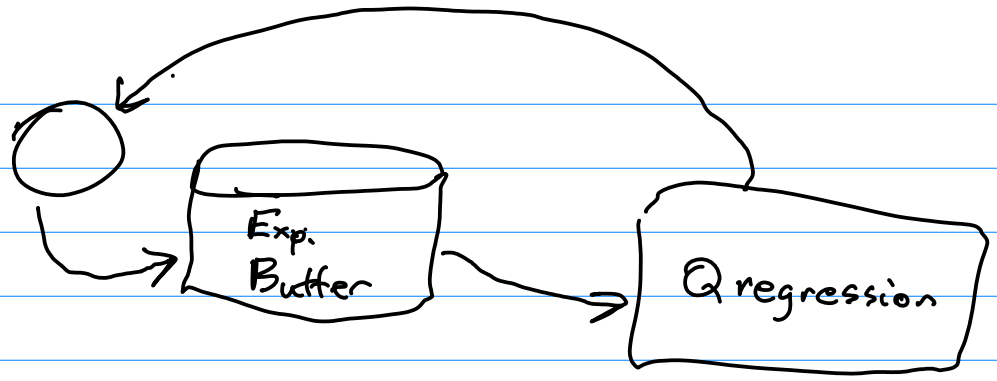


Last-Time

- Discrete Bayesian Filter

This Time

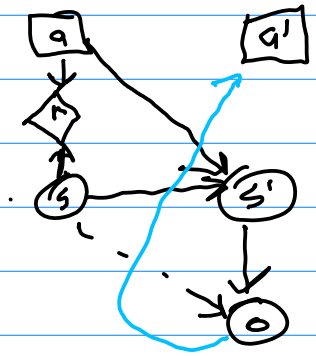
POMDPs
Particle Filters
Value of Information

Outcome	Aleatory		MDP (known model)
Model	Epistemic	Static	RL
State	Epistemic	Dynamic	POMDP

MDP (S, A, T, R, γ)

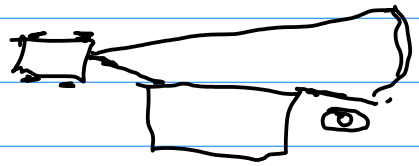
POMDP $(S, A, T, R, O, Z, \gamma)$

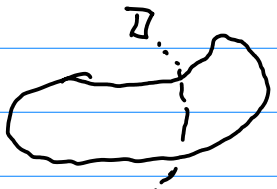
$Z(o \mid a, s')$ could be $Z(o \mid s, a, s')$

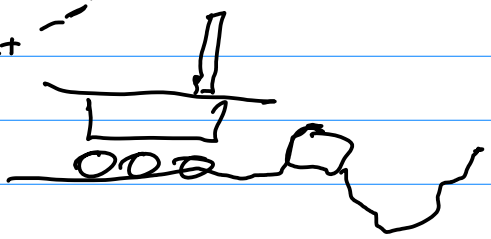


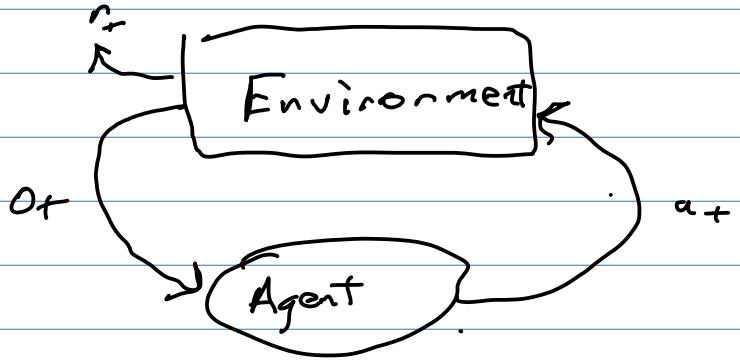
Examples

- Automated Driving
- ACAS X
- Space Exploration
- Disease Monitoring/Treatment
- HVAC
- Water Management
- Wildfire
- HRI









So far $o_t = s_t$ $\pi(s_t)$

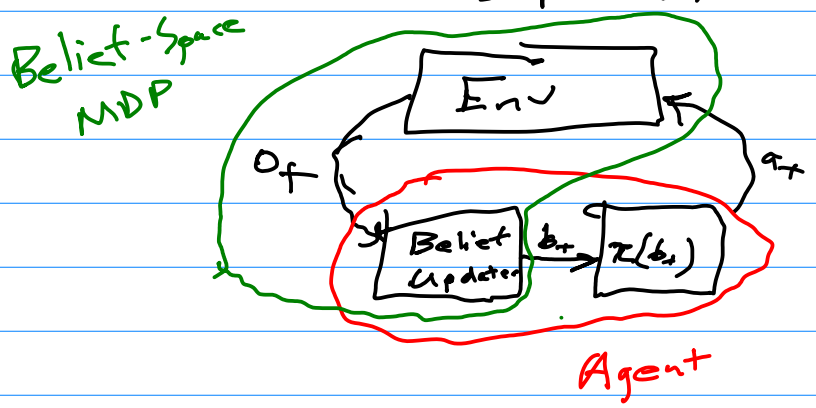
now $o_t \sim Z(a_{t-1}, s_t)$

MDPs: optimal policy based on last state

POMDPs: optimal policy based on last obs

- history: $h_t = (o_0, a_0, o_1, \ldots, a_{t-1}, o_t)$
- belief: $b_t(s) = P(s_t = s \mid h_t)$

