Recap

Recap

- POMDP

Recap

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

Recap

- POMDP
- Belief Updates

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

Recap

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

$$b_t(s) = P(s_t = s \mid h_t)$$

Recap

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

$$b_t(s) = P(s_t = s \mid h_t)$$

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

Recap

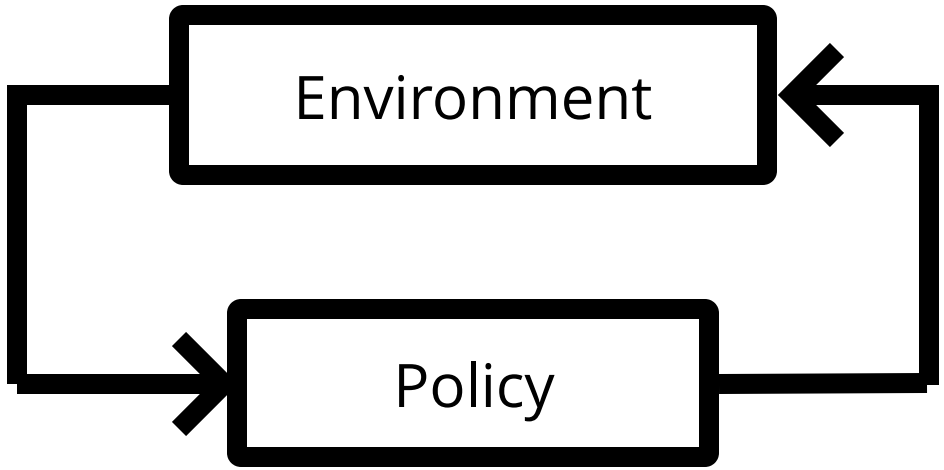
- POMDP $(S, A, O, R, T, Z, \gamma)$
- Belief Updates

$$b_t(s) = P(s_t = s \mid h_t)$$

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

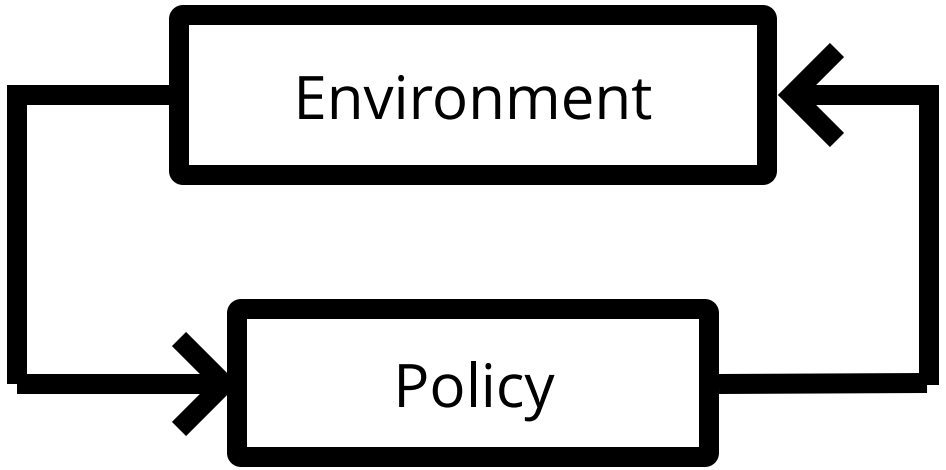
$$b'(s') \propto Z(o \mid a, s') \sum_s T(s' \mid s, a)\, b(s)$$

