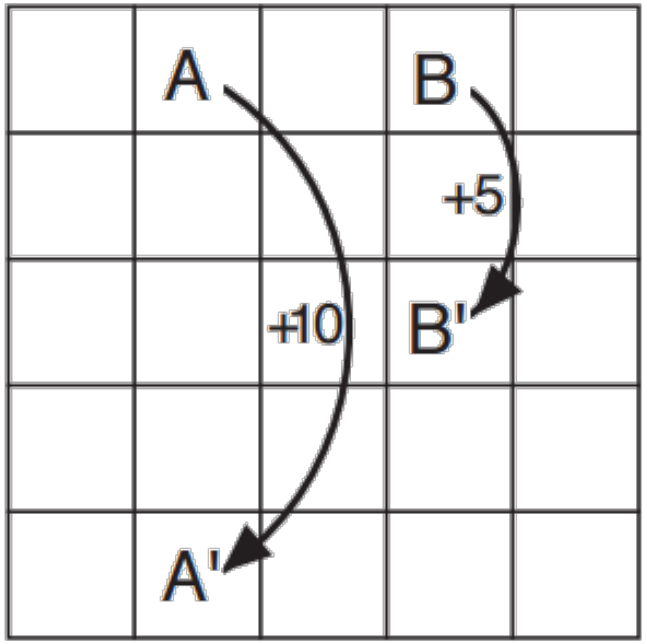
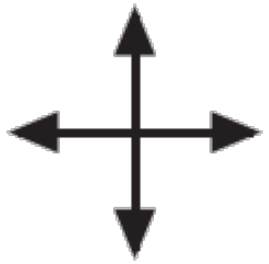
# Bellman Equation

State-value function
for equiprobable
random policy;
$\gamma = 0.9$



Actions

| | | | | |
|---|---|---|---|---|
| 3.3 | 8.8 | 4.4 | 5.3 | 1.5 |
| 1.5 | 3.0 | 2.3 | 1.9 | 0.5 |
| 0.1 | 0.7 | 0.7 | 0.4 | -0.4 |
| -1.0 | -0.4 | -0.4 | -0.6 | -1.2 |
| -1.9 | -1.3 | -1.2 | -1.4 | -2.0 |

- Actions that would take the agent off the grid leave its location unchanged, but also result in a rewa
- Other

$$= \sum_a \pi(a|s) \sum_{s',r} p(s', r \,|\, s, a) \Big[ r + \gamma v_\pi(s') \Big], \quad \text{for all } s \in \mathcal{S},$$

s A and B.