

$\langle S, A, R, T \rangle$

$a_t = (\uparrow, \downarrow, \leftarrow, \rightarrow, \times)$

$s_t = (x_t, y_t)$

$r(s_t, a_t) = \begin{cases} 1 & s_t = \text{GOAL} \\ 0 & \text{o.w.} \end{cases}$

$O_t = (\text{LIGHT} / \text{NO LIGHT})$

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

$P(O_t^{\text{LIGHT}} | s_t)$

$P(s_0) = \text{PRIOR}$

$= \begin{cases} 0.9 & \text{if } (s_t - s_{\text{prev}}) < \delta \\ 0.1 & \text{o.} \end{cases}$

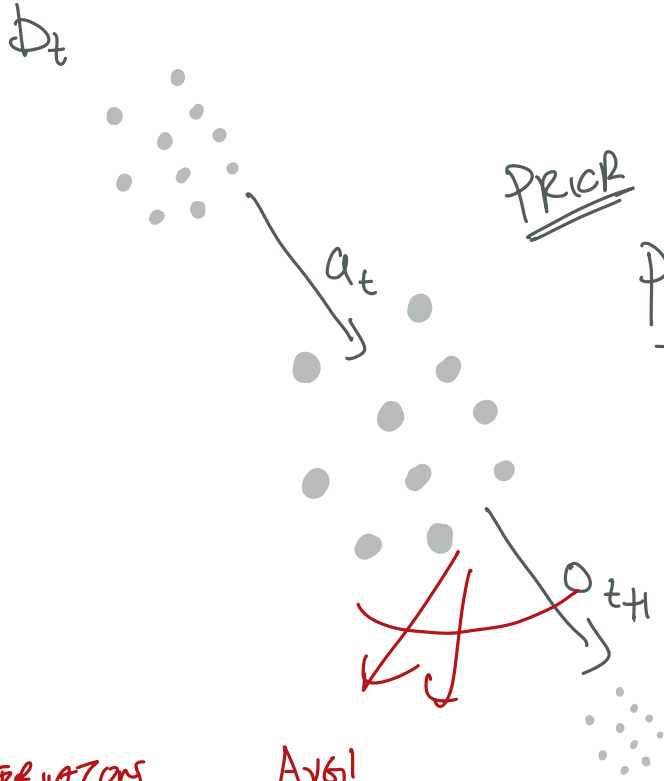
TECHNIQUE : CONFIDENTLY CONVERT TO BELIEF MDP

$\langle S, A, R, T \rangle \rightarrow \langle B, A, R_B, T_B \rangle$

$b_t = P(s_t) =$

$a_t = \text{THE SAME} = (\uparrow, \downarrow, \leftarrow, \rightarrow, \times)$

$$r(b_t, a_t) = \sum_{s_t} r(s_t, a_t) b_t(s_t)$$



$$\frac{P(s_{t+1})}{b_{t+1}} = \sum_{s_t} \frac{P(s_{t+1} | s_t, a_t)}{\text{MDP TRANSITION}} \frac{P(s_t)}{D_t}$$

$$D_{t+1} = \tilde{\tau}_b(b_t, a_t, O_{t+1})$$

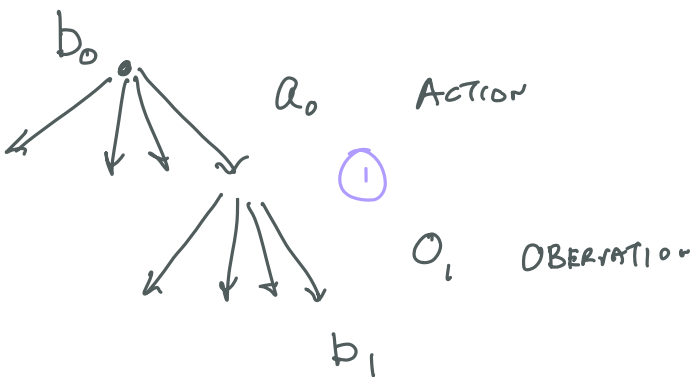
POSTERIOR

CAN WE CHOOSE OBSERVATIONS.

AVG!