MDP Sense-Plan-Act Loop

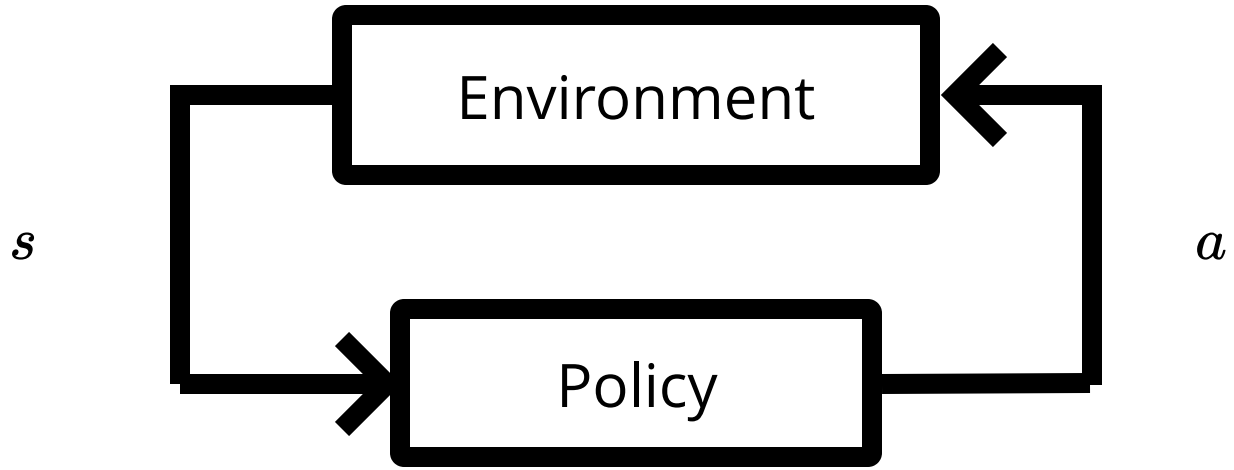


MDP Sense-Plan-Act Loop

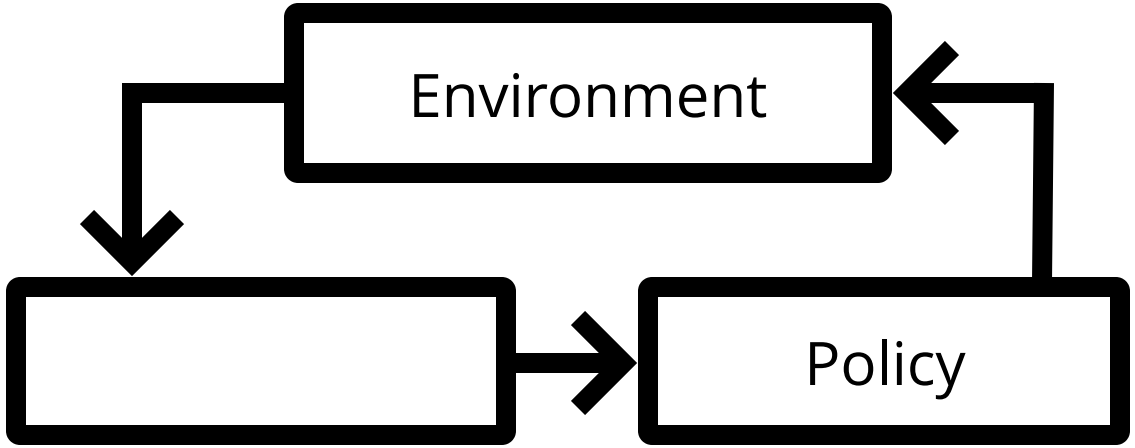
s



MDP Sense-Plan-Act Loop



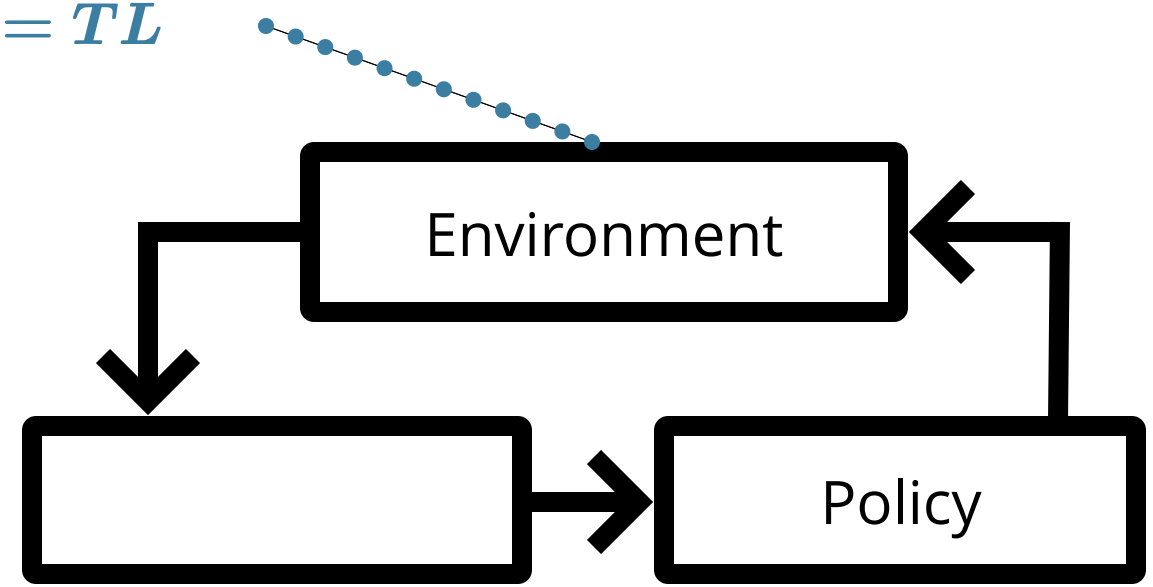
POMDP Sense-Plan-Act Loop



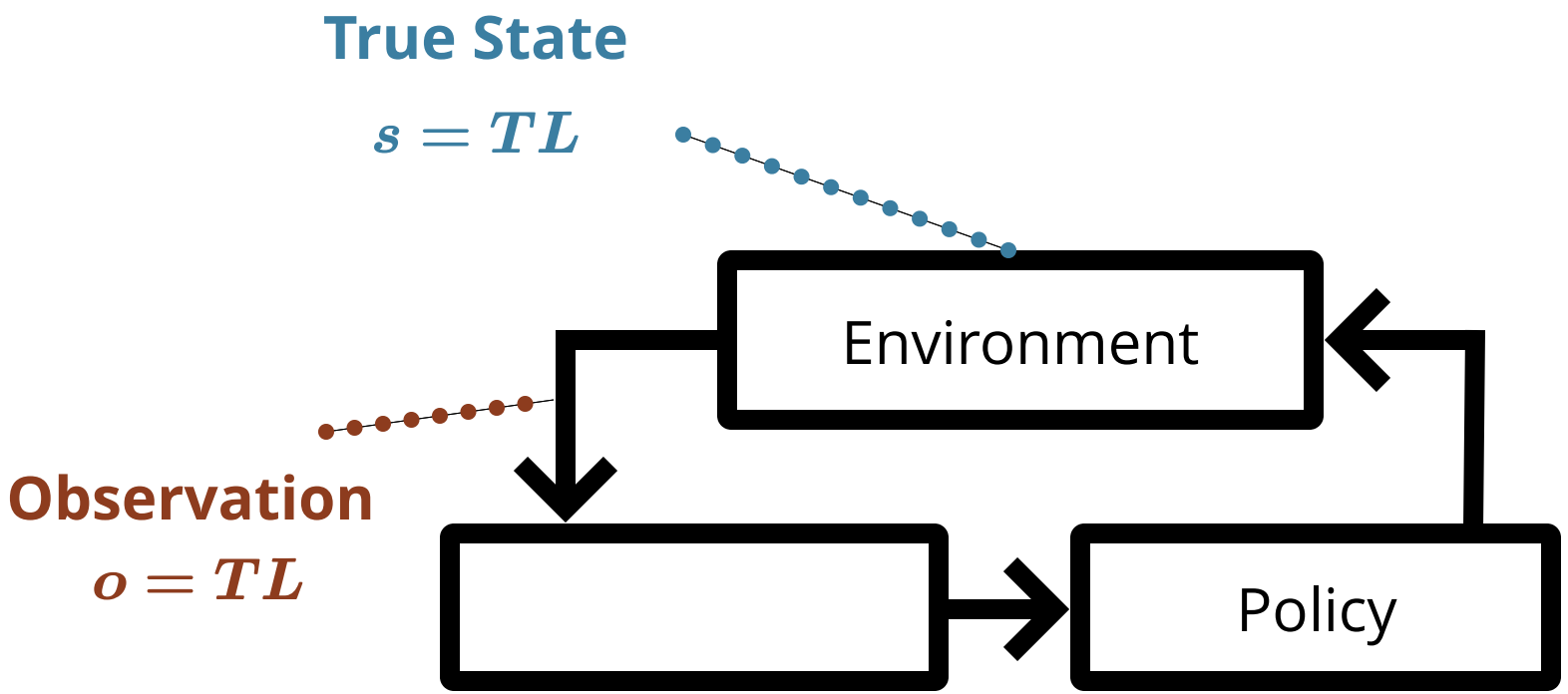
POMDP Sense-Plan-Act Loop

True State

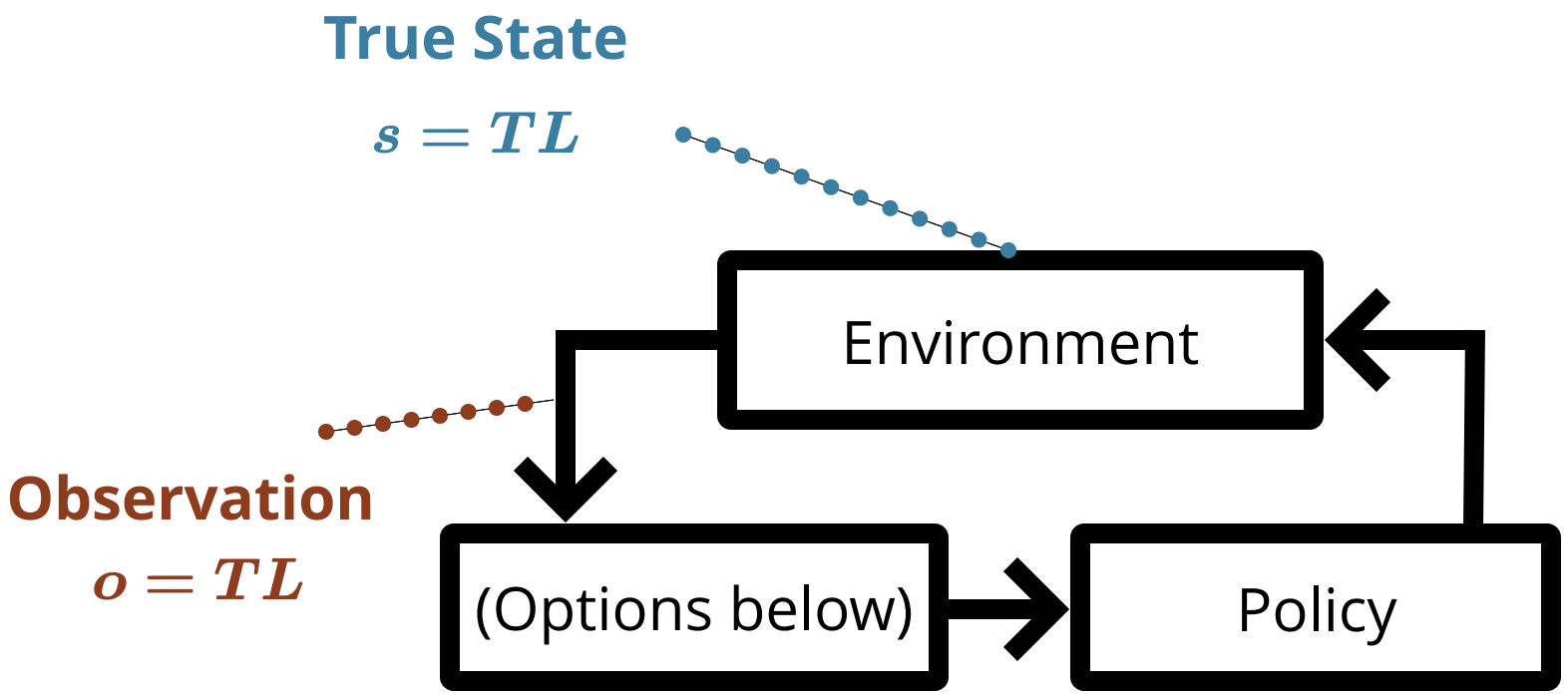
$$s = TL$$



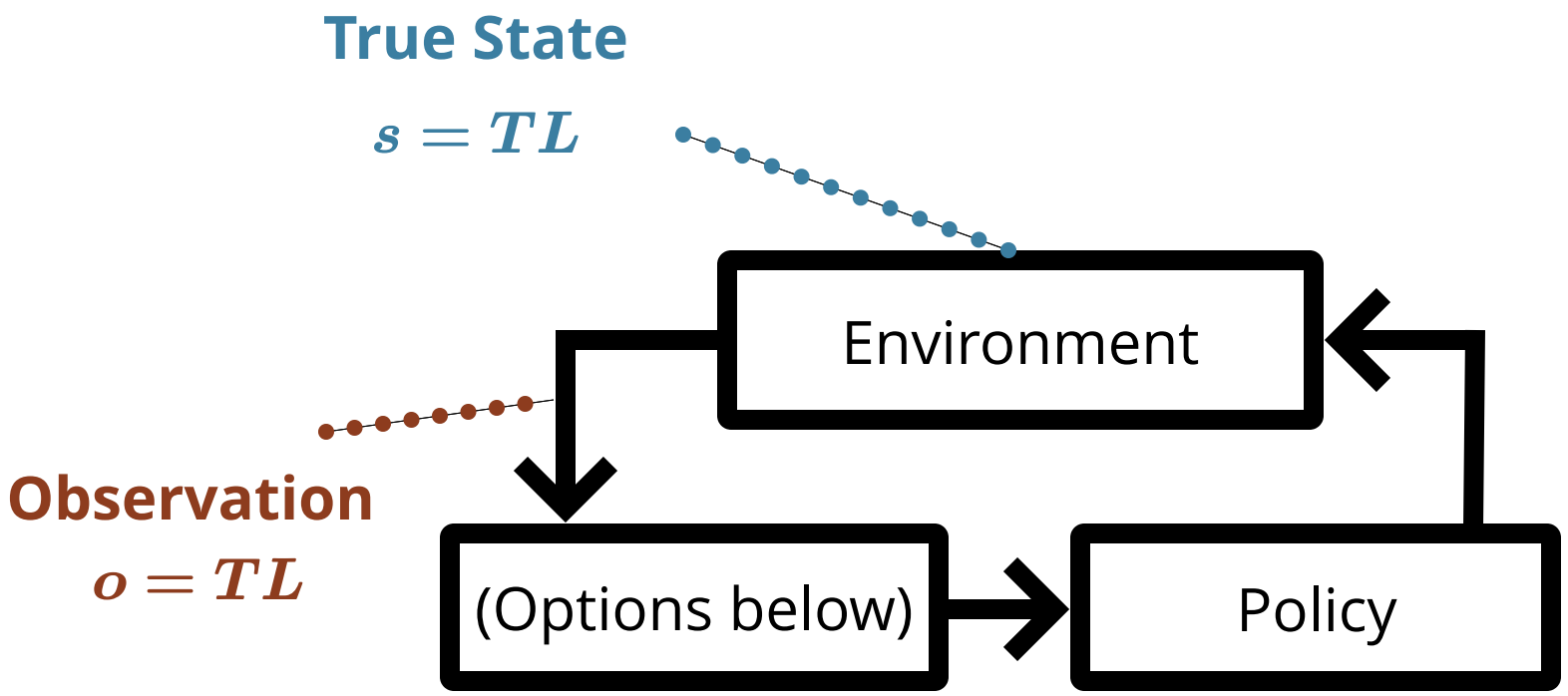
POMDP Sense-Plan-Act Loop



POMDP Sense-Plan-Act Loop

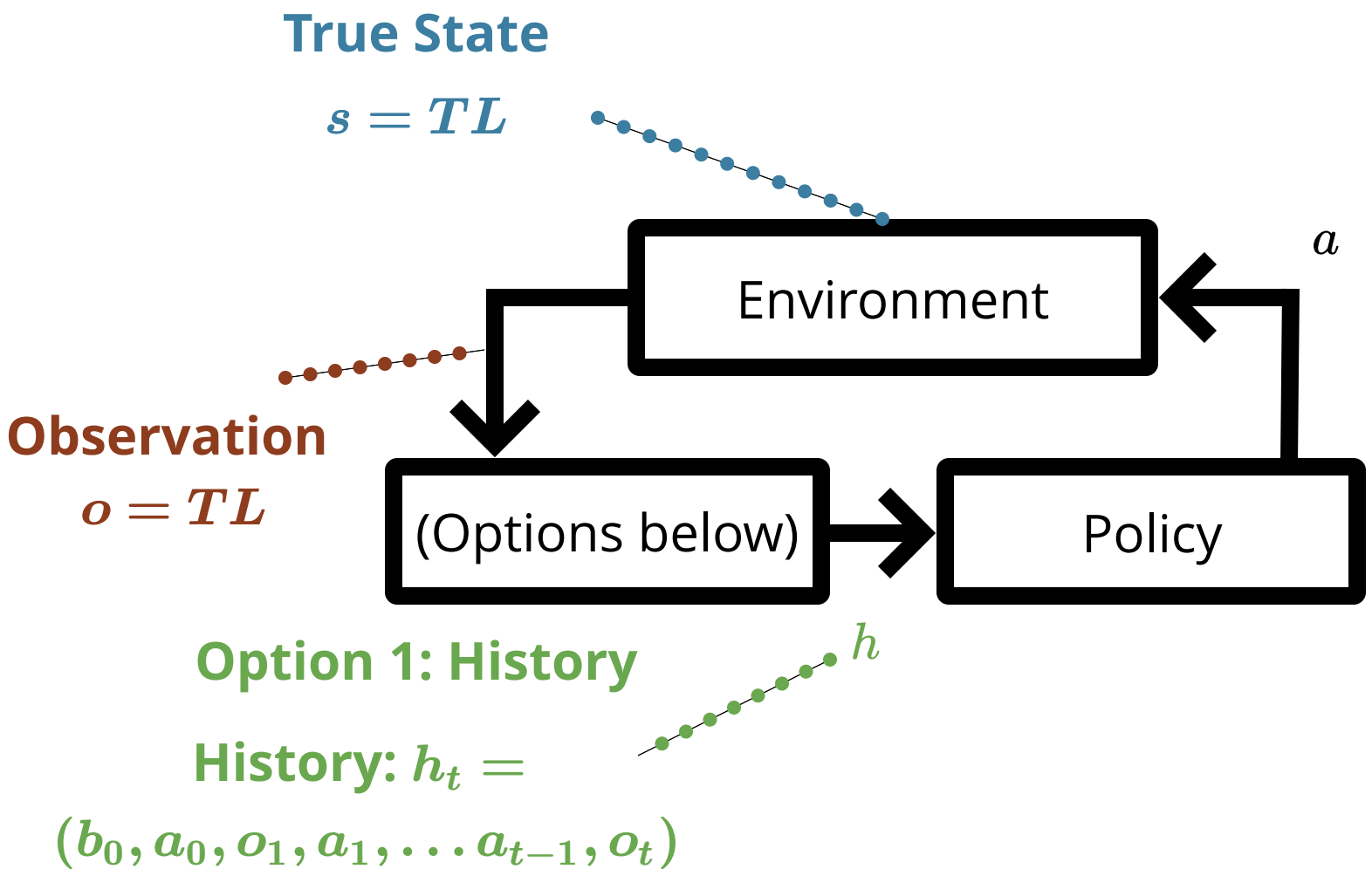


POMDP Sense-Plan-Act Loop

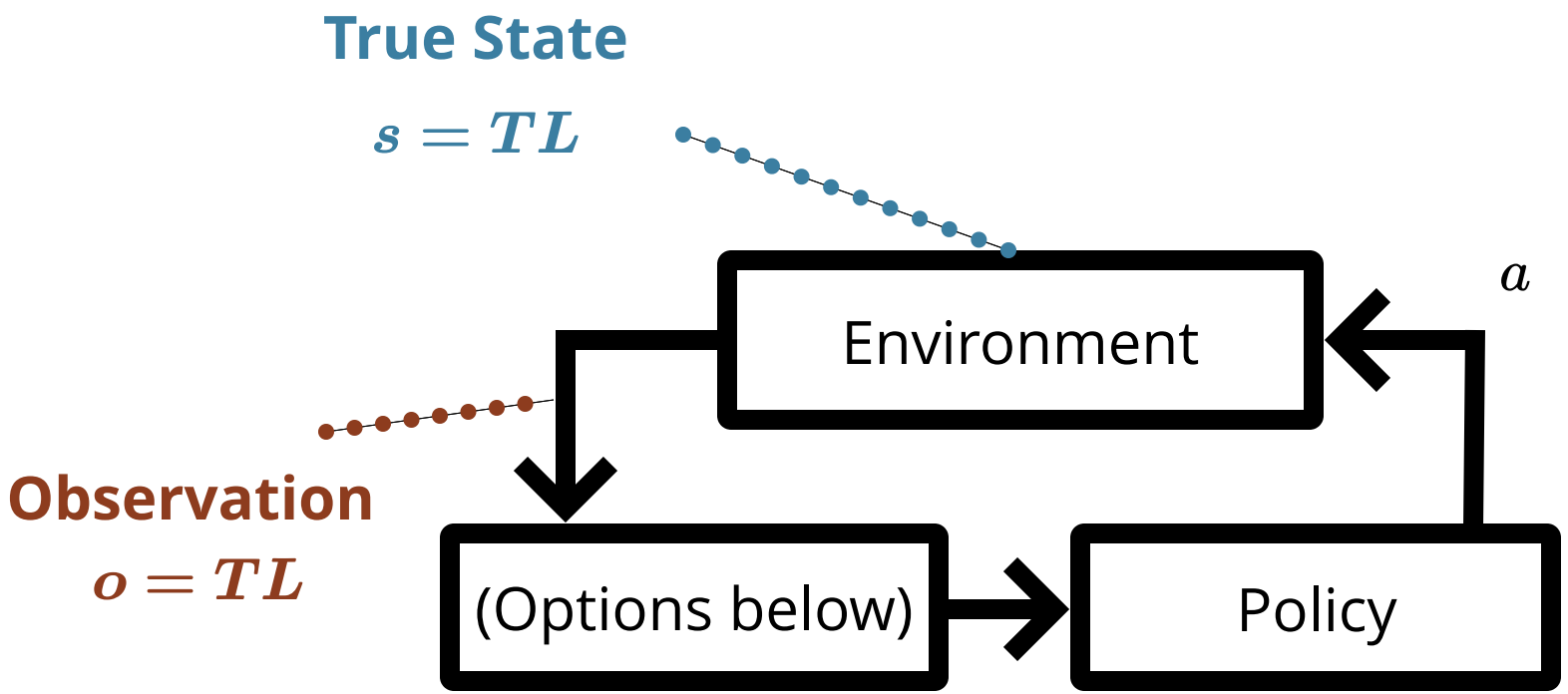


Option 1: History

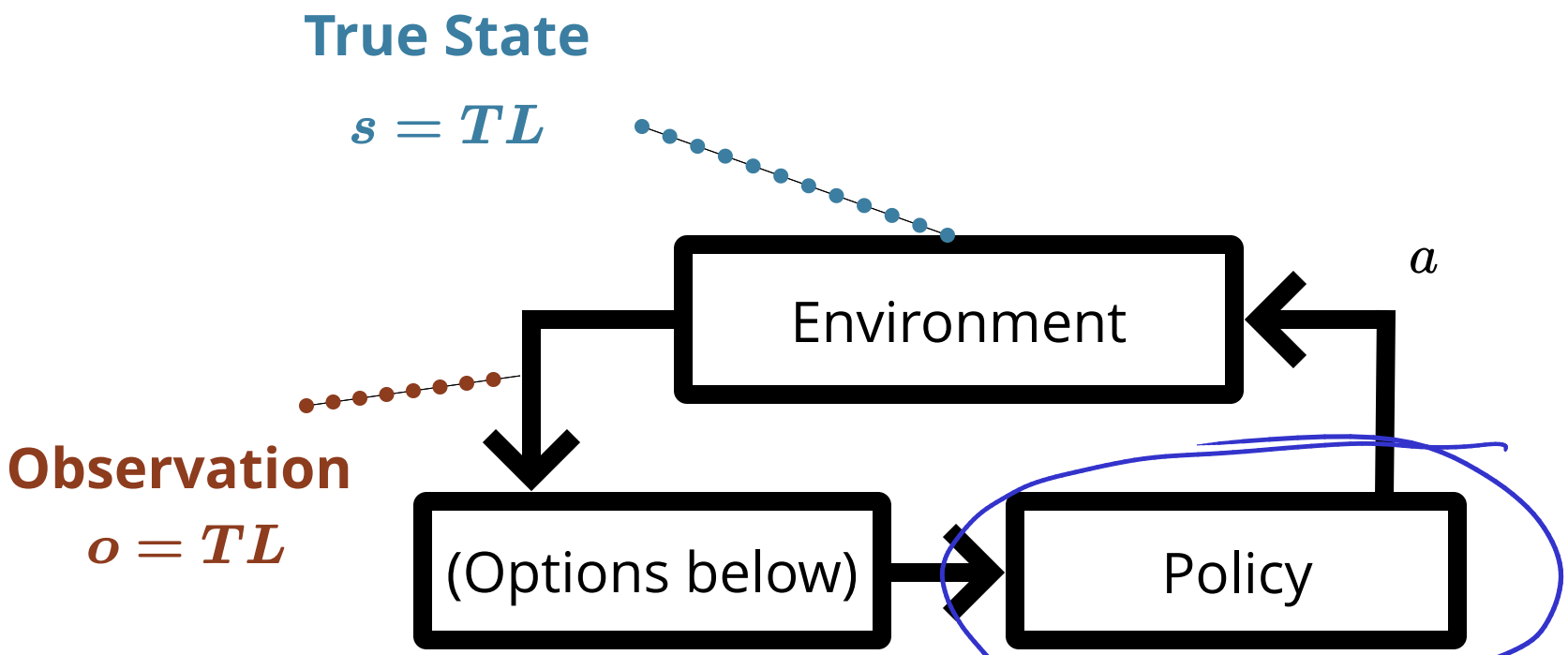
POMDP Sense-Plan-Act Loop



POMDP Sense-Plan-Act Loop



POMDP Sense-Plan-Act Loop



Exercise 1: Crying Baby Belief Update

Exercise 1: Crying Baby Belief Update

$S = \{h, \neg h\}$

$A = \{f, \neg f\}$

$O = \{c, \neg c\}$

Exercise 1: Crying Baby Belief Update

$S = \{h, \neg h\}$

$A = \{f, \neg f\}$

$O = \{c, \neg c\}$

$R(s, a) = R(s) + R(a)$

Exercise 1: Crying Baby Belief Update

$$S = \{h, \neg h\}$$

$$A = \{f, \neg f\}$$

$$O = \{c, \neg c\}$$

$$R(s, a) = R(s) + R(a)$$

$$R(s) = \begin{cases} -10 \text{ if } s = h \\ 0 \text{ otherwise} \end{cases}$$

Exercise 1: Crying Baby Belief Update

$$S = \{h, \neg h\}$$

$$A = \{f, \neg f\}$$

$$O = \{c, \neg c\}$$

$$R(s, a) = R(s) + R(a)$$

$$R(s) = \begin{cases} -10 \text{ if } s = h \\ 0 \text{ otherwise} \end{cases}$$

$$R(a) = \begin{cases} -5 \text{ if } a = f \\ 0 \text{ otherwise} \end{cases}$$

Exercise 1: Crying Baby Belief Update

$S = \{h, \neg h\}$ $\quad$ $T(h \mid h, \neg f) = 1.0$

$A = \{f, \neg f\}$ $\quad$ $T(h \mid \neg h, \neg f) = 0.1$

$O = \{c, \neg c\}$ $\quad$ $T(\neg h \mid \cdot, f) = 1.0$

$$R(s, a) = R(s) + R(a)$$

$$R(s) = \begin{cases} -10 \text{ if } s = h \\ 0 \text{ otherwise} \end{cases}$$

$$R(a) = \begin{cases} -5 \text{ if } a = f \\ 0 \text{ otherwise} \end{cases}$$

Exercise 1: Crying Baby Belief Update

$S = \{h, \neg h\}$

$A = \{f, \neg f\}$

$O = \{c, \neg c\}$

$T(h \mid h, \neg f) = 1.0$

$T(h \mid \neg h, \neg f) = 0.1$

$T(\neg h \mid \cdot, f) = 1.0$

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

$$R(s, a) = R(s) + R(a)$$

$$R(s) = \begin{cases} -10 \text{ if } s = h \\ 0 \text{ otherwise} \end{cases}$$

$$R(a) = \begin{cases} -5 \text{ if } a = f \\ 0 \text{ otherwise} \end{cases}$$

$Z(c \mid \cdot, h) = 0.8)$

$Z(c \mid \cdot, \neg h) = 0.1$

Exercise 1: Crying Baby Belief Update

$S = \{h, \neg h\}$

$A = \{f, \neg f\}$

$O = \{c, \neg c\}$

$T(h \mid h, \neg f) = 1.0$

$T(h \mid \neg h, \neg f) = 0.1$

$T(\neg h \mid \cdot, f) = 1.0$

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

$$R(s, a) = R(s) + R(a)$$

$$R(s) = \begin{cases} -10 \text{ if } s = h \\ 0 \text{ otherwise} \end{cases}$$

$$R(a) = \begin{cases} -5 \text{ if } a = f \\ 0 \text{ otherwise} \end{cases}$$

$Z(c \mid \cdot, h) = 0.8)$

$Z(c \mid \cdot, \neg h) = 0.1$

$\gamma = 0.9$

Exercise 1: Crying Baby Belief Update

$S = \{h, \neg h\}$

$A = \{f, \neg f\}$

$O = \{c, \neg c\}$

$T(h \mid h, \neg f) = 1.0$

$T(h \mid \neg h, \neg f) = 0.1$

$T(\neg h \mid \cdot, f) = 1.0$

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

$$R(s, a) = R(s) + R(a)$$

$$R(s) = \begin{cases} -10 \text{ if } s = h \\ 0 \text{ otherwise} \end{cases}$$

$$R(a) = \begin{cases} -5 \text{ if } a = f \\ 0 \text{ otherwise} \end{cases}$$

$Z(c \mid \cdot, h) = 0.8)$

$Z(c \mid \cdot, \neg h) = 0.1$

$\gamma = 0.9$

Starting at a $b(h) = 0$, calculate b' with $a = \neg f$ and $o = c$.

