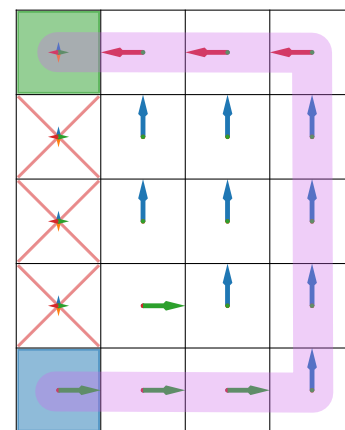


# Visualizing the learned policy at different steps

Sub-optimal, but safest policy



Optimal, but riskiest policy

