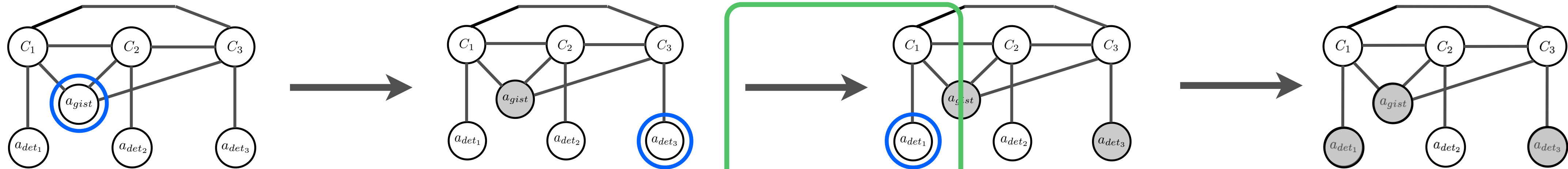
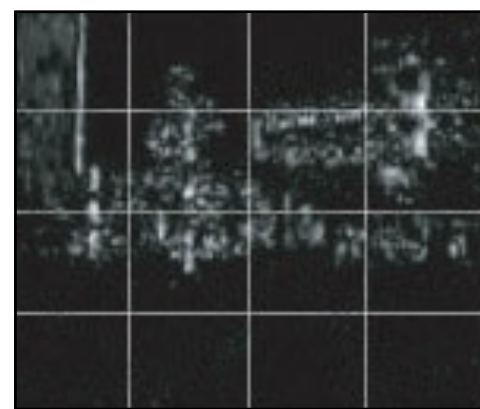


Belief state s



action selection
maximize expected value

Action $a \in \mathcal{A}$

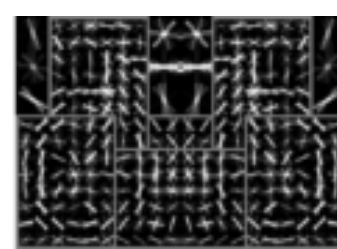


scene context

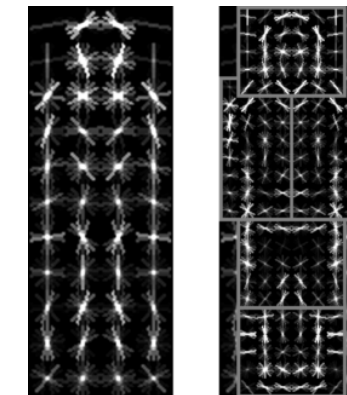
execute action
takes time;
receive observations



