

Recap

Recap

- POMDP

Recap

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

Recap

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

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

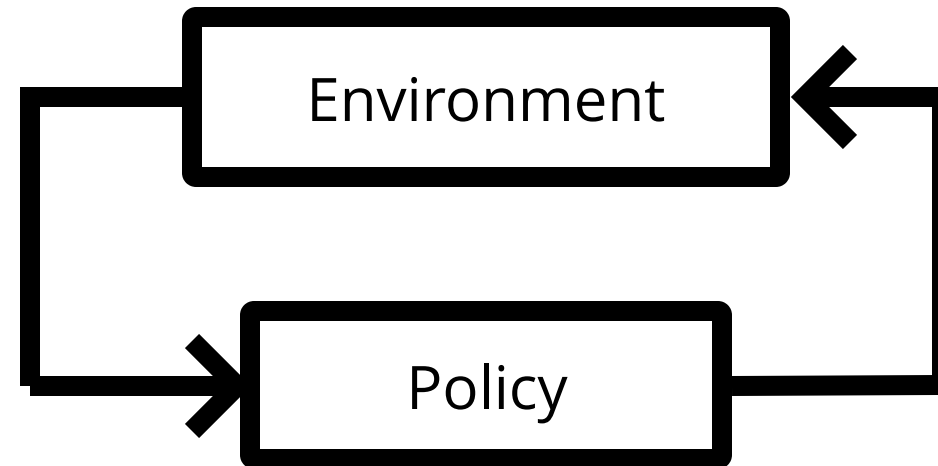
- POMDP $(S, A, O, R, T, Z, \gamma)$
- Belief Updates