$b'(h) \propto 0.8\,(0.1\,(1.0) + 1.0\,(0.0)) = 0.08$

$b'(h) \propto Z(c \mid \neg f, h)\,(T(h \mid \neg h, \neg f)\, b(\neg h) + T(h \mid h, \neg f)\, b(h))$

$b'(h) \propto 0.08$

$b'(\neg h) \propto Z(c \mid \neg f, \neg h)\,(T(\neg h \mid \neg h, \neg f)\, b(\neg h) + T(\neg h \mid h, \neg f)\, b(h))$

$= 0.09$

$b'(h) = 0.08/(0.08 + 0.09) = 47\%$

$b'(\neg h) = 0.09/(0.08 + 0.09) = 53\%$

$b'(h) =$

$b'(\neg h) =$

Belief Dynamics

Belief Dynamics

$$b'(s') \propto Z(o \mid a, s') \sum_{s} T(s' \mid s, a)\, b(s)$$

Belief Dynamics

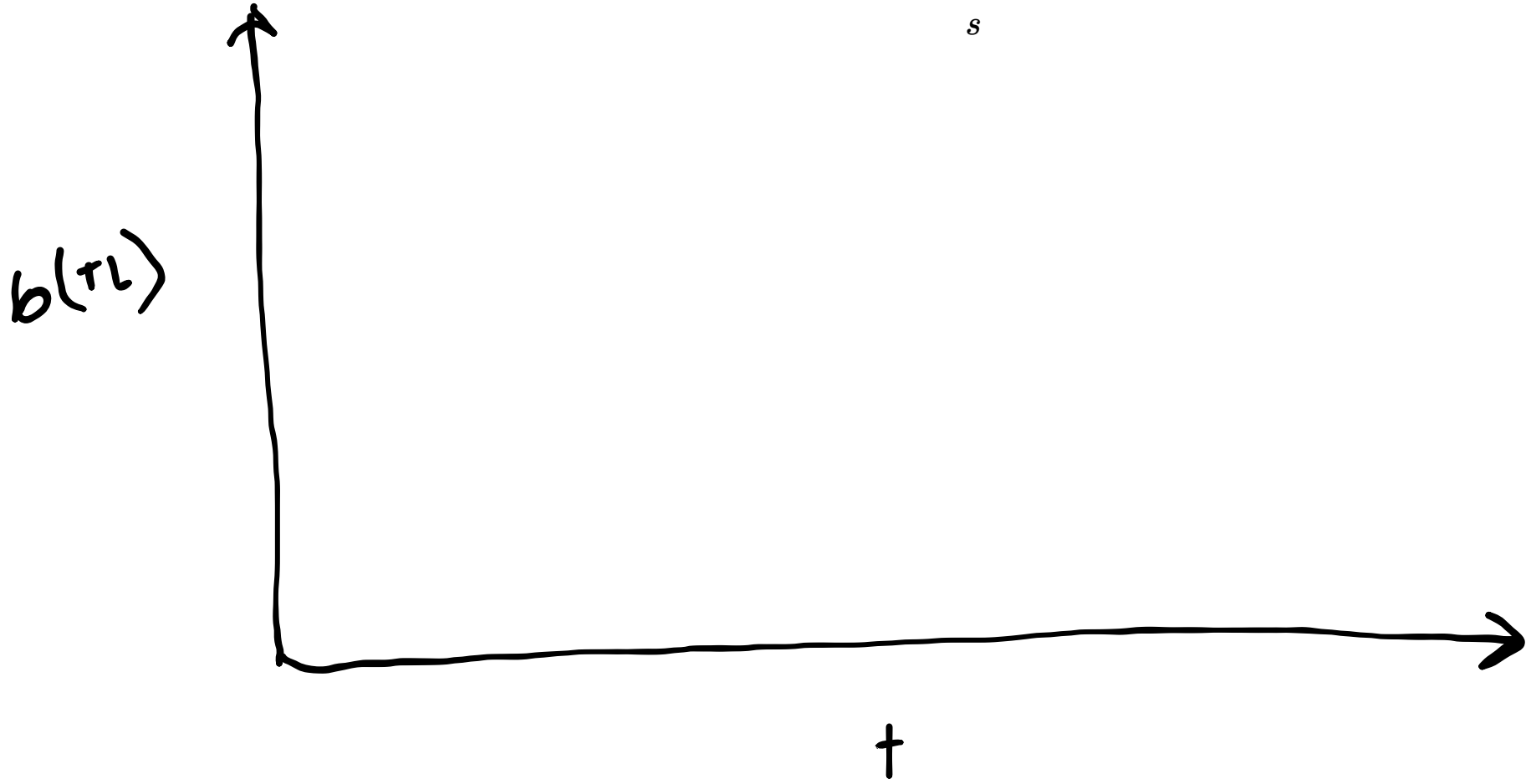
$$b'(s') \propto Z(o \mid a, s') \sum_{s} T(s' \mid s, a)\, b(s)$$

