$$val_*(s) = \max_{a \in A} \sum_{\substack{s' \in S \\ r \in \mathbb{R}}} \Delta(s, a)(s', r) \left(r + \gamma \, val_*(s') \right)$$

(optimal) Bellman equations

Theorem: val_* is the unique solution to the following equations:

$$v(s) = \max_{a \in A} \sum_{\substack{s' \in S \\ r \in \mathbb{R}}} \Delta(s, a)(s', r) \left(r + \gamma \, v(s') \right)$$

All* algorithms are based on solving Bellman equations.

Value iteration

Fixed point computations

Tarski's theorem

$v_0 : S \to \mathbb{R}$ $\quad v_0(s) = 0$

$v : S \to \mathbb{R}$

For $k = 0$ to infinity:

$$v_{k+1}(s) = \max_{a \in A} \sum_{\substack{s' \in S \\ r \in \mathbb{R}}} \Delta(s, a)(s', r) \left(r + \gamma \, v_k(s') \right)$$

if for all $s \in S$ $|v_{k+1}(s) - v_k(s)| \leq \varepsilon$

return $v_{k+1} \approx_\varepsilon val_*$

(*) $v_k(s)$ is the optimal total reward we get in k steps.

$$v_k(s) = \max_{\sigma} \mathbb{E}_{\substack{\sigma, s \\ k \text{ steps}}} \left[r_0 + \gamma r_1 + \gamma^2 r_2 + \dots + \gamma^{k-1} r_{k-1} \right]$$

$$\sigma_*(s) = \arg\max_{a \in A}$$

$$\sum_{\substack{s' \in S \\ r \in \mathbb{R}}} \Delta(s, a)(s', r) \left(r + \gamma \, val_*(s') \right)$$

$q_*(s,a)$

$$\sigma_*(s) = \underset{a \in A}{\operatorname{argmax}}\ q_*(s,a)$$