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

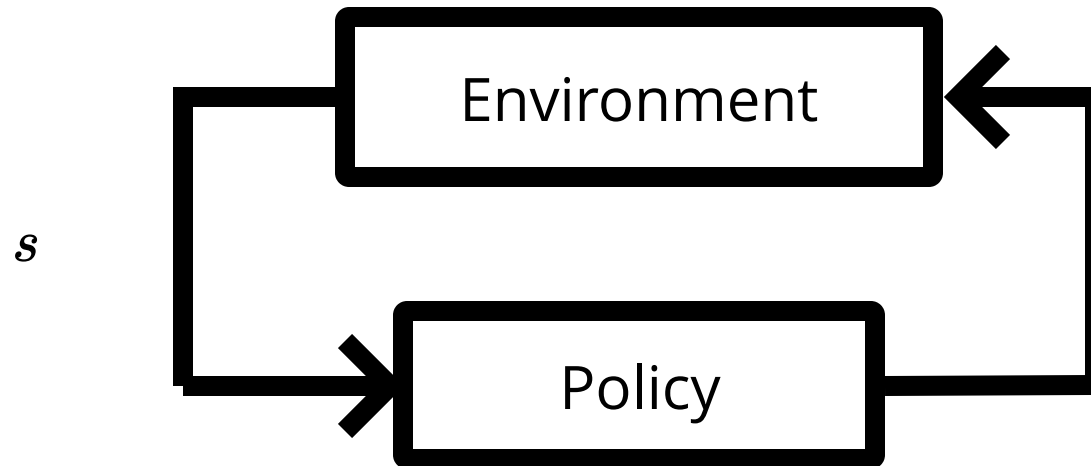
$$b' = \tau(b, a, o)$$

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


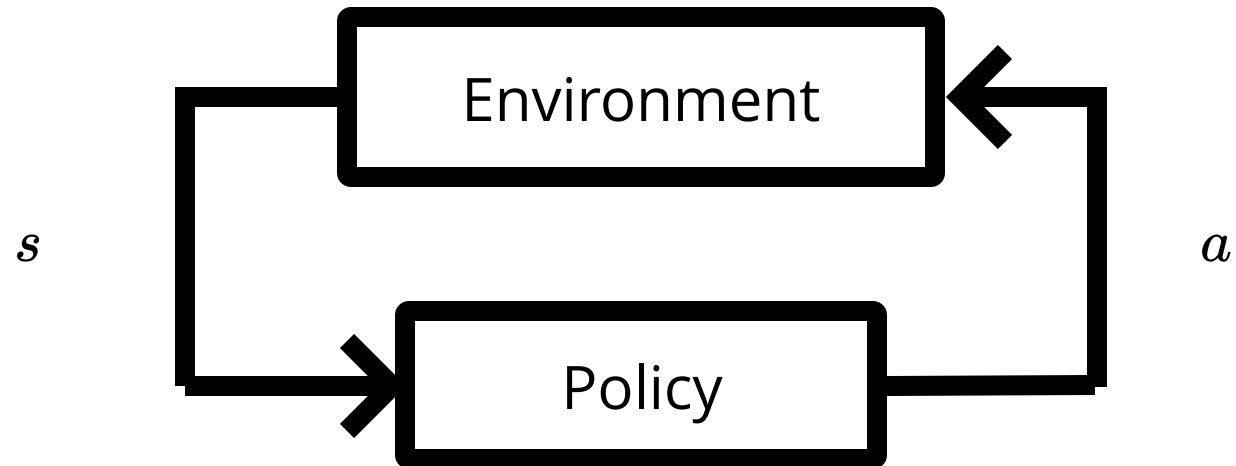
MDP Sense-Plan-Act Loop