$b(\text{TL})$

†

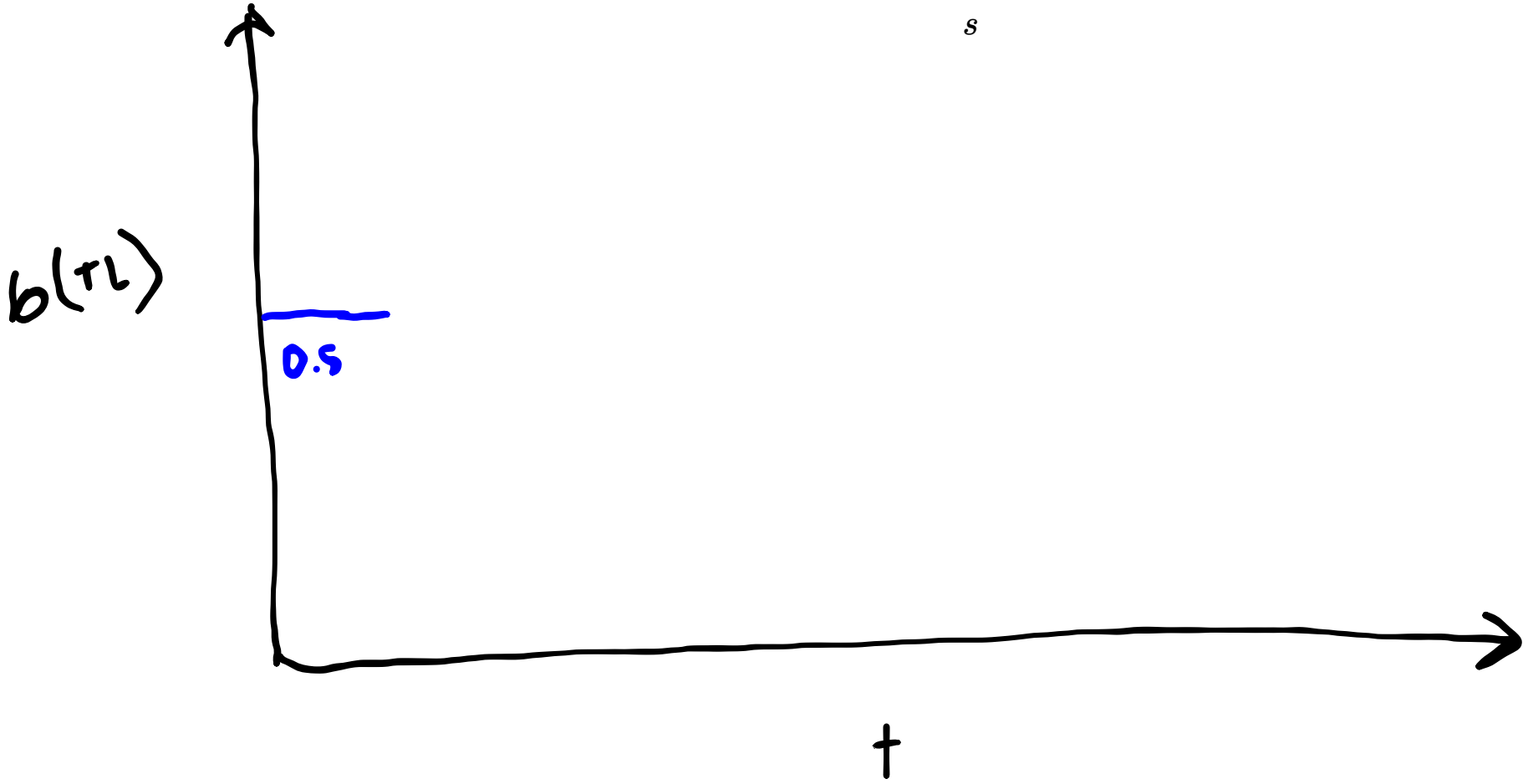
Belief Dynamics

$$b'(s') \propto Z(o \mid a, s') \sum_{s} T(s' \mid s, a)\, b(s)$$



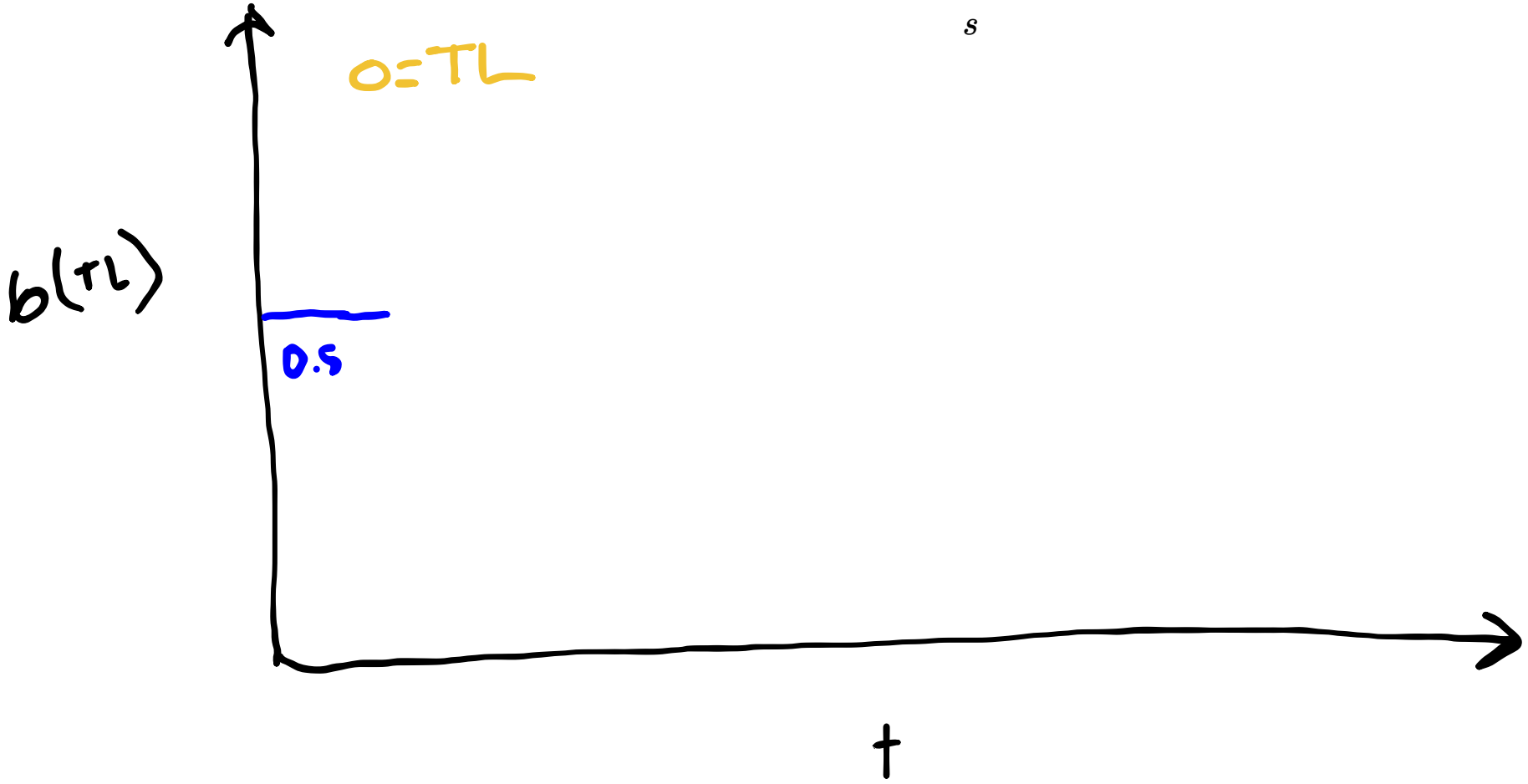
Belief Dynamics

$$b'(s') \propto Z(o \mid a, s') \sum_{s} T(s' \mid s, a)\, b(s)$$



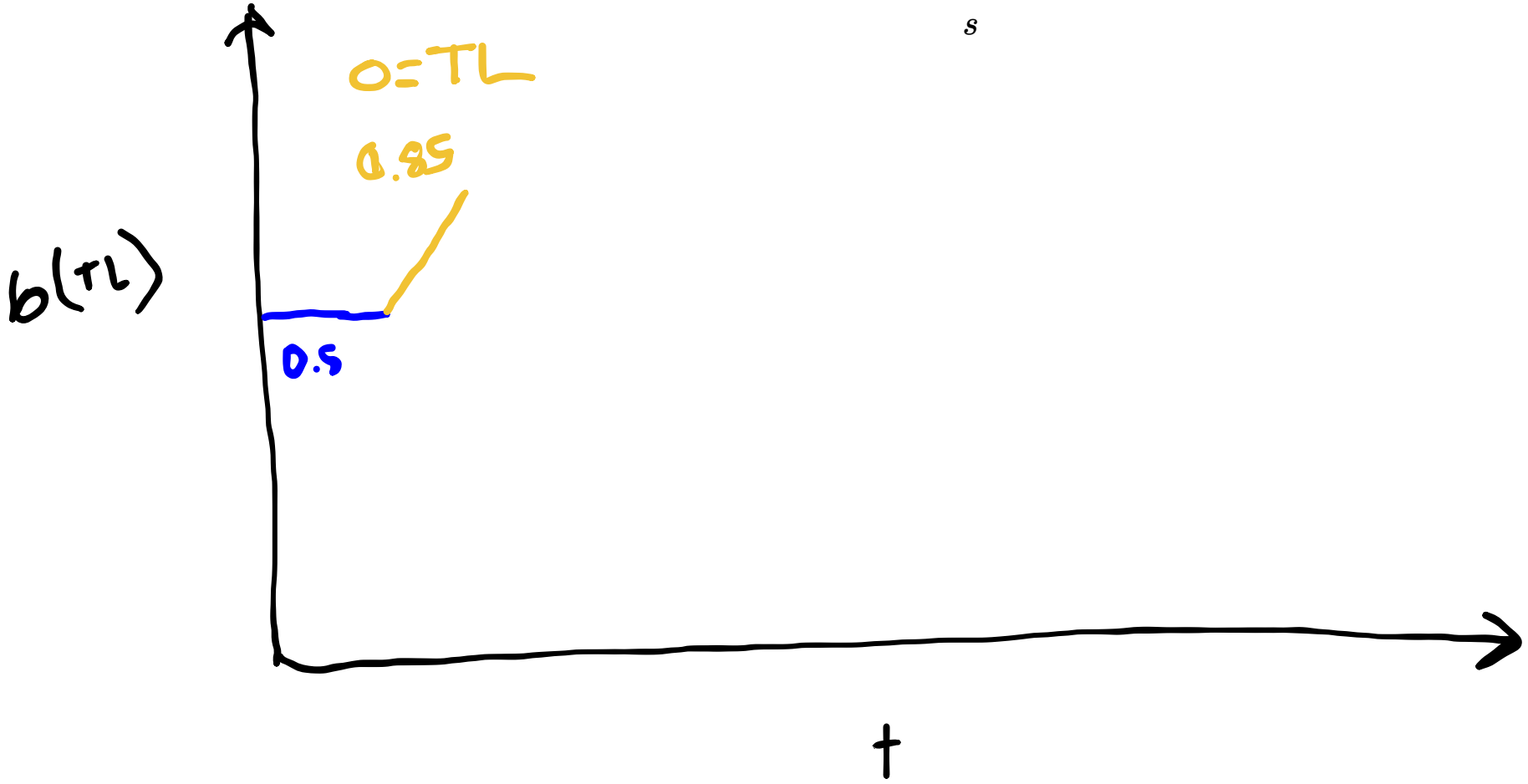
Belief Dynamics

$$b'(s') \propto Z(o \mid a, s') \sum_{s} T(s' \mid s, a)\, b(s)$$



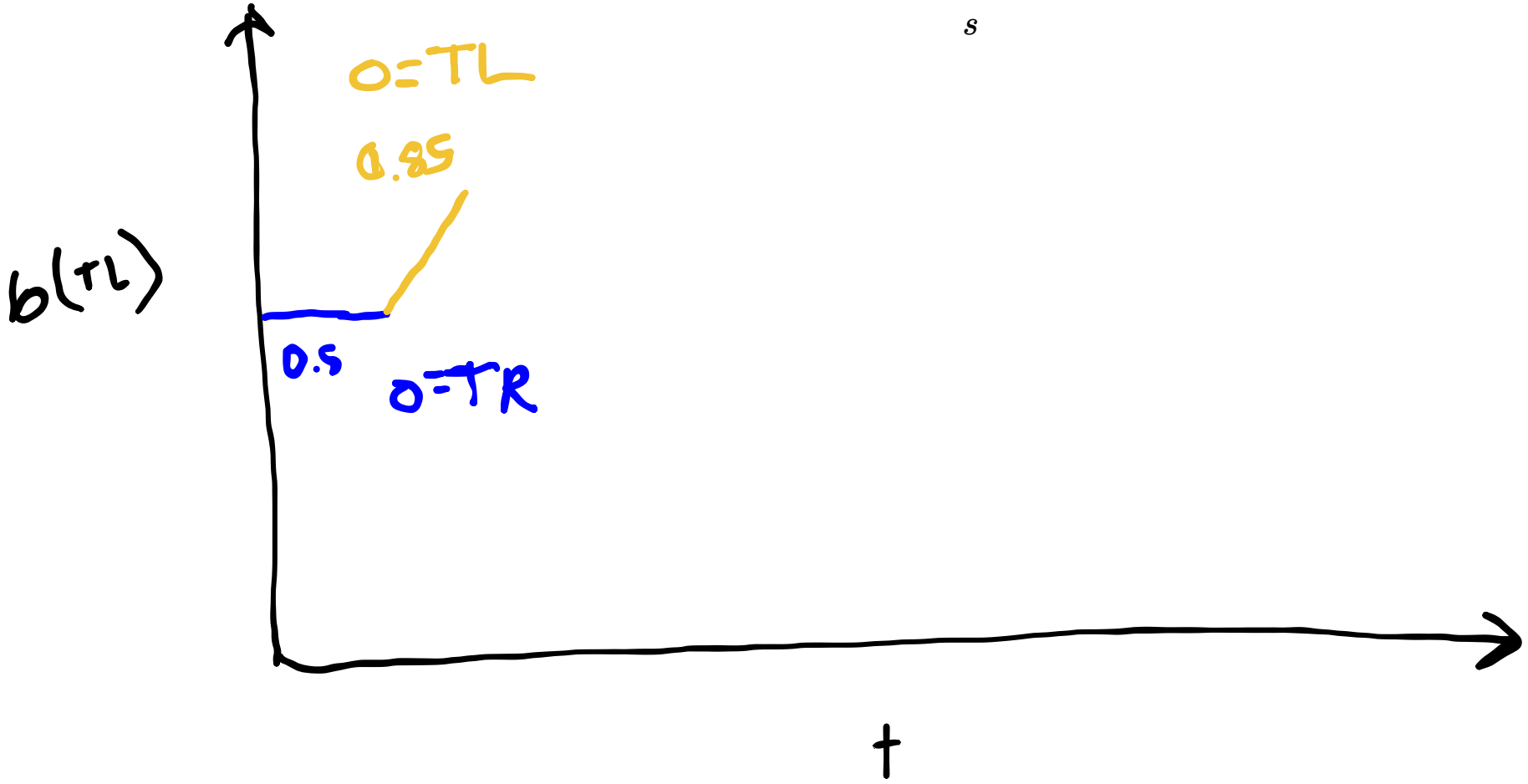
Belief Dynamics

$$b'(s') \propto Z(o \mid a, s') \sum_{s} T(s' \mid s, a)\, b(s)$$



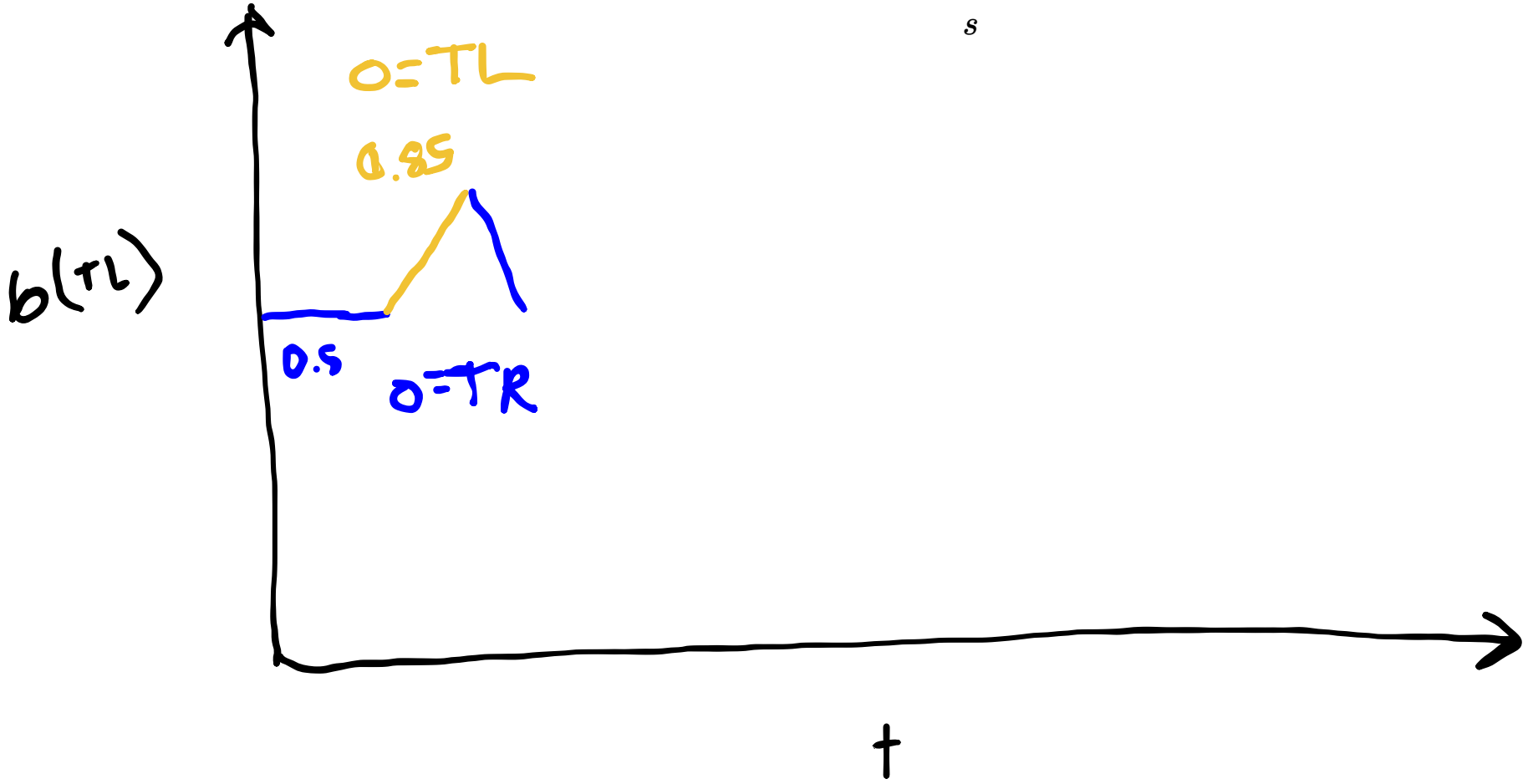
Belief Dynamics

$$b'(s') \propto Z(o \mid a, s') \sum_{s} T(s' \mid s, a)\, b(s)$$



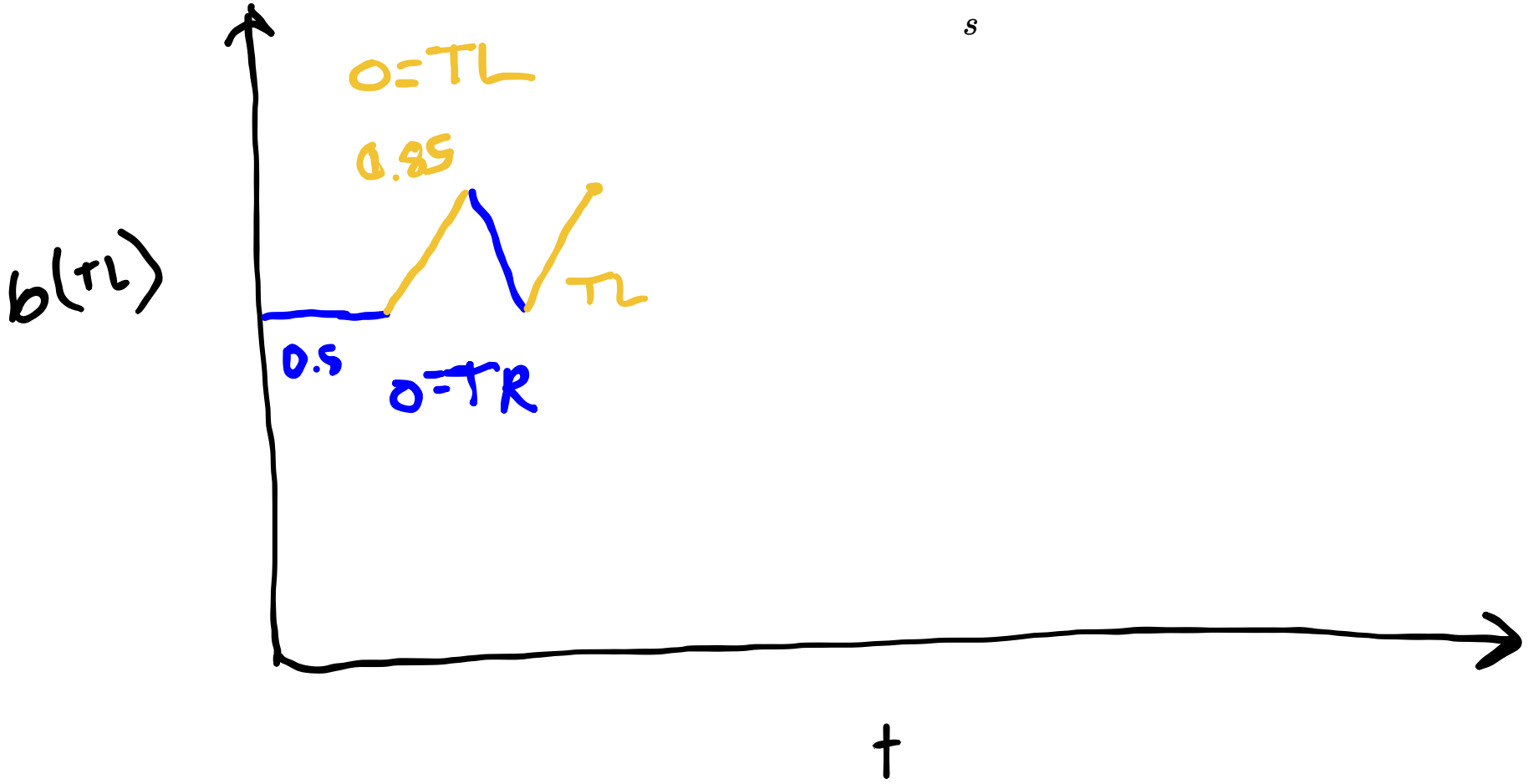
Belief Dynamics

$$b'(s') \propto Z(o \mid a, s') \sum_{s} T(s' \mid s, a)\, b(s)$$



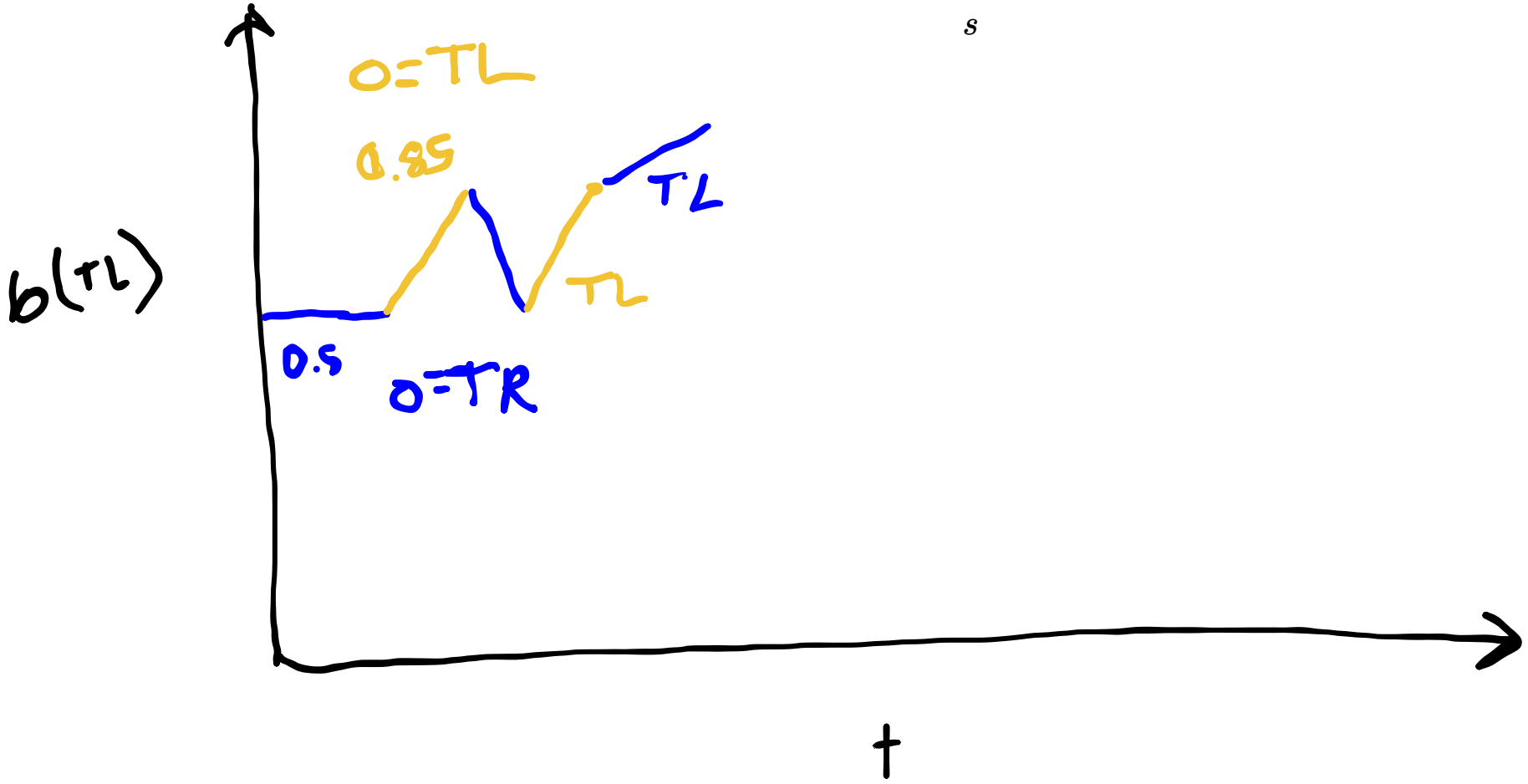
Belief Dynamics

$$b'(s') \propto Z(o \mid a, s') \sum_{s} T(s' \mid s, a)\, b(s)$$



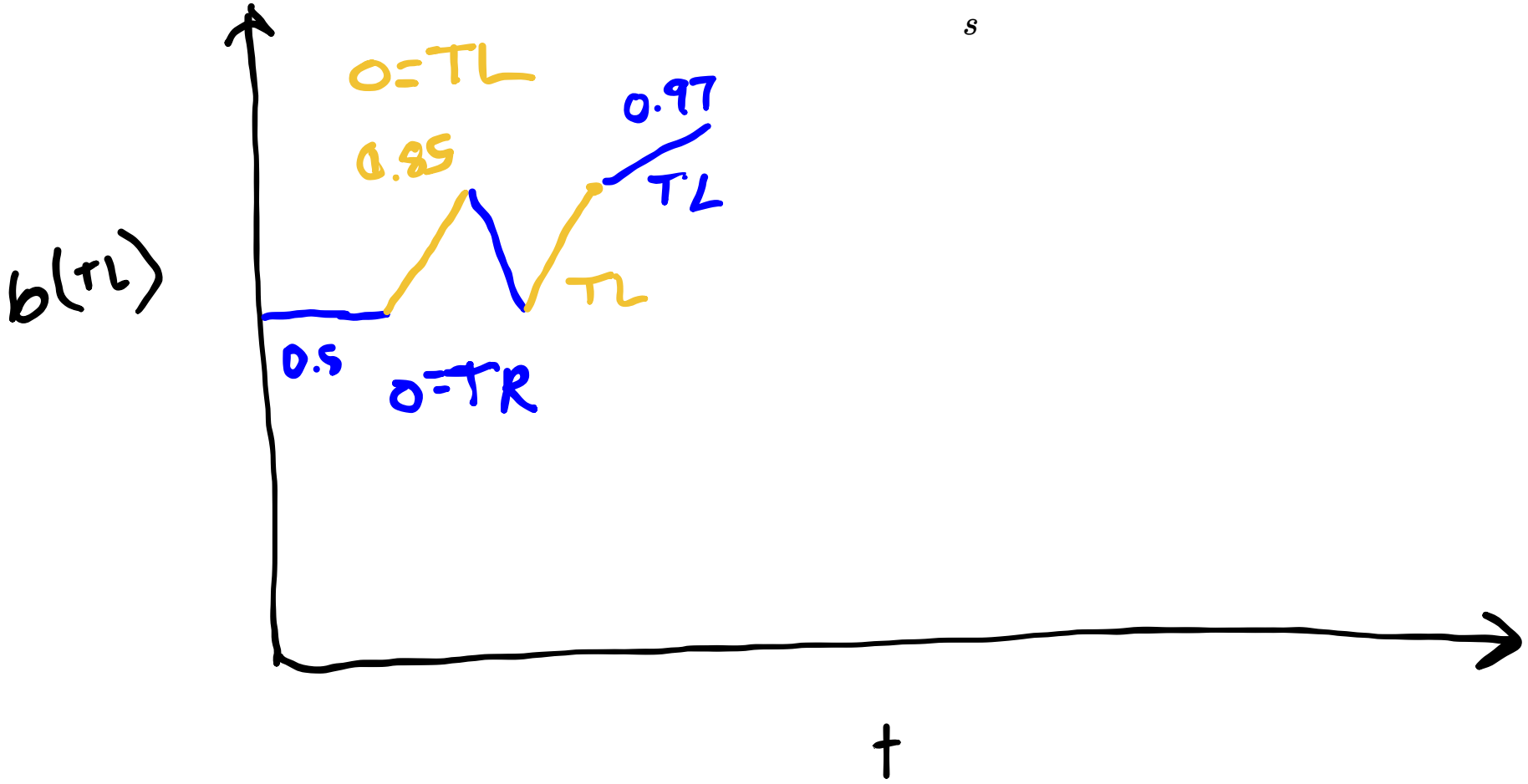
Belief Dynamics

$$b'(s') \propto Z(o \mid a, s') \sum_{s} T(s' \mid s, a)\, b(s)$$



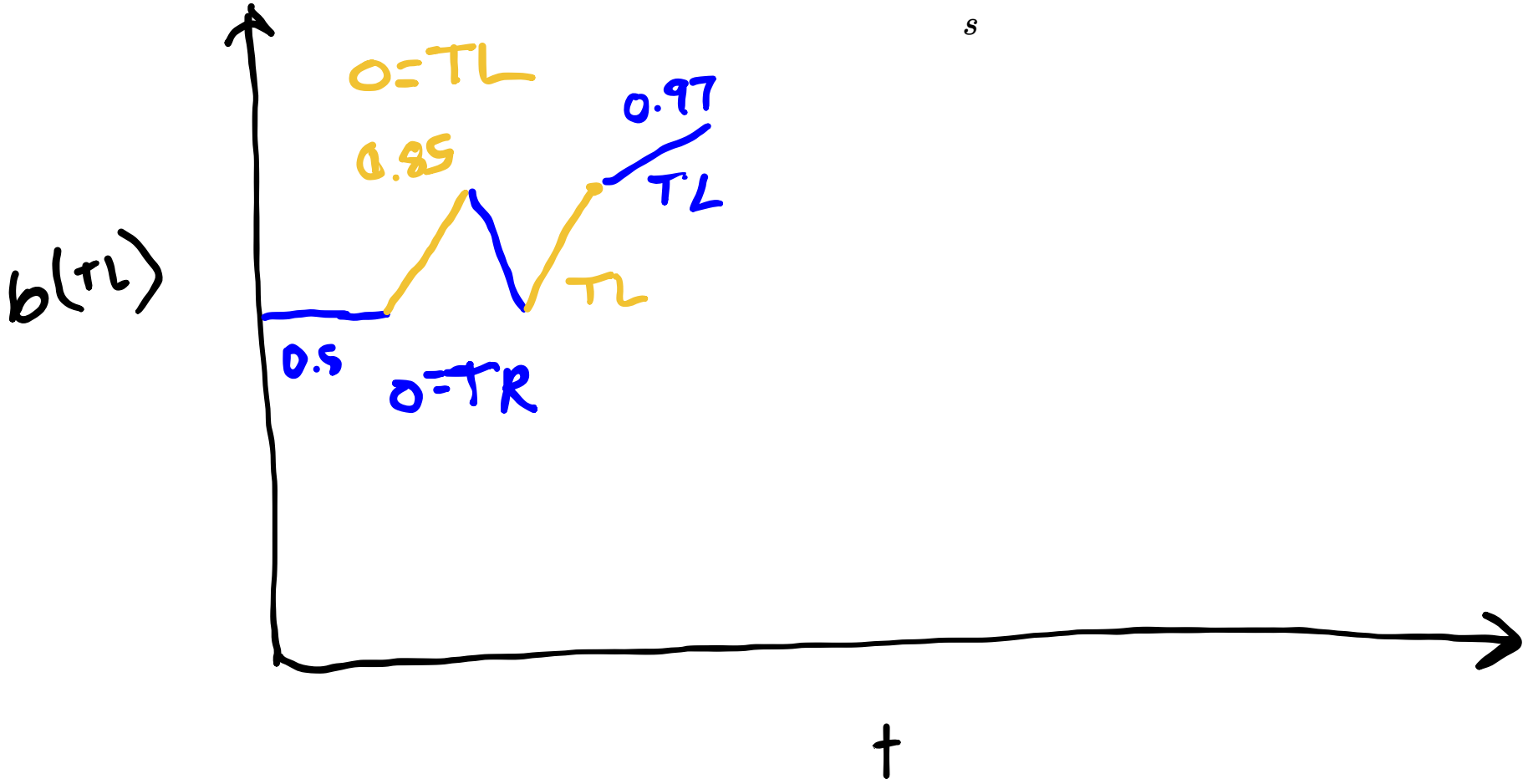
Belief Dynamics

$$b'(s') \propto Z(o \mid a, s') \sum_{s} T(s' \mid s, a)\, b(s)$$



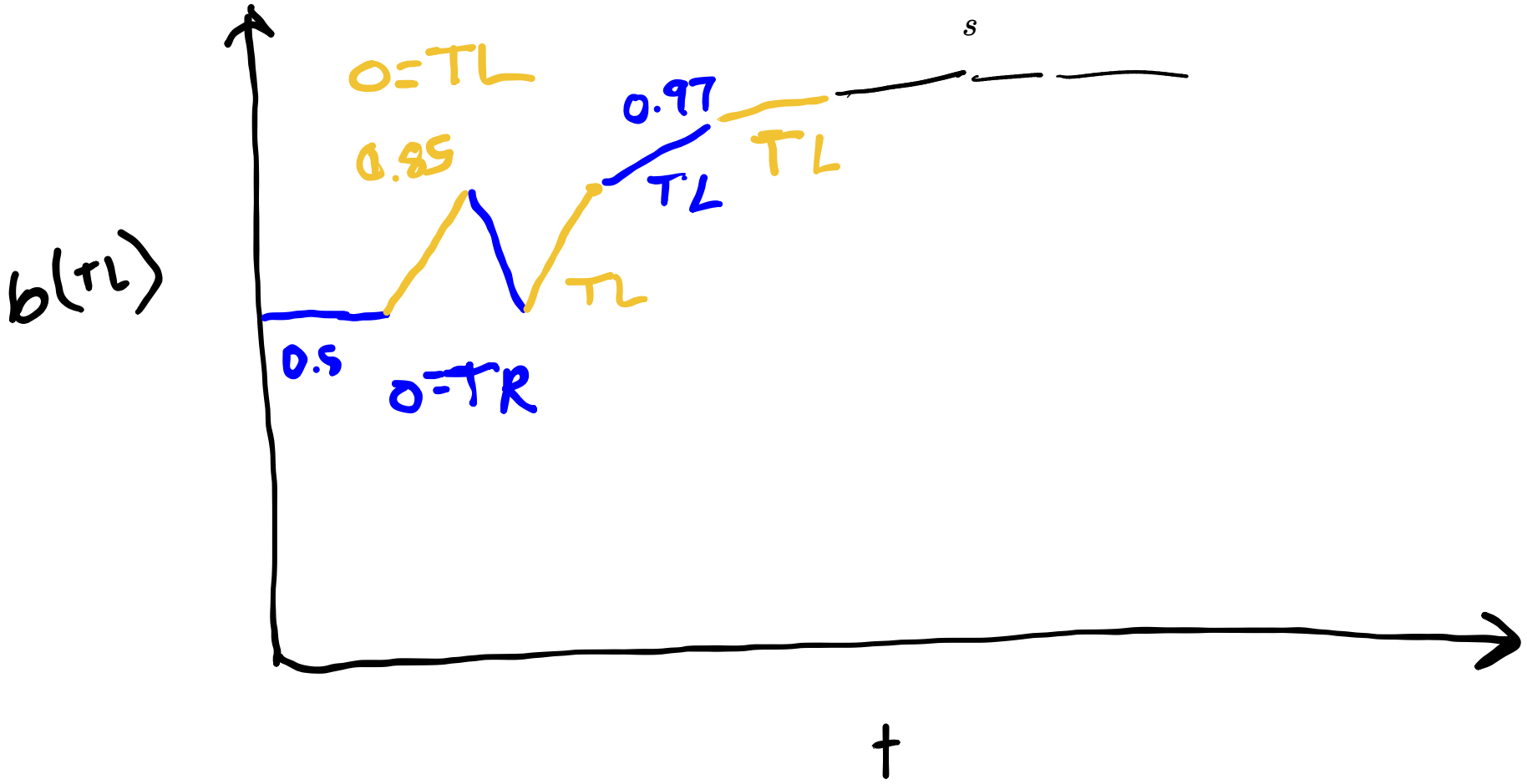
Belief Dynamics

$$b'(s') \propto Z(o \mid a, s') \sum_{s} T(s' \mid s, a)\, b(s)$$

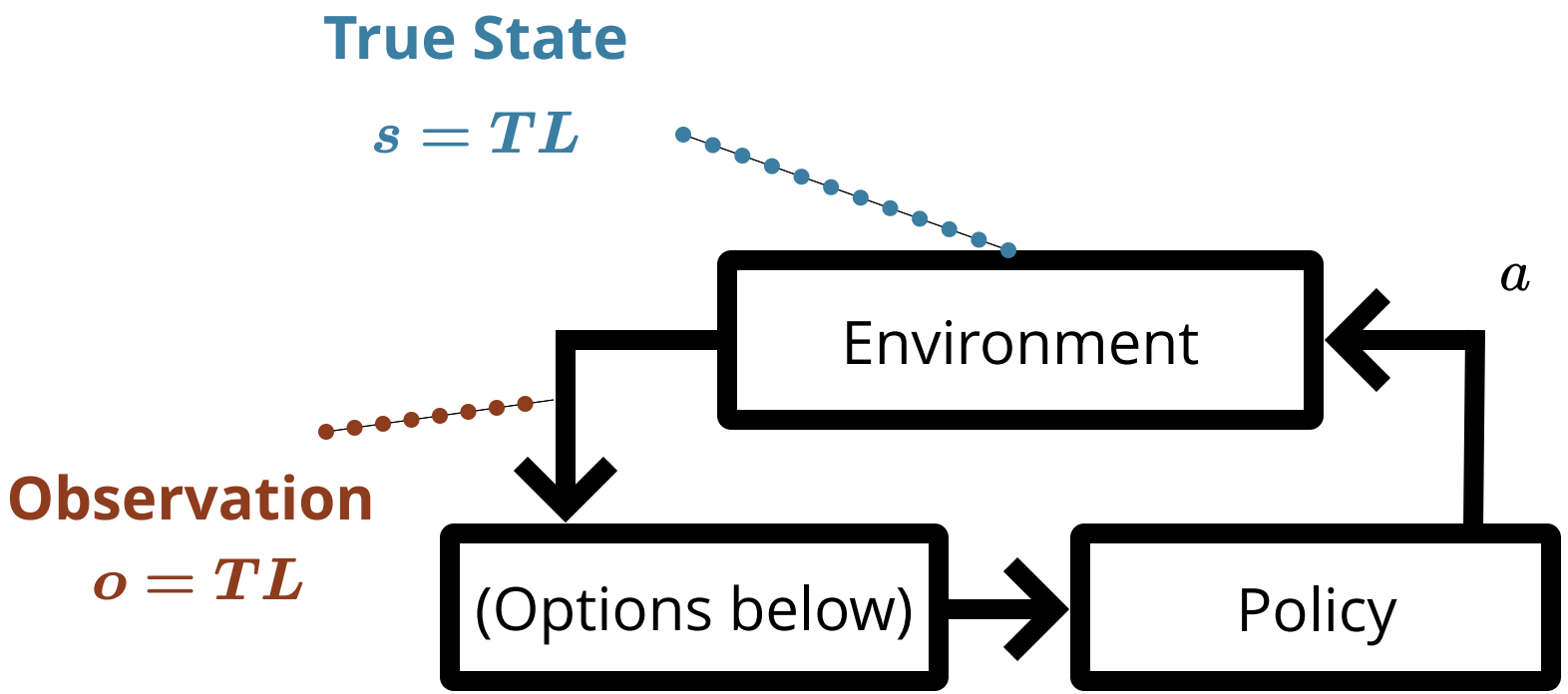


Belief Dynamics

$$b'(s') \propto Z(o \mid a, s') \sum_{s} T(s' \mid s, a)\, b(s)$$

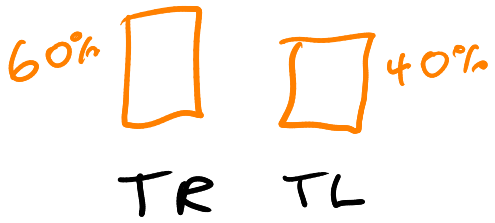


POMDP Sense-Plan-Act Loop



Guiding Quesiton

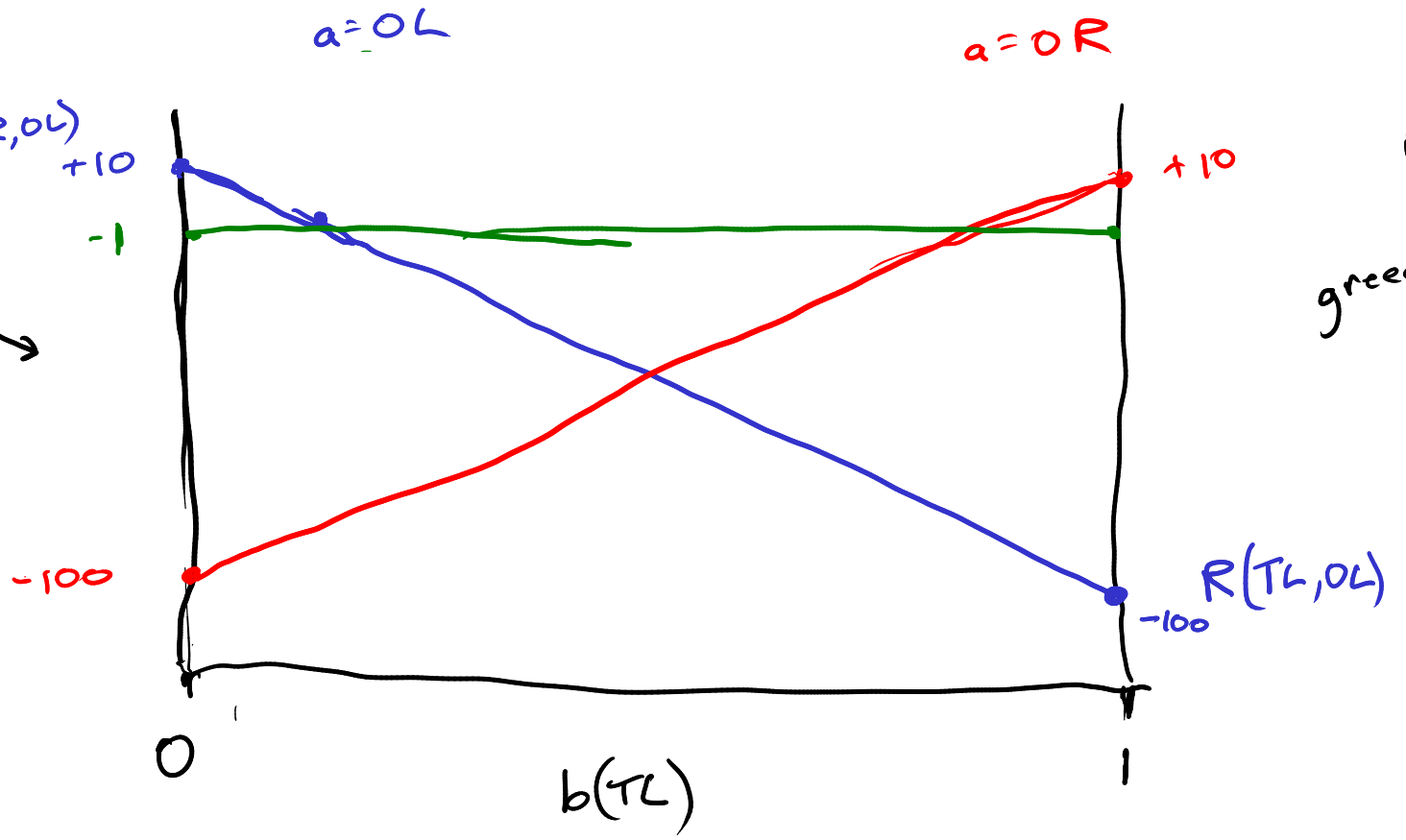
How do we calculate the optimal action in a POMDP?



