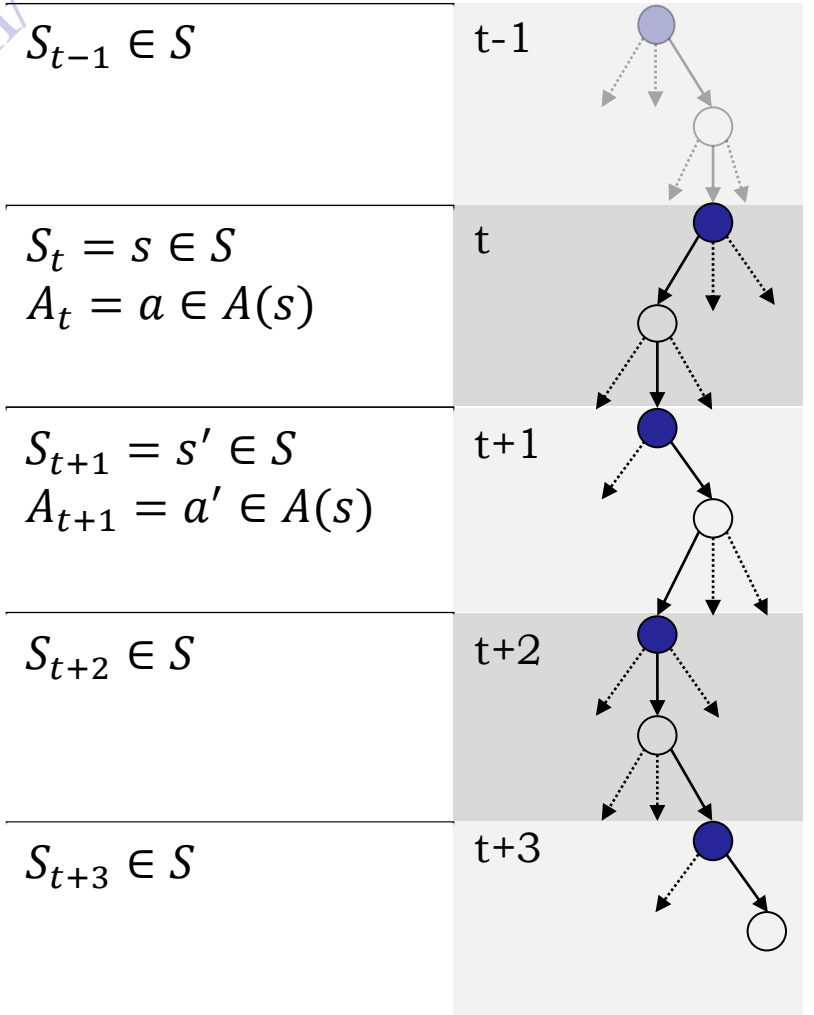
Value functions & Bellman equations

$State, s \in S$
$Action, a \in A(s)$
$Reward, r \in R(s, a, s')$
$Transition\ Probability, P(s, a, s')$
$Optimal\ Policy, \pi^*(s)$
$Return, G_t = \sum_{k=0}^{\infty} \gamma^k r_{t+k}$
$State\ value, V^\pi(s)$
$Action\ value, Q^\pi(s, a)$
$V^*(s) = \max_a \sum_{s'} P(s'|s, a)\big(r + \gamma V^*(s')\big)$
$Q^*(s, a) = \sum_{s'} p(s'|s, a)\left[r + \gamma \max_{a'} Q^*(s', a')\right]$

- Value of a state s under policy π
  - $V^\pi(s) = \mathbb{E}(G_t \mid S_t = s)$
  - $V^\pi(s) = \sum_a \pi(a|s) \sum_{s'} P(s'|s, a)\big(r + \gamma V^\pi(s')\big)$
- Value of a state s under optimal policy π^*
  - $V^*(s) = \max_a \sum_{s'} P(s'|s, a)\big(r + \gamma V^*(s')\big)$
- Value of taking action a in state s then following policy π
  - $Q^\pi(s) = \mathbb{E}(G_t \mid S_t = s, A_t = a)$
  - $Q^\pi(s, a) = \sum_{s'} P(s'|s, a)[r + \gamma V^\pi(s')]$
- Value of taking action a in state s then following optimal policy π^*
  - $Q^*(s, a) = \sum_{s'} p(s'|s, a)\left[r + \gamma \max_{a'} Q^*(s', a')\right]$

Standard Notations



http://www.profrkumar.co.in/