bicycle detector

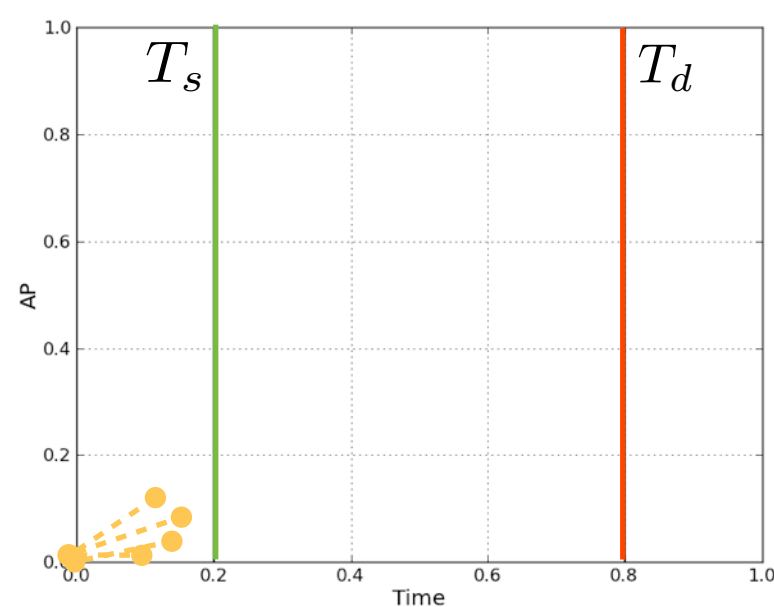


person detector

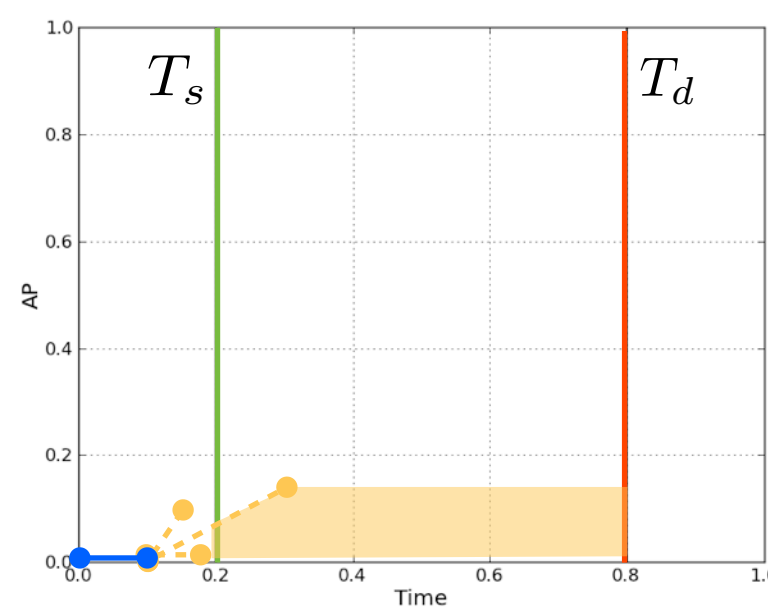


etc.

Reward $R(s, a)$

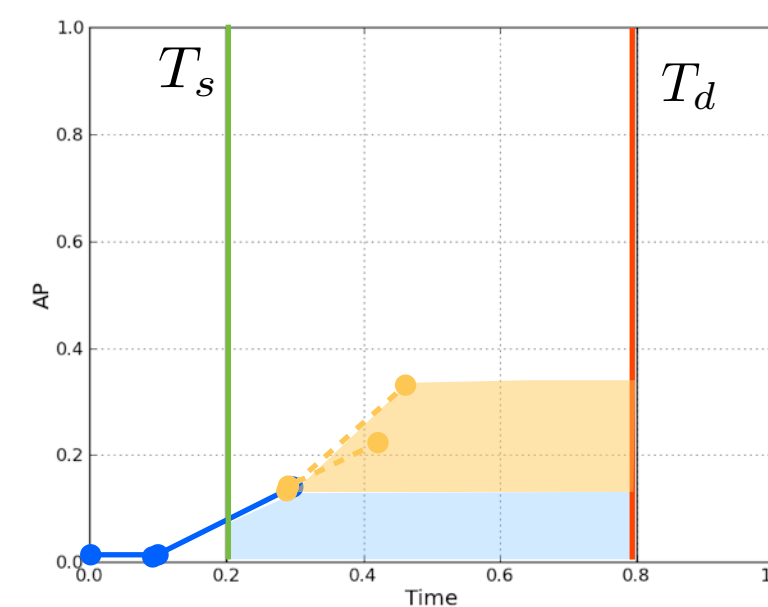


$t = 0$

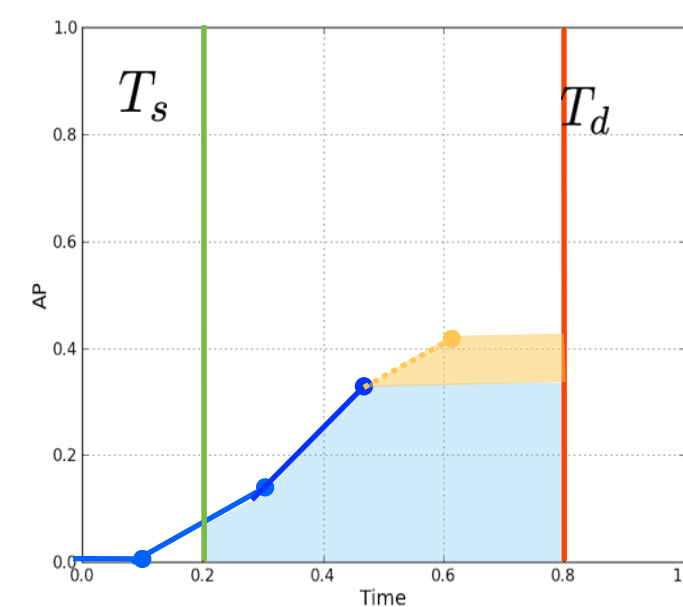


$t = 0.1$

time



$t = 0.3$



$t = 0.5$