One-step utility

Reward
+10 empty
-1 Listen
-100 tiger

$R(b,a) = E_{s\sim b}[R(s,a)]$



$R(b,a) = \bar{r}^a \cdot b$

greedy action
$\operatorname{argmax}_a \bar{r}^a \cdot b$

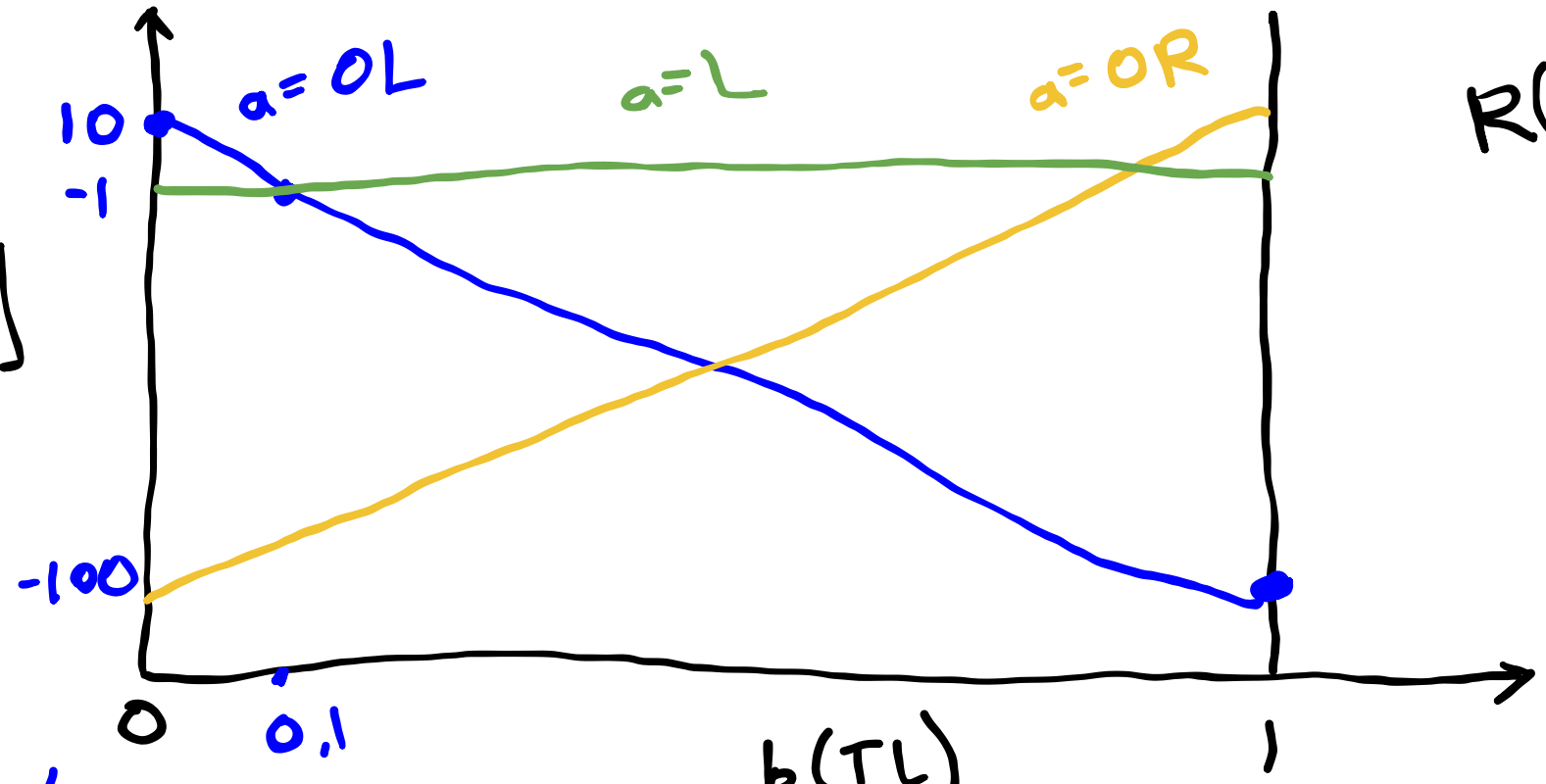
1-step
α-vector

$R(b, a{=}OL) = +10\,(1 - b(TL)) + (-100)\,b(TL)$

$b(TL) = 0.1 \;:\; 10 \cdot 0.9 + -100 \cdot 0.1$ (9 - 10)

One-step utility

Reward: +10 empty door
−1 Listen
−100 Tiger



$R(b,a) = E_{s\sim b}[R(s,a)]$

$R(b,a) = \bar{r}_a \cdot b$

α-vector

$R(b,a) = 10 \cdot (1 - b(TL)) - 100\, b(TL) = -110\, b + 10$

Exercise 2: Crying Baby 1-Step Utility

Exercise 2: Crying Baby 1-Step Utility

$S = \{h, \neg h\}$

$A = \{f, \neg f\}$

$O = \{c, \neg c\}$

Exercise 2: Crying Baby 1-Step Utility

$S = \{h, \neg h\}$

$A = \{f, \neg f\}$

$O = \{c, \neg c\}$

$R(s, a) = R(s) + R(a)$

Exercise 2: Crying Baby 1-Step Utility

$S = \{h, \neg h\}$

$A = \{f, \neg f\}$

$O = \{c, \neg c\}$

$$R(s, a) = R(s) + R(a)$$

$$R(s) = \begin{cases} -10 \text{ if } s = h \\ 0 \text{ otherwise} \end{cases}$$

Exercise 2: Crying Baby 1-Step Utility

$S = \{h, \neg h\}$

$A = \{f, \neg f\}$

$O = \{c, \neg c\}$

$$R(s, a) = R(s) + R(a)$$

$$R(s) = \begin{cases} -10 \text{ if } s = h \\ 0 \text{ otherwise} \end{cases}$$

$$R(a) = \begin{cases} -5 \text{ if } a = f \\ 0 \text{ otherwise} \end{cases}$$

Exercise 2: Crying Baby 1-Step Utility

$S = \{h, \neg h\}$

$A = \{f, \neg f\}$

$O = \{c, \neg c\}$

$T(h \mid h, \neg f) = 1.0$

$T(h \mid \neg h, \neg f) = 0.1$

$T(\neg h \mid \cdot, f) = 1.0$

$$R(s, a) = R(s) + R(a)$$

$$R(s) = \begin{cases} -10 \text{ if } s = h \\ 0 \text{ otherwise} \end{cases}$$

