

BLINDFOLDED ROBOT

STATE: JOINT ANGLES OF THE ARM

ACTION: TORQUE

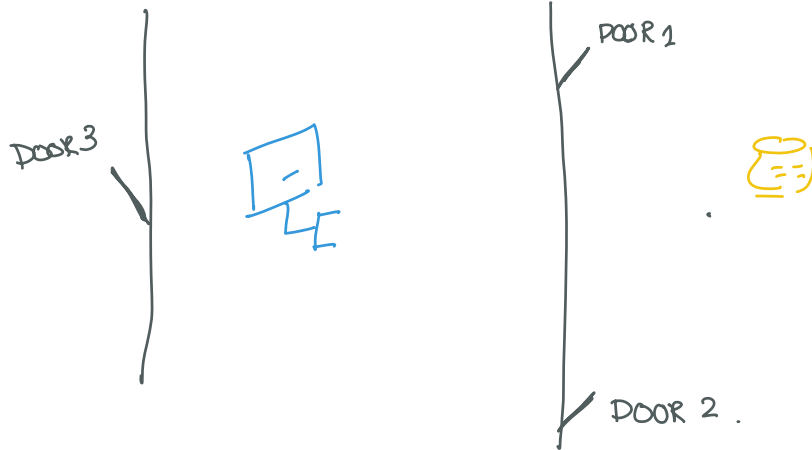
COST: +1 TILL YOU REACH THE GOAL

TRANSITION:

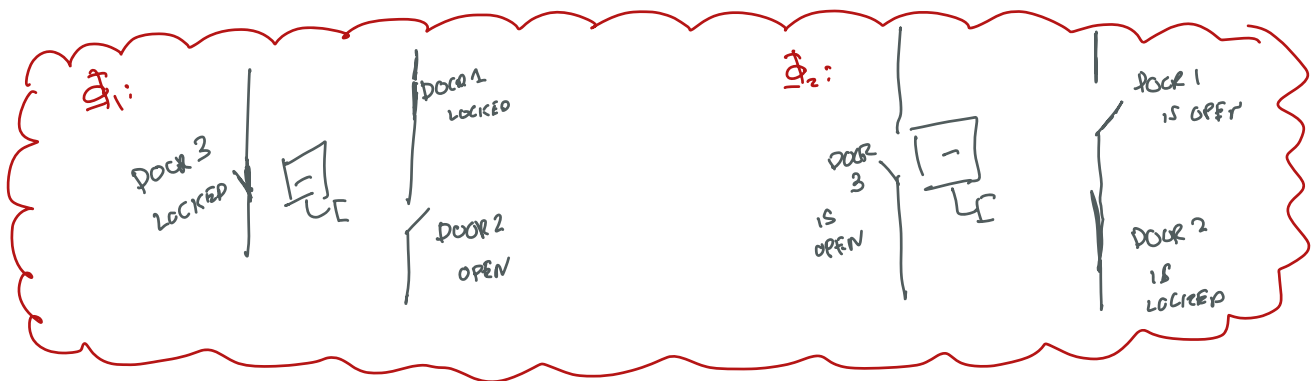
$$P(s_{t+1} | s_t, a_t, \Phi)$$

UNCERTAIN

CONFIGURATIONS OF THE WORLD
(LOCATION, GEOMETRY OF FRIDGE)



Φ : WHICH DOORS ARE LOCKED?