POMDP Simulation




```
b ← p0
s ← rand(p0)
loop
    a ← π(b)
    s, o, r ← G(s, a)
    b ← τ(b, a, o)
```

A POMDP is an MDP on the Belief Space

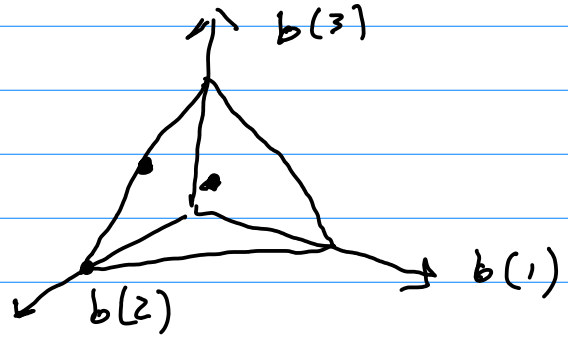
$$POMDP(S, A, T, R, O, Z, \gamma) = MDP(B(S), A, \tau, R', \gamma)$$

$$B(S) \subset [0,1]^{|S|}$$

$$B(S) = \{b \in [0,1]^{|S|} : \sum_i b(i) = 1\}$$

for $S = \{1, 2\}$ represent b with 1 number

for $S = \{1, 2, 3\}$



τ "Belief Update"

$b' = \tau(b, a, o)$

$b'(s') \propto Z(o|a, s') \sum_s T(s'|s, a)\, b(s)$

replace with integral

$O(|S|^2)$

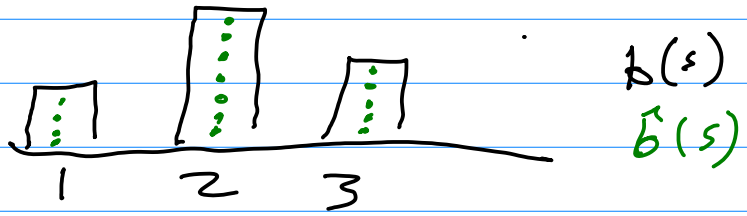
$R'(b, a) = \mathop{E}_{s \sim b}[R(s, a)]$

Particle Filter

"Rejection / Unweighted" PF

$\hat{b} = [\hat{s}_1, \hat{s}_2, \hat{s}_3, \ldots \hat{s}_K]$

$\hat{b}(s) = \frac{\sum_k 1(\hat{s}_k = s)}{|\hat{b}|}$



$b(s)$

$\hat{b}(s)$

Rejection PF update

receive o

$\hat{b}' = [\,]$

for $\hat{s}$ in $\hat{b}$ ←

$\hat{s}', \hat{o} \leftarrow G(\hat{s}, a)$

if $\hat{o} = o$

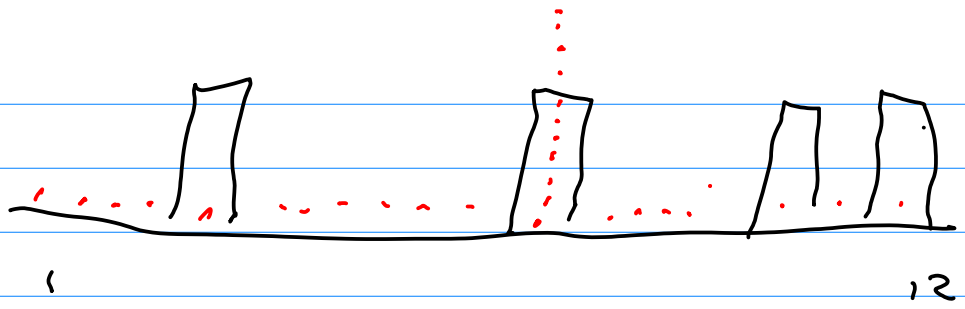
push! $(\hat{b}', \hat{s}')$

output $\hat{b}'$

Requires

$G(s, a)$

← replenish randomly or with prior $\hat{b}$ based on o



Particle Depletion

Preventing Particle Depletion
- Add noise
- Initialize Random particles and add to b

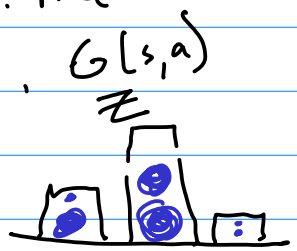
Bootstrap Filter

receive o
$\tilde{b} = [\ ]$ (s, w)
for $\hat{s}$ in $\hat{b}$
$\hat{s}' \leftarrow G(\hat{s}, a)$
$w \leftarrow Z(o \mid a, \hat{s}')$
push! $(\tilde{b}, (\hat{s}', w))$
$\hat{b}' = \text{rand}(\tilde{b}, |\hat{b}|)$
↖ samples from $\tilde{b}$ according to w

~~$O(n^2)$~~ → $O(n)$

Weighted PF
Importance Sampling
SIR

Requires
$G(s, a)$
Z



$b'(s) \approx \sum_k \hat{s}_k w_k$ ($\tilde{b}(k)$)