$$R(a) = \begin{cases} -5 \text{ if } a = f \\ 0 \text{ otherwise} \end{cases}$$

Exercise 2: Crying Baby 1-Step Utility

$S = \{h, \neg h\}$

$A = \{f, \neg f\}$

$O = \{c, \neg c\}$

$T(h \mid h, \neg f) = 1.0$

$T(h \mid \neg h, \neg f) = 0.1$

$T(\neg h \mid \cdot, f) = 1.0$

$$R(s, a) = R(s) + R(a)$$

$$R(s) = \begin{cases} -10 \text{ if } s = h \\ 0 \text{ otherwise} \end{cases}$$

$$R(a) = \begin{cases} -5 \text{ if } a = f \\ 0 \text{ otherwise} \end{cases}$$

$Z(c \mid \cdot, h) = 0.8)$

$Z(c \mid \cdot, \neg h) = 0.1$

Exercise 2: Crying Baby 1-Step Utility

$$S = \{h, \neg h\} \qquad T(h \mid h, \neg f) = 1.0$$
$$A = \{f, \neg f\} \qquad T(h \mid \neg h, \neg f) = 0.1$$
$$O = \{c, \neg c\} \qquad T(\neg h \mid \cdot, f) = 1.0$$

$$R(s, a) = R(s) + R(a)$$

$$R(s) = \begin{cases} -10 \text{ if } s = h \\ 0 \text{ otherwise} \end{cases}$$

$$R(a) = \begin{cases} -5 \text{ if } a = f \\ 0 \text{ otherwise} \end{cases}$$

$$Z(c \mid \cdot, h) = 0.8)$$
$$Z(c \mid \cdot, \neg h) = 0.1$$
$$\gamma = 0.9$$

Exercise 2: Crying Baby 1-Step Utility

$S = \{h, \neg h\}$

$A = \{f, \neg f\}$

$O = \{c, \neg c\}$

$T(h \mid h, \neg f) = 1.0$

$T(h \mid \neg h, \neg f) = 0.1$

$T(\neg h \mid \cdot, f) = 1.0$

$$R(s, a) = R(s) + R(a)$$

$$R(s) = \begin{cases} -10 \text{ if } s = h \\ 0 \text{ otherwise} \end{cases}$$

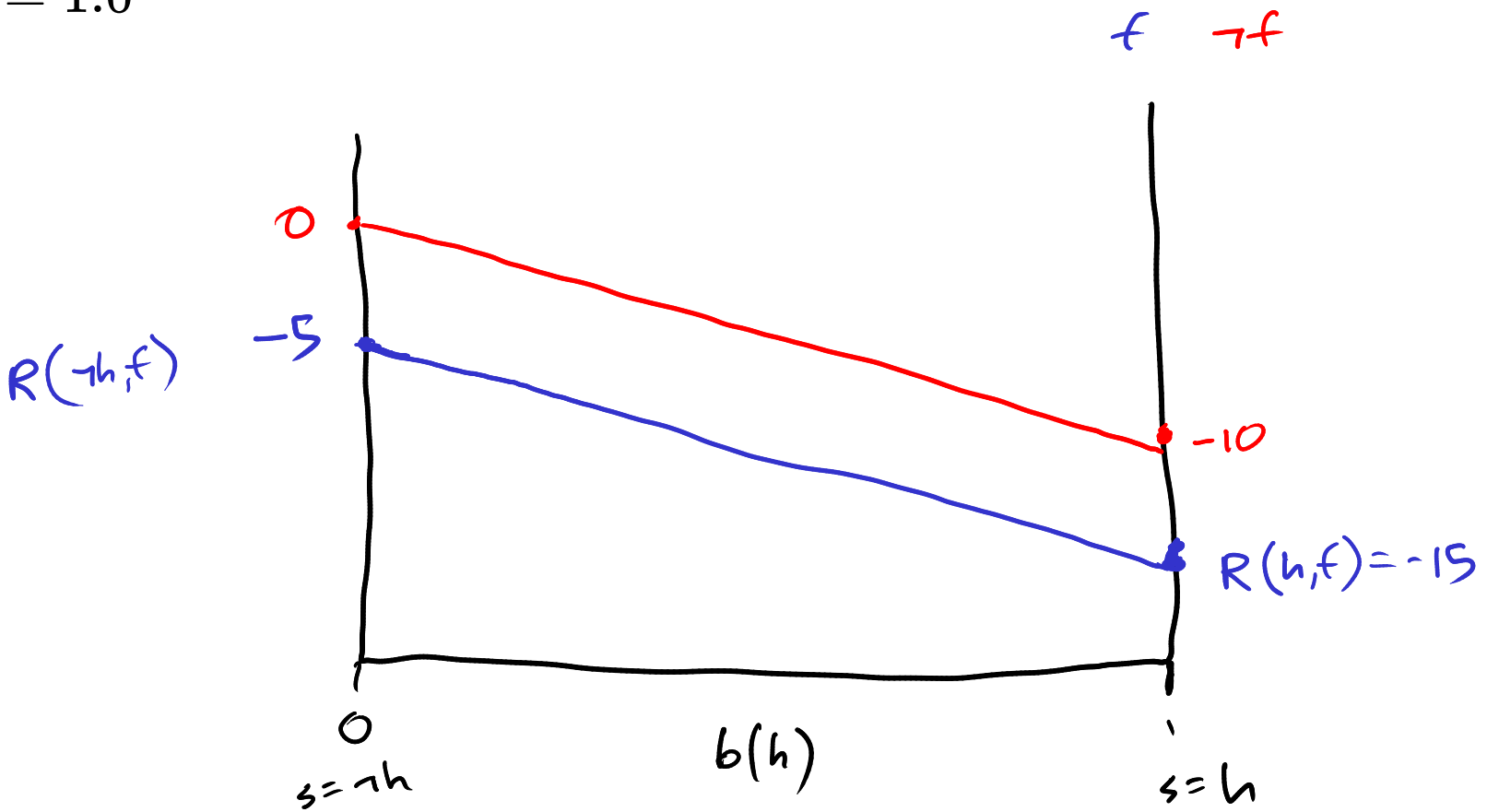
$$R(a) = \begin{cases} -5 \text{ if } a = f \\ 0 \text{ otherwise} \end{cases}$$

$$Z(c \mid \cdot, h) = 0.8)$$

$$Z(c \mid \cdot, \neg h) = 0.1$$

$$\gamma = 0.9$$

Draw the 1-step utility α-vectors for the Crying Baby problem.



