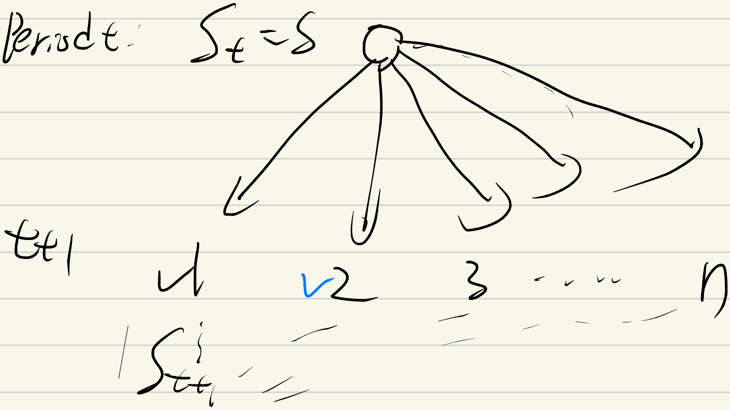
$S_t = s \in \{1, 2, \dots, n\}$ 26.01.27

Period t: $S_t = s$

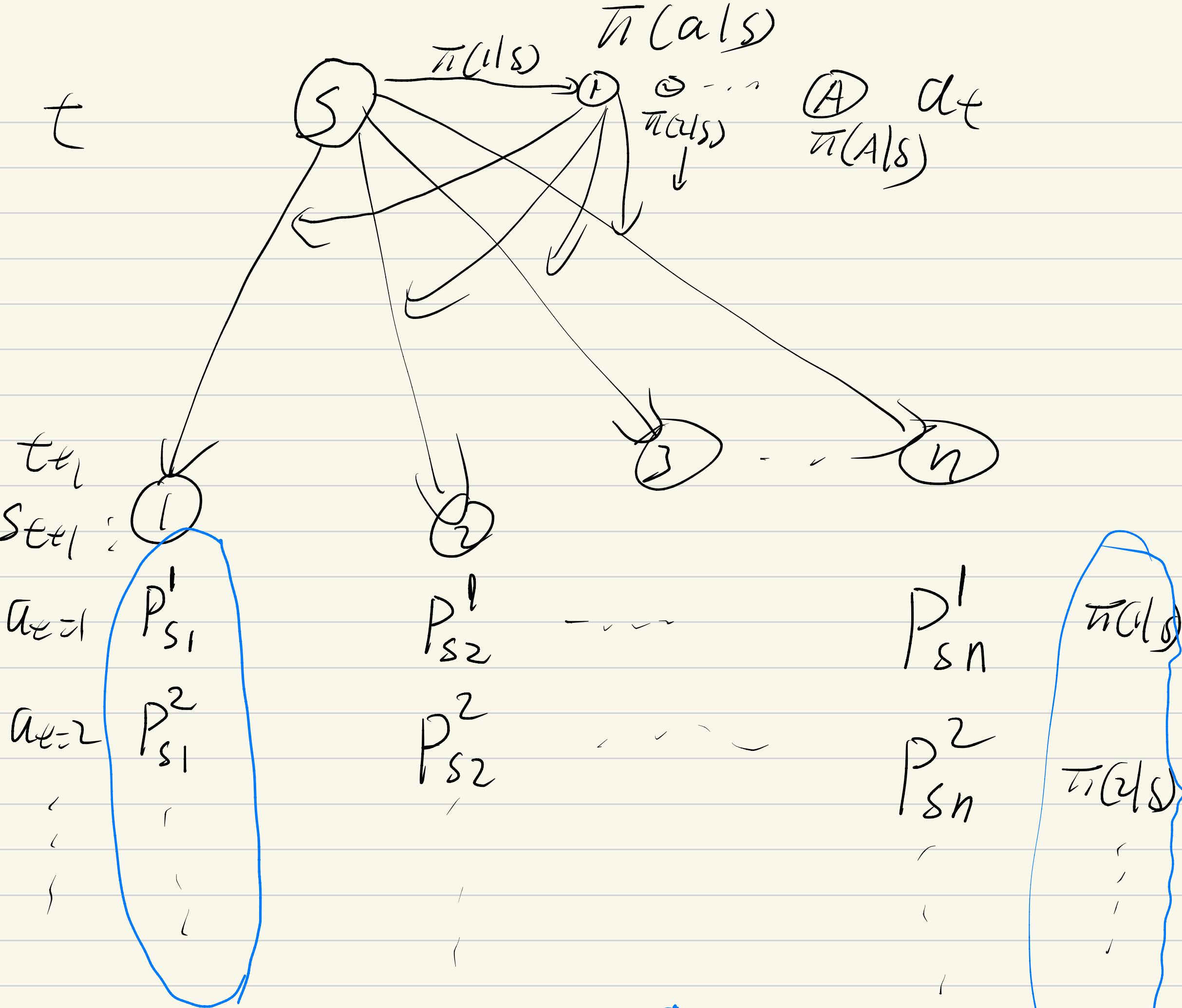


$$v(s) = P_{s1}\cdot\left(R_s + \gamma\cdot v(1)\right) + P_{s2}\cdot\left(R_s + \gamma v(2)\right) + \dots + P_{sn}\left(R_s + \gamma\cdot v(n)\right)$$

Bellman Equation

$$v(s) = \underbrace{\left(\sum_{i=1}^{n} P_{s_i}\right)}_{1}\cdot R_s + \sum_{i=1}^{n} P_{s_i}\cdot\gamma\cdot v(i)$$

$$= R_s + \sum_{i=1}^{n} P_{s,i}\cdot\gamma\cdot v(i)$$



$$P_{s_1}^{\pi} = \sum_{a=1}^{A} \pi(a|s) \cdot P_{s_1}^{a}$$

⋮

$$P_{ss'}^{\pi} = \sum_{a=1}^{A} \pi(a|s) P_{ss'}^{a}$$

$\vec{v}_k$ s.t. $\vec{v}_0 = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$, $\vec{v}_{k+1} = \text{Bellman Oper}^{\pi}(\vec{v}_k)$

$$v_\pi = \text{Bellman}^{\pi}(v_\pi)$$

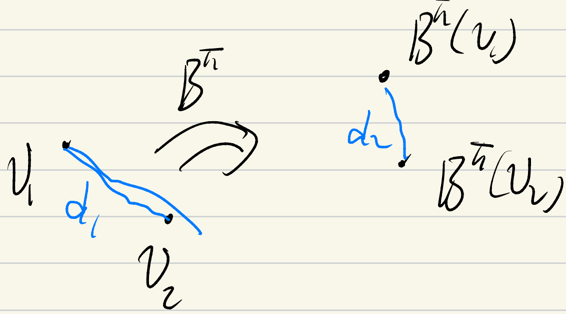
$\|\vec{v}_{k+1} - \vec{v}_k\|_\infty \le \varepsilon$

$\Downarrow$

Output $\vec{v}_k$

Bellman^{π} is a contraction mapping in L^∞-norm

$$\|\text{Bellman}^{\pi}(v_1) - B^{\pi}(v_2)\|_\infty \le \gamma \|v_1 - v_2\|_\infty$$



$d_2 \le \gamma \cdot d_1$