$$|S| = n \cdot m$$

$$S = \{(0,0), (0,1), \ldots, (1,0) \ldots\}$$

$$= \{0, \ldots n-1\} \times \{0, \ldots, m-1\}$$

$$= \{(x,y) \mid x \in \{0, \ldots m-1\} \quad y \in \{0, \ldots, m-1\}\}$$

$$A = \{\text{"L"}, \text{"R"}, \text{"U"}, \text{"D"}\}$$

$$= \{0, 1, 2, 3\}$$

$$= \left\{ \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right\}$$

$R = \{0\}$

P:

$p(s', r \mid s, a)$

$f'(s, \text{"R"}) = \begin{cases} s & \text{if } s \text{ is right edge} \\ s+(1) & \text{else} \end{cases}$

$\cdots$

$p(s' \mid s, a) = \begin{cases} 1 & s' = f(s,a) \\ 0 & \text{else} \end{cases}$

$\gamma = ?$

Inputs:

- Policy π
- Tolerance θ
- (Initial value function V_1)

$V(s) \leftarrow 0$ for all $s \in S$

While TRUE:

$\Delta \leftarrow 0$

For s in S:

$v \leftarrow V(s)$

$V(s) \leftarrow \mathbb{E}_{\pi}(R_{t+1} + \gamma V(S_{t+1}) | S_t)$

$\Delta \leftarrow \max(|v - V(s)|, \Delta)$

If $\Delta \leq \theta$:

Break

$$S = \{0, 1, \ldots, 99, 100\}$$

$$A(s) = \{1, \ldots, \min(s, 100 - s)\}$$

$$R = \{0, 1\}$$

$$p(a + s \mid s, a) = 0,4$$

$$p(s - a \mid s, a) = 0,6$$

if $s' = 100$, then $r = 1$

else $r = 0$