Alpha Vectors for Conditional Plans

Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

1 Step:

Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

1 Step: 

Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

1 Step: 

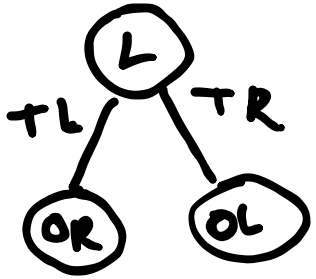
2 Step:

Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

1 Step:

2 Step:

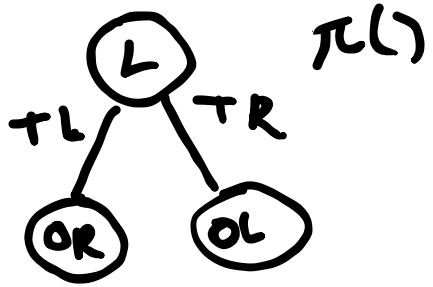


Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

1 Step:

2 Step:

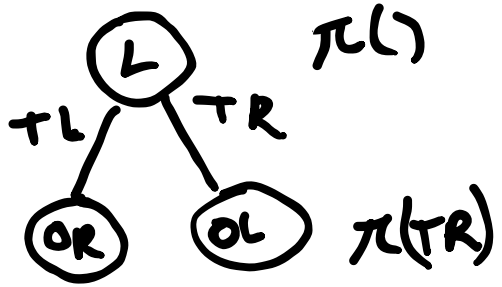


Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

1 Step:

2 Step:

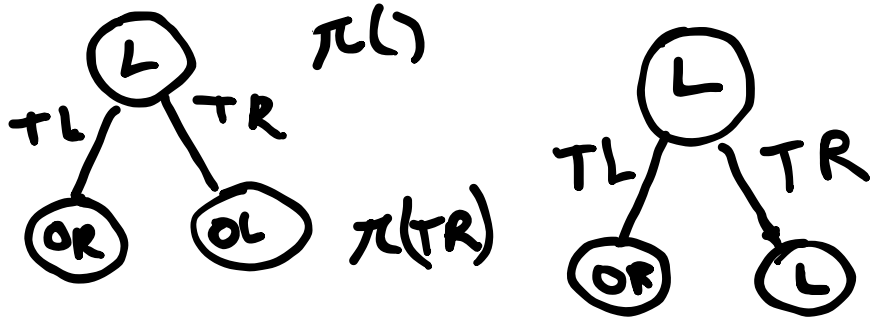


Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

1 Step:

2 Step:

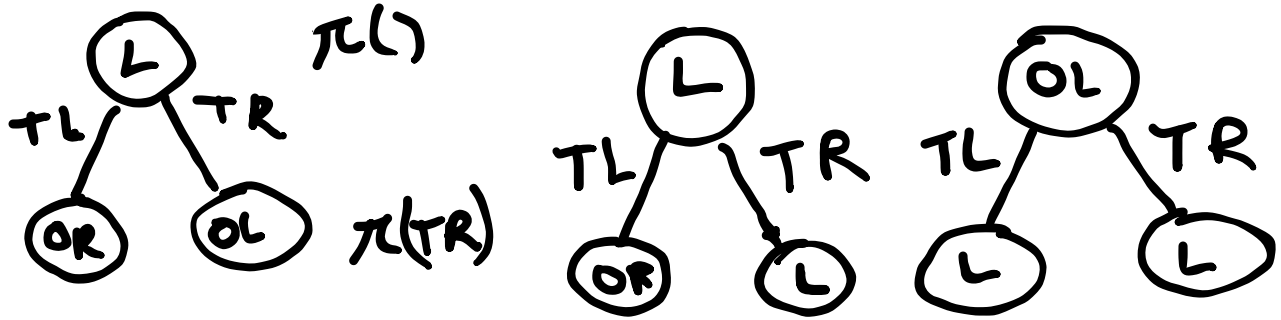


Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

1 Step:

2 Step:

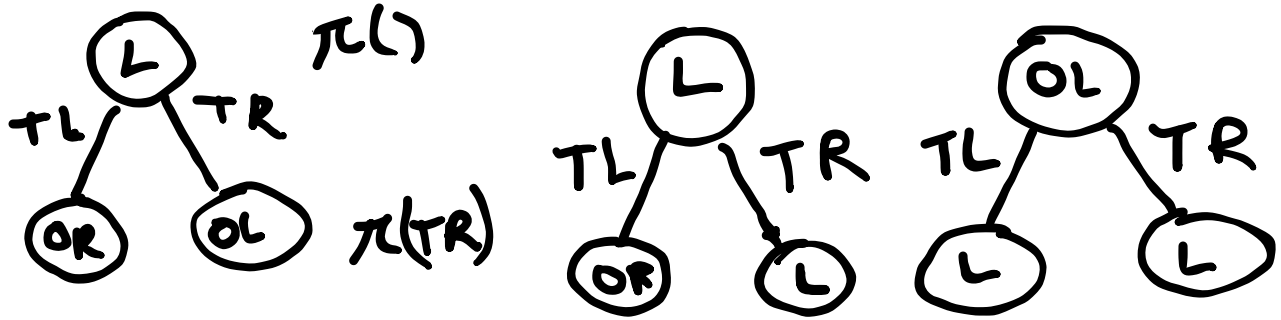


Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

1 Step:

2 Step:



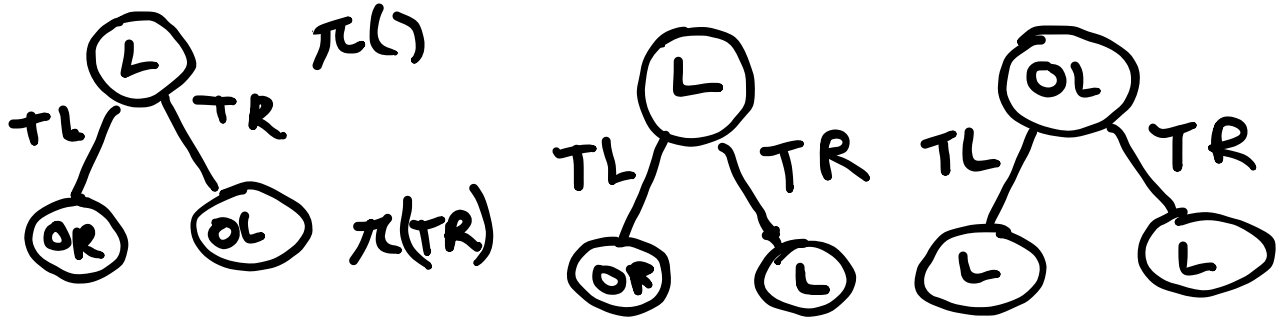
$$|A|^{\frac{(|O|^h-1)}{(|O|-1)}}$$

Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

1 Step:

2 Step:



$|A|^{\frac{(|O|^h-1)}{(|O|-1)}}$

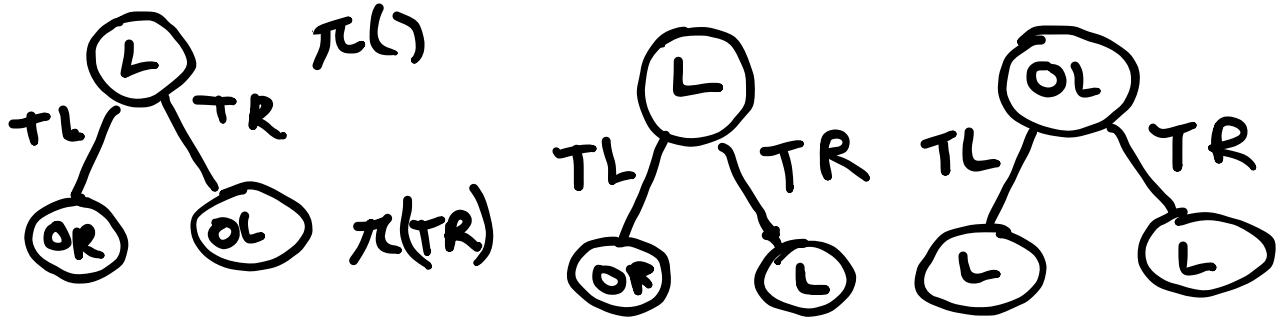
27 two step plans!

Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

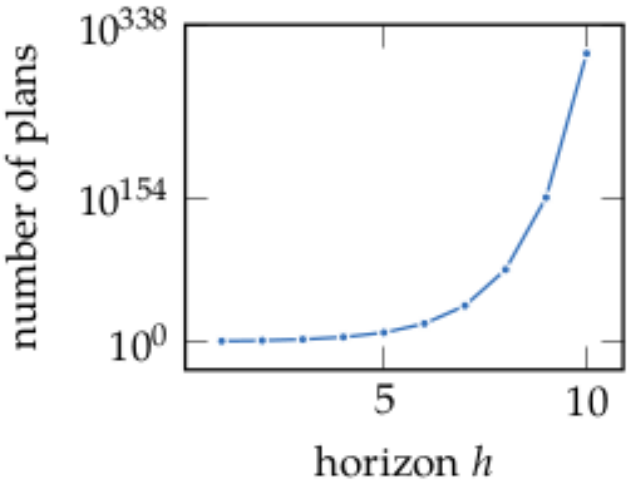
1 Step:

2 Step:



$$|A|^{\frac{(|O|^h-1)}{(|O|-1)}}$$

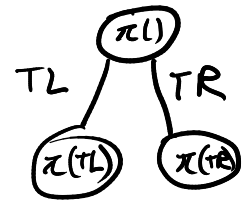
27 two step plans!



Alpha Vectors for Conditional Plans

Alpha Vectors for Conditional Plans

$$U^{\pi}(s) = R(s, \pi()) + \gamma \left[\sum_{s'} T(s' \mid s, \pi()) \sum_{o} O(o \mid \pi(), s') U^{\pi(o)}(s') \right]$$

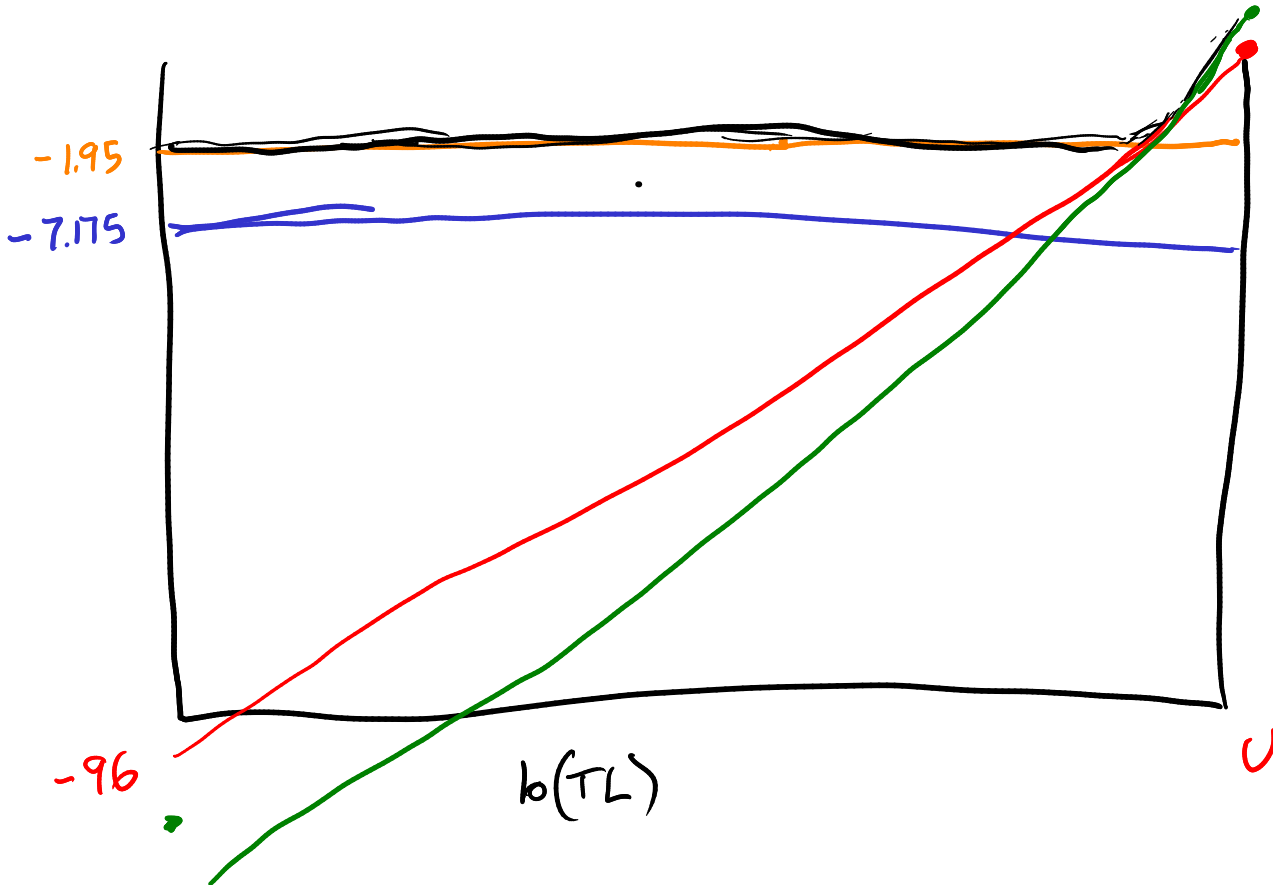


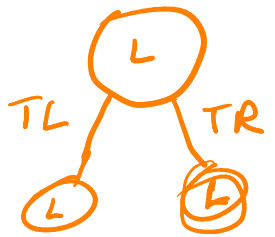
Alpha Vectors for Conditional Plans

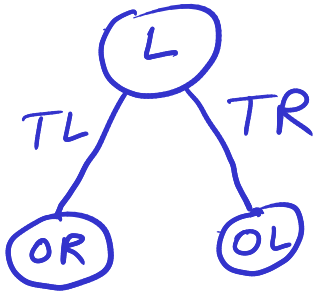
$$U^{\pi}(s) = R(s, \pi()) + \gamma \left[\sum_{s'} T(s' \mid s, \pi()) \sum_{o} O(o \mid \pi(), s') U^{\pi(o)}(s') \right]$$

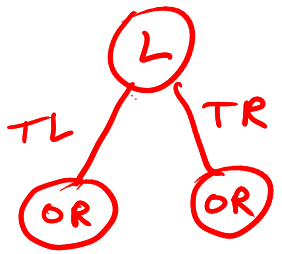
$U^*(b) = \max_{\pi} \alpha^{\pi} \cdot b$

For 1-step: $U^{\pi}(s) = R(s, \pi())$

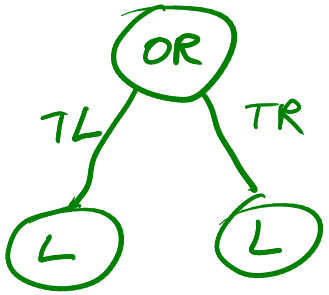








$U^{\pi}(TL) = -1 + \gamma 10$
$= 8.5$
$U^{\pi}(TR) = -1 + \gamma -100$
-96



$U^{\pi}(TL) = 10 + \gamma(-1)$
$= 9.5$
$U^{\pi}(TR) = -100 + \gamma(-1)$
$= -100.95$

