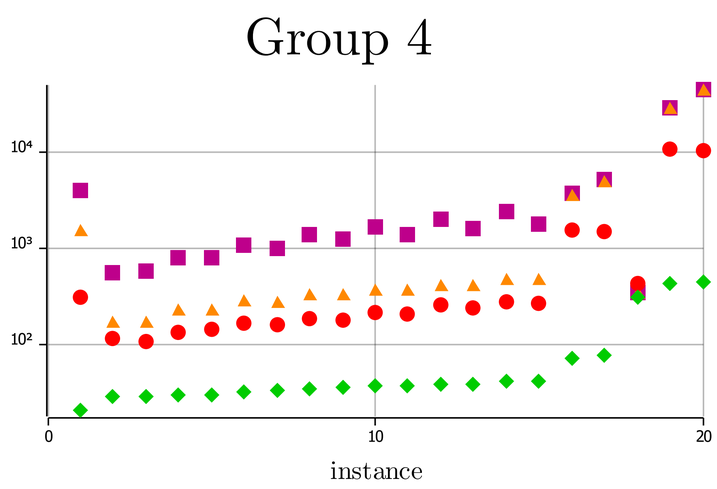
Group 1: 170 instances
All algorithms take less than 0.1 secs

Group 2: 135 instance
Fastest and slowest algorithms are only at most 1.10 times of each other

Group 3: 139 instances
Guessing VI is not the fastest

Group 4: 30 instances
The rest, our improvement





Acknowledgments: ERC CoG 863818