S: 2ⁿ LOCATION

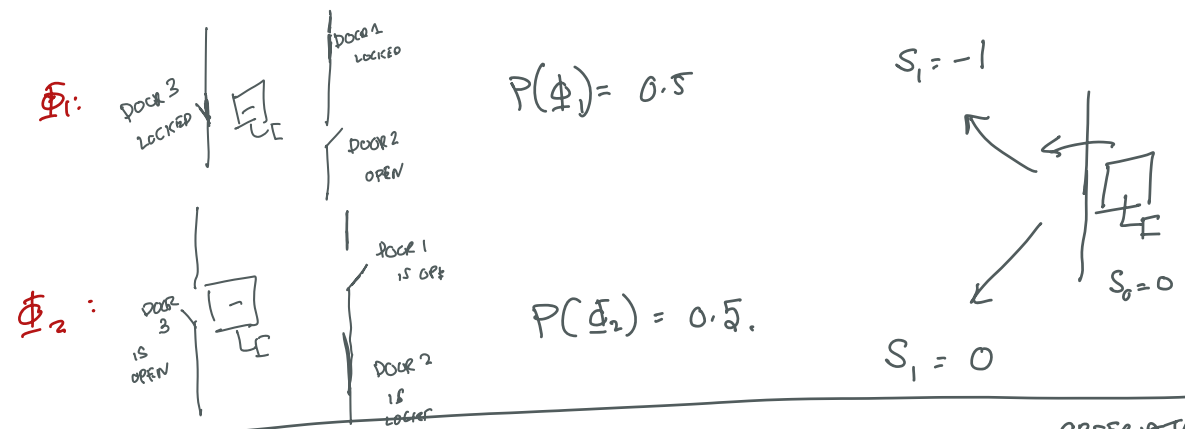
A: 4 ACTIONS

C: +1 TILL GOAL REACHED

T: $P(s_{t+1} | s_t, a_t, \Phi)$

$$\tilde{\gamma}(s_t, a_t, s_{t+1}) := \sum_{\Phi} P(\Phi) P(s_{t+1} | s_t, a_t, \Phi)$$

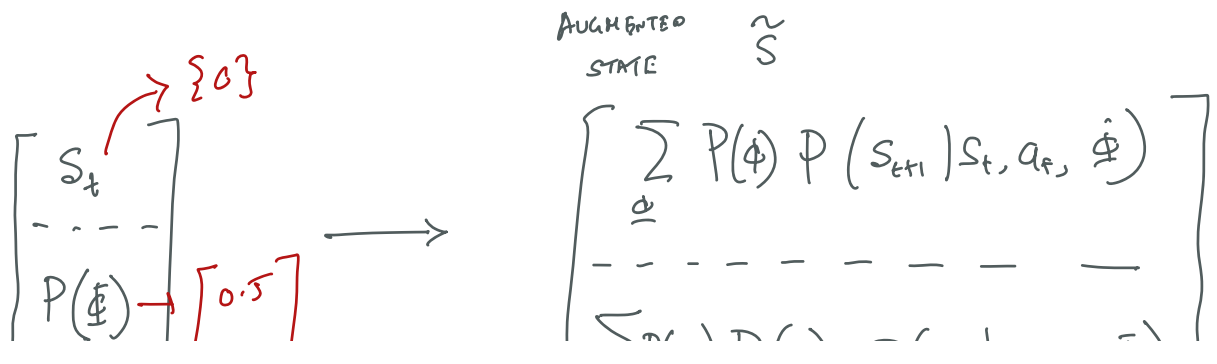
How DOES MY BELIEF $P(\Phi)$ EVOLVE?



POSTERIOR	PRIOR.	OBSERVATION
$P(\Phi s_0, a_0, s_1) \propto P(\Phi) \cdot P(s_1 s_0, a_0, \Phi)$		

MDP $\langle S, A, C, T \rangle \longrightarrow$ BELIEF MDP $\langle \begin{bmatrix} S \\ P(\Phi) \end{bmatrix}, A, C, \tilde{\gamma} \rangle$

$P(\tilde{s}_{t+1} | \tilde{s}_t, a_t)$



\sum_t

$$\sqrt{[0.5]}$$

$$\left\{ \sum_{s_{t+1}} P(s_{t+1}) P(\phi) \cdot P(s_{t+1} | s_t, a_t, \phi) \right\}$$

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} - \\ + \\ - \end{bmatrix}$$

$$S_0 = \{0\}$$

$$a_0 = \{\text{LEFT}\}$$

$$\begin{bmatrix} S_t \rightarrow \{0\} \\ \dots \\ P_t(\phi) \rightarrow \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} \end{bmatrix}$$

\sum

50%

$$\begin{bmatrix} S_t \rightarrow \{-1\} \\ \dots \\ P_t(\phi) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{bmatrix}$$

50%

$$\begin{bmatrix} S_t = \{0\} \\ \dots \\ P_t(\phi) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{bmatrix}$$