POMDP Value Functions

POMDP Value Functions

$$V^*(b) = \max_{\alpha \in \Gamma} \alpha^\top b$$

Exercise: 2-Step Crying Baby α Vectors

Exercise: 2-Step Crying Baby α Vectors

$$S = \{h, \neg h\}$$

$$A = \{f, \neg f\}$$

$$O = \{c, \neg c\}$$

Exercise: 2-Step Crying Baby α Vectors

$$S = \{h, \neg h\}$$

$$A = \{f, \neg f\}$$

$$O = \{c, \neg c\}$$

$$R(s, a) = R(s) + R(a)$$

Exercise: 2-Step Crying Baby α Vectors

$$S = \{h, \neg h\}$$

$$A = \{f, \neg f\}$$

$$O = \{c, \neg c\}$$

$$R(s, a) = R(s) + R(a)$$

$$R(s) = \begin{cases} -10 \text{ if } s = h \\ 0 \text{ otherwise} \end{cases}$$

Exercise: 2-Step Crying Baby α Vectors

$$S = \{h, \neg h\}$$

$$A = \{f, \neg f\}$$

$$O = \{c, \neg c\}$$

$$R(s, a) = R(s) + R(a)$$

$$R(s) = \begin{cases} -10 \text{ if } s = h \\ 0 \text{ otherwise} \end{cases}$$

$$R(a) = \begin{cases} -5 \text{ if } a = f \\ 0 \text{ otherwise} \end{cases}$$

Exercise: 2-Step Crying Baby α Vectors

$$S = \{h, \neg h\} \qquad T(h \mid h, \neg f) = 1.0$$
$$A = \{f, \neg f\} \qquad T(h \mid \neg h, \neg f) = 0.1$$
$$O = \{c, \neg c\} \qquad T(\neg h \mid \cdot, f) = 1.0$$

$$R(s, a) = R(s) + R(a)$$

$$R(s) = \begin{cases} -10 \text{ if } s = h \\ 0 \text{ otherwise} \end{cases}$$

$$R(a) = \begin{cases} -5 \text{ if } a = f \\ 0 \text{ otherwise} \end{cases}$$

Exercise: 2-Step Crying Baby α Vectors

$$S = \{h, \neg h\} \qquad T(h \mid h, \neg f) = 1.0$$
$$A = \{f, \neg f\} \qquad T(h \mid \neg h, \neg f) = 0.1$$
$$O = \{c, \neg c\} \qquad T(\neg h \mid \cdot, f) = 1.0$$

$$R(s, a) = R(s) + R(a)$$

$$R(s) = \begin{cases} -10 \text{ if } s = h \\ 0 \text{ otherwise} \end{cases}$$

$$R(a) = \begin{cases} -5 \text{ if } a = f \\ 0 \text{ otherwise} \end{cases}$$

$$Z(c \mid \cdot, h) = 0.8)$$
$$Z(c \mid \cdot, \neg h) = 0.1$$

Exercise: 2-Step Crying Baby α Vectors

$$S = \{h, \neg h\} \qquad T(h \mid h, \neg f) = 1.0$$
$$A = \{f, \neg f\} \qquad T(h \mid \neg h, \neg f) = 0.1$$
$$O = \{c, \neg c\} \qquad T(\neg h \mid \cdot, f) = 1.0$$

$$R(s, a) = R(s) + R(a)$$

$$R(s) = \begin{cases} -10 \text{ if } s = h \\ 0 \text{ otherwise} \end{cases}$$

$$R(a) = \begin{cases} -5 \text{ if } a = f \\ 0 \text{ otherwise} \end{cases}$$

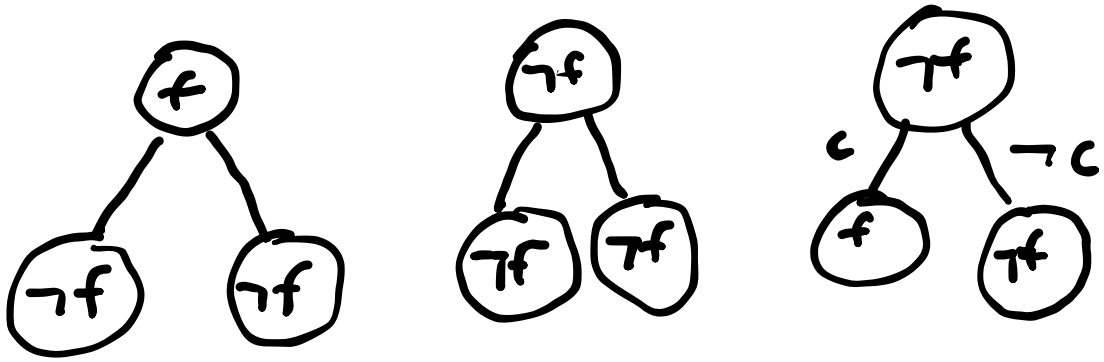
$$Z(c \mid \cdot, h) = 0.8)$$
$$Z(c \mid \cdot, \neg h) = 0.1$$

$$\gamma = 0.9$$

Exercise: 2-Step Crying Baby α Vectors

$$S = \{h, \neg h\}$$
$$A = \{f, \neg f\}$$
$$O = \{c, \neg c\}$$

$$T(h \mid h, \neg f) = 1.0$$
$$T(h \mid \neg h, \neg f) = 0.1$$
$$T(\neg h \mid \cdot, f) = 1.0$$



$$R(s, a) = R(s) + R(a)$$

$$R(s) = \begin{cases} -10 \text{ if } s = h \\ 0 \text{ otherwise} \end{cases}$$

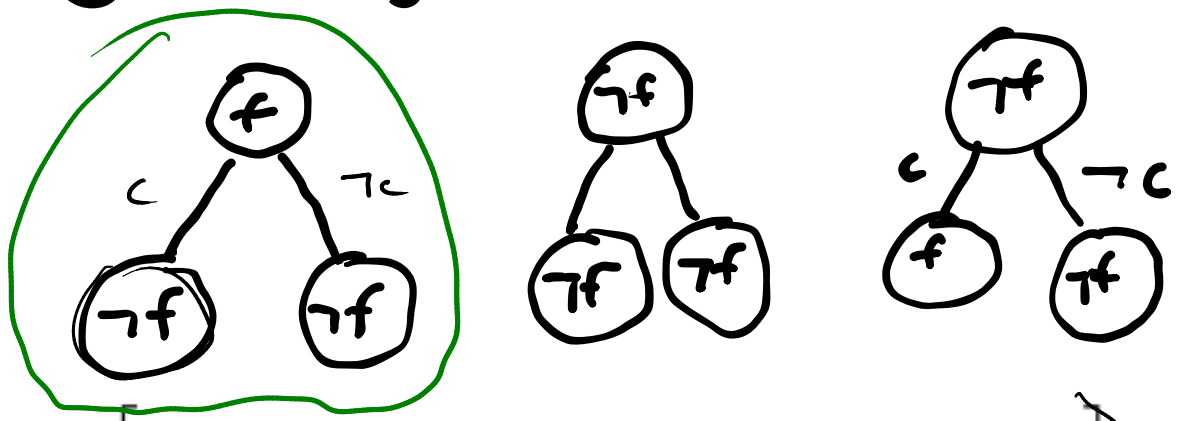
$$R(a) = \begin{cases} -5 \text{ if } a = f \\ 0 \text{ otherwise} \end{cases}$$

$$Z(c \mid \cdot, h) = 0.8)$$
$$Z(c \mid \cdot, \neg h) = 0.1$$

$$\gamma = 0.9$$

Exercise: 2-Step Crying Baby α Vectors

$$S = \{h, \neg h\} \qquad T(h \mid h, \neg f) = 1.0$$
$$A = \{f, \neg f\} \qquad T(h \mid \neg h, \neg f) = 0.1$$
$$O = \{c, \neg c\} \qquad T(\neg h \mid \cdot, f) = 1.0$$



$$R(s,a) = R(s) + R(a)$$

$$R(s) = \begin{cases} -10 \text{ if } s = h \\ 0 \text{ otherwise} \end{cases}$$

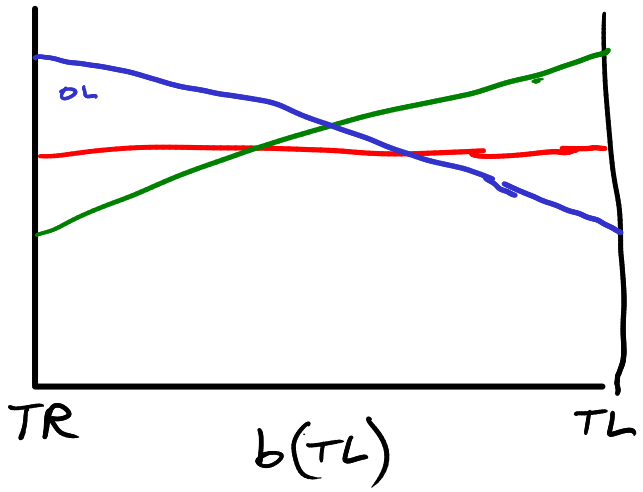
$$R(a) = \begin{cases} -5 \text{ if } a = f \\ 0 \text{ otherwise} \end{cases}$$

$$Z(c \mid \cdot, h) = 0.8)$$
$$Z(c \mid \cdot, \neg h) = 0.1$$
$$\gamma = 0.9$$

$$U^{\pi}(s) = R(s, \pi()) + \gamma \left[\sum_{s'} T(s' \mid s, \pi()) \sum_{o} O(o \mid \pi(), s') U^{\pi(o)}(s') \right]$$

$U^{\pi(c)}(h), U^{\pi(\neg c)}(h)$

$U^{\pi}(h) = R(h, f)$
$+ \gamma \left(T(h \mid h, f) \left(Z(c \mid f, h) U^{\pi(c)}(h) + Z(\neg c \mid f, h) U^{\pi(\neg c)}(h) \right) \right.$
$\left. + T(\neg h \mid h, f) \left(Z(c \mid f, h) U^{\pi(c)}(\neg h) + Z(\neg c \mid f, \neg h) U^{\pi(\neg c)}(\neg h) \right) \right)$
$-15 + 0.9 \left(1.0 \left(\overset{0.1}{\cancel{0.1}} \cdot 0 + 0.9 \cdot 0 \right) \right) = -15$

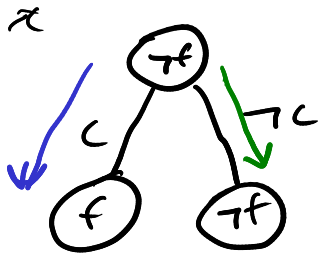


$$\alpha[s] = U^{\pi}(s)$$

To calculate α vectors n-step conditional plans, start with n-1 step conditional plans

$U^{f}(h) = -15$ $\quad$ $U^{\neg f}(h) = -10$

$U^{f}(\neg h) = -5$ $\quad$ $U^{\neg f}(\neg h) = 0$



$$U^{\pi}(s) = R(s,\neg f) + \gamma\Big(T(h|s,\neg f)\big(Z(c|\neg f,h)\,U^{f}(h) + Z(\neg c|\neg f,h)\,U^{\neg f}(h)\big) + T(\neg h|s,\neg f)\big(Z(c|\neg f,\neg h)\,U^{f}(\neg h) + Z(\neg c|\neg f,\neg h)\,U^{\neg f}(\neg h)\big)\Big)$$

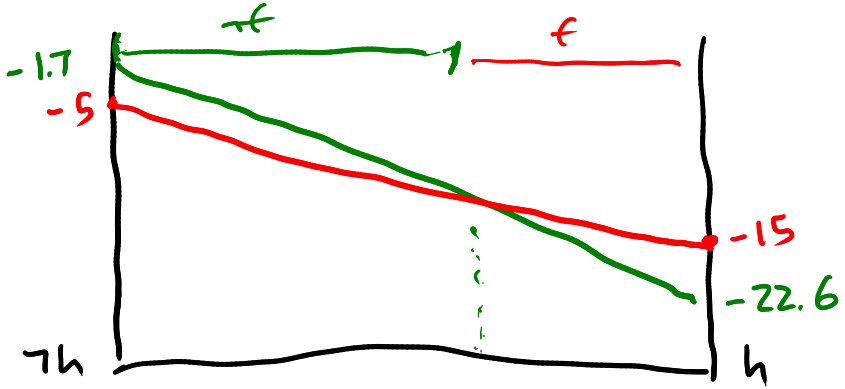
$0.8 \cdot -15 + 0.2 \cdot -10 = -14$

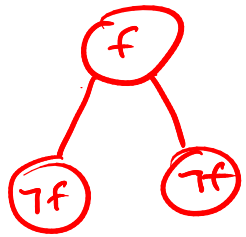
$0.1 \cdot -5 + 0.9 \cdot 0 = -0.5$

$U^{\pi}(h) = -10 + 0.9\,(1.0 \cdot -14 + 0 \cdot -0.5) = -22.6$

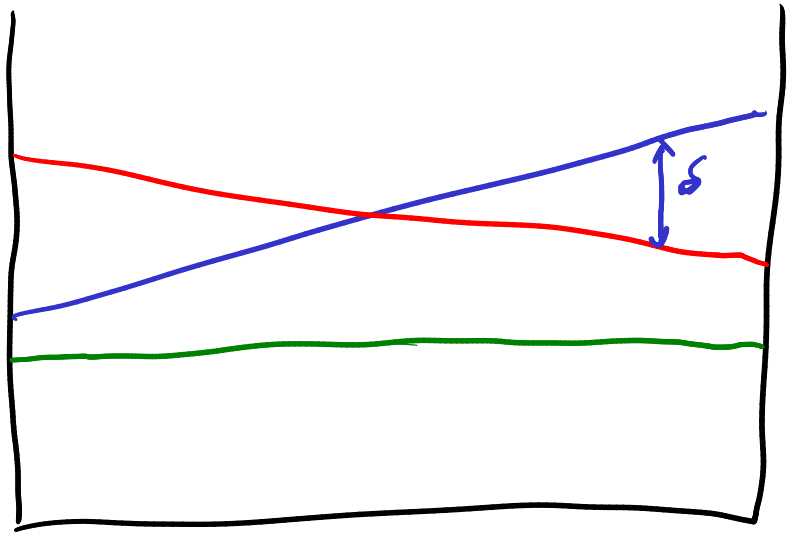
$U^{\pi}(\neg h) = 0 + 0.9\,(0.1 \cdot -14 + 0.9 \cdot -0.5) = -1.7$

$\alpha^{\pi} = [-22.6, -1.7]$





α-Vector Pruning



$$\begin{aligned} &\underset{\delta, b}{\text{maximize}} \ \delta \\ &\text{subject to} \quad b \geq 0 \\ &\qquad\qquad \mathbf{1}^T b = 1 \\ &\qquad\qquad \alpha^T b \geq \alpha'^T b + \delta \quad \forall \alpha' \in \Gamma \end{aligned}$$

- If there is a positive solution for δ then α is not dominated
- b is sometimes called a "witness"

Alpha Vector Expansion

POMDP Value Iteration (horizon d)

$\Gamma^0 \leftarrow \emptyset$

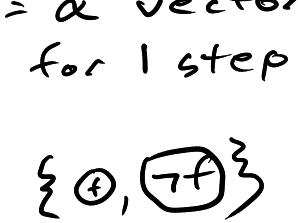
for $n \in 1 \ldots d$

Construct Γ^n by expanding with Γ^{n-1}

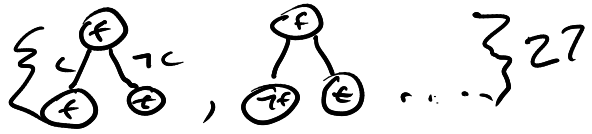
Prune Γ^n

$\Gamma^0 = \{\}$

Γ^1 = α vectors for 1 step conditional plans



Γ^2 = α vectors for 2-step conditional plans



Γ^3

Recap

Recap

- A POMDP is an MDP on the _____

Recap

- A POMDP is an MDP on the belief space

Recap

- A POMDP is an MDP on the belief space
- The value function of a discrete POMDP can be represented by a set of _____

Recap

- A POMDP is an MDP on the belief space
- The value function of a discrete POMDP can be represented by a set of α-vectors

Recap

- A POMDP is an MDP on the belief space
- The value function of a discrete POMDP can be represented by a set of α-vectors
- Each α vector corresponds to a _____

Recap

- A POMDP is an MDP on the belief space
- The value function of a discrete POMDP can be represented by a set of α-vectors
- Each α vector corresponds to a conditional plan