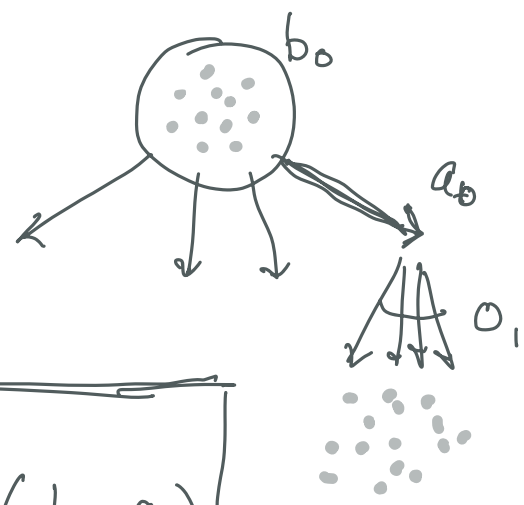
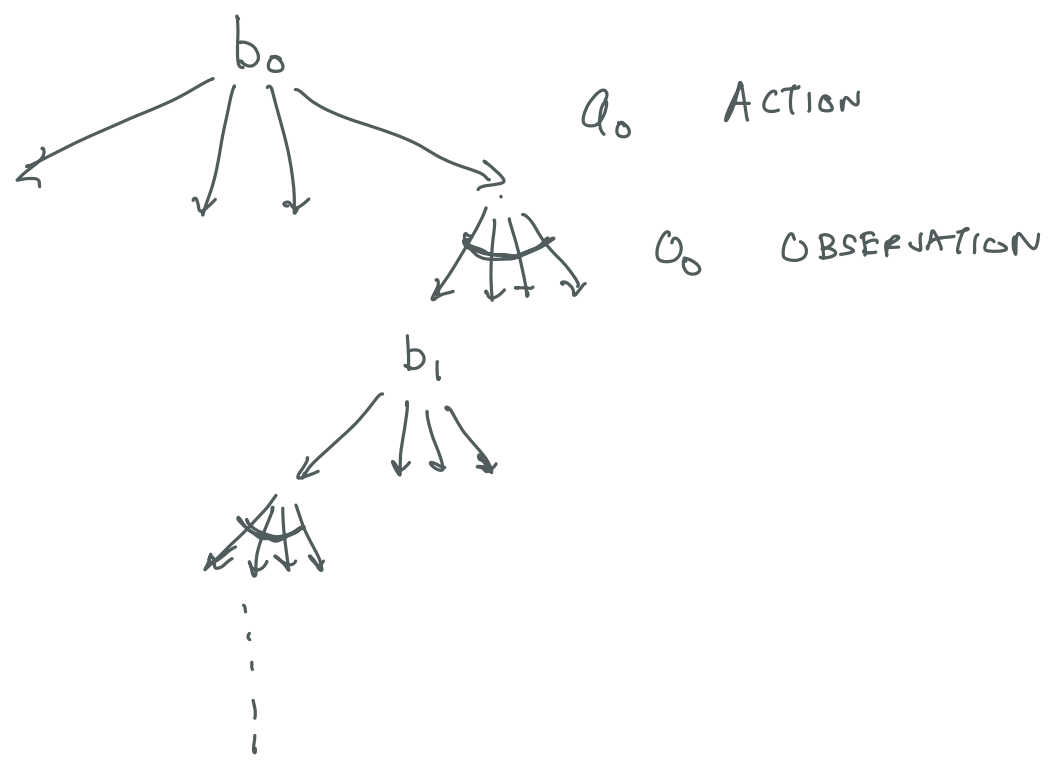
$$P(b_{t+1} | b_t, a_t) = \sum_{O_{t+1}} \tau_b(b_t, a_t, O_{t+1}) \times P(O_{t+1} | b_t, a_t)$$

$$\frac{P(s_{t+1} | O_{t+1}, a_t, b_t)}{b_{t+1}} \propto P(O_{t+1} | s_{t+1}) \sum_{s_t} P(s_{t+1} | s_t, a_t) P(s_t)$$



$$b_{t+1} = \tau_b(b_t, a_t, O_{t+1})$$

Discrete
state \rightarrow BELIEF
SPACE
IS
CONTINUOUS



- ① Sample a_t
- ② Average O_{t+1}
- ③

$$\max_{\pi} \mathbb{E} \sum_{t=1}^{\tau} r(b_t, a_t)$$

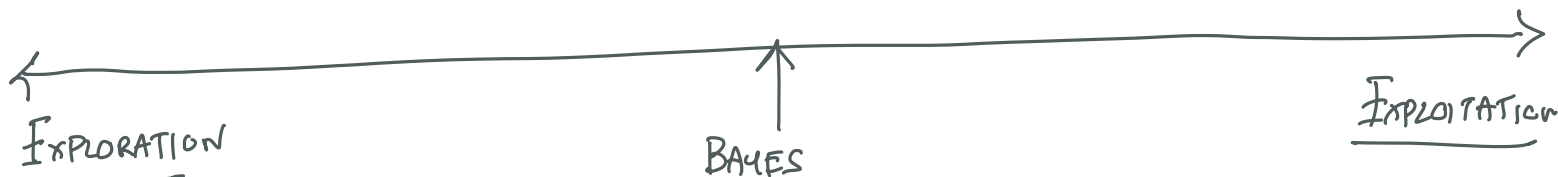
$$\pi: b_t \rightarrow a_t$$

CONCLUSION : BELIEF SPACE PLANNING IS HARD. WHAT ARE

GOOD HEURISTICS?

Given your belief b_t

$$\alpha \max_{a_t} f(b_t)$$



INFORMATION GAIN

$$\max_{a_t} H(b_t) - H(b_{t+1} | a_t)$$

\downarrow

$$- \sum_{s_t} P(s_t) \log P(s_t)$$

QMDP

(HINDSIGHT OPT)

$$\max_{a_t} E_{s_t \sim b_t} Q(s_t, a_t)$$