$V(4,2) = 1\,[-1 + 1 \cdot 5] = 4$

↑ P ↑ R ↑ γ ↑ $V(4,1)$

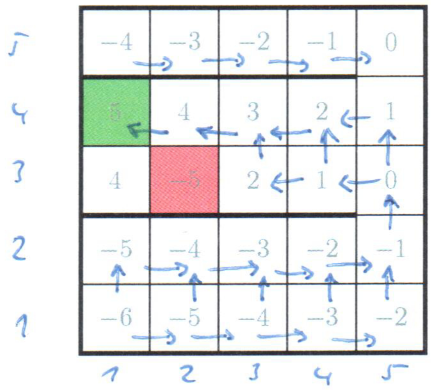
$\boxed{R = -1}$ ← empirically ermittelt!

~~R = 0 or R = 1 does not work!~~

$V(4,3) = 1\,[-1 + 1 \cdot 4] = 3$

$V(4,4) = 1\,[-1 + 1 \cdot 3] = 2$

(b) $V(3,4) \overset{\vdots}{=} \sum \begin{matrix} 0.5[-1+1\cdot 2] \\ 0.5[-1+1\cdot 2] \end{matrix} = \sum \begin{matrix} 0.5 \\ 0.5 \end{matrix} = 1$ u.s.w.



(c) No. Changing γ changes the value function but not the relative order. ✓

(d) The value function would change but the policy would not.
$\gamma = 0.8$, $R = 1$

4 Frozen Lake MDP

(c) Stochasticity generally increases the number of iterations required to converge. In the stochastic frozen lake environment, the number of iterations for value iteration increases. For policy iteration, depending on the implementation method, the number of iterations could remain unchanged; or policy iteration might not even converge at all. The stochasticity would also change the optimal policy. In this environment, the optimal policy of the stochastic frozen lake is different from the one of the deterministic frozen lake.

$V(4,2) = 1\,[\underbrace{-1+1}_{} + 0.8 \cdot 5] = 4$

$V(4,3) = 1\,[0 + 0.8 \cdot 4] = 3.2$

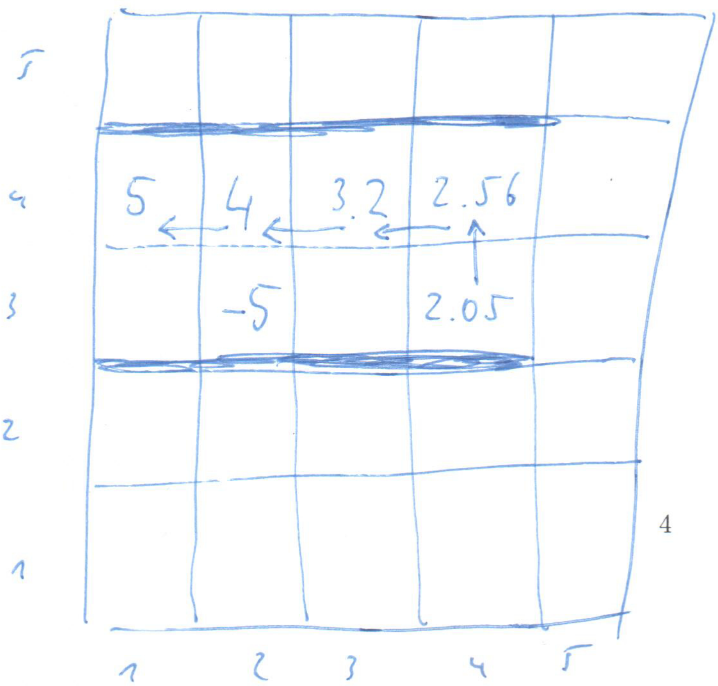
$V(4,4) = 1\,[0 + 0.8 \cdot 3.2] = 2.56$

$\vdots$

$V(3,4) = \sum \begin{matrix} 0.5\,[0 + 0.8 \cdot 2.56] \\ 0.5\,[0 + 0.8 \cdot 2.56] \end{matrix}$

$= \sum \begin{matrix} 1.024 \\ 1.024 \end{matrix}$

$= 2.048$



↝ Value function changes!

↝ Policy does not change!