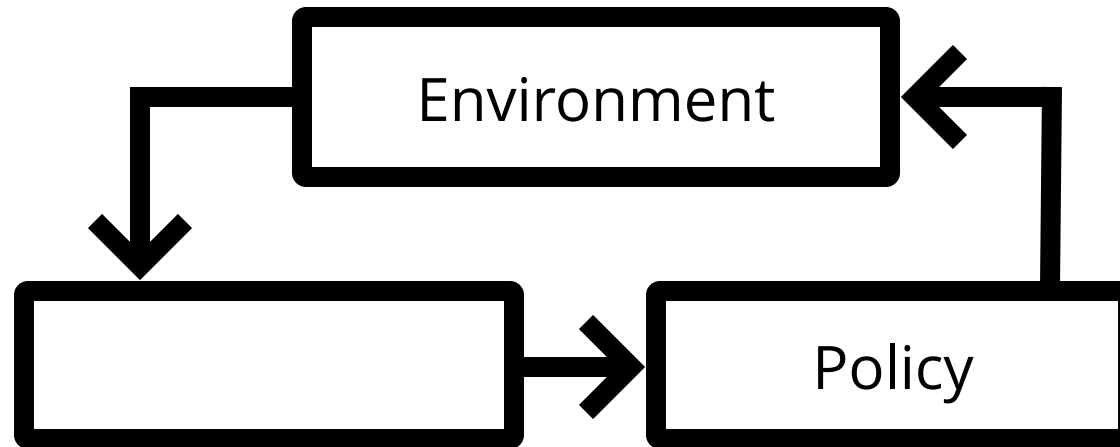
MDP Sense-Plan-Act Loop



MDP Sense-Plan-Act Loop



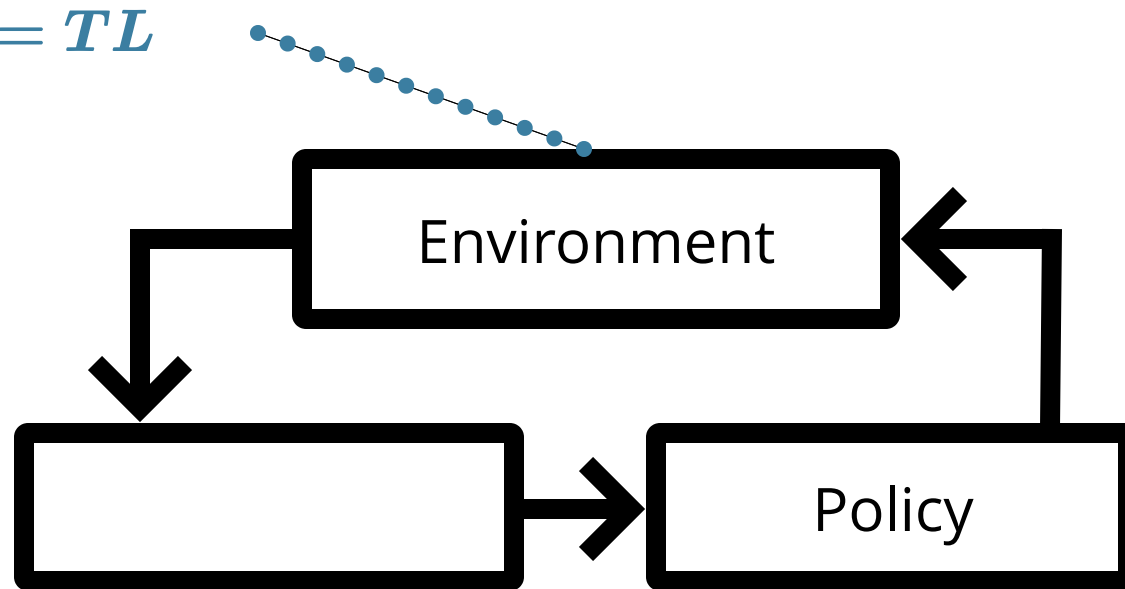
POMDP Sense-Plan-Act Loop



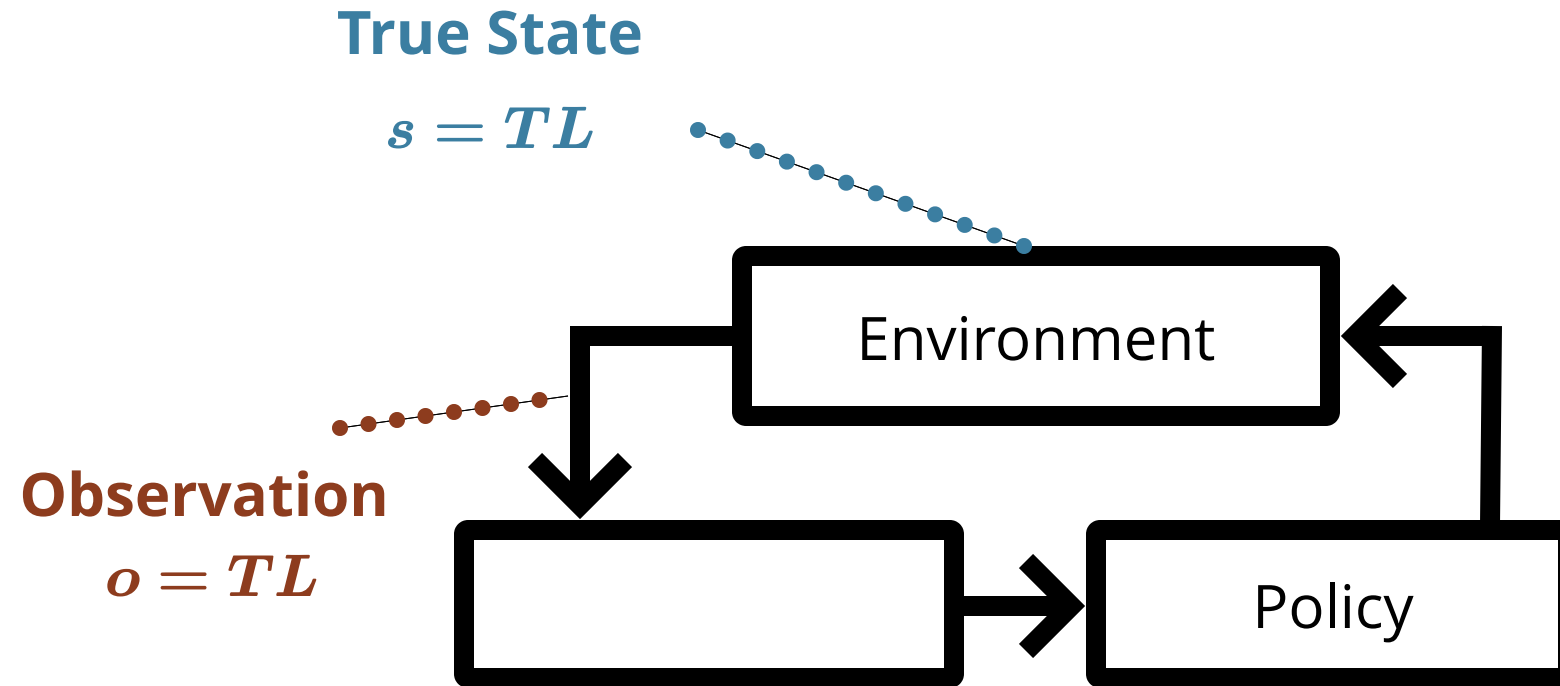
POMDP Sense-Plan-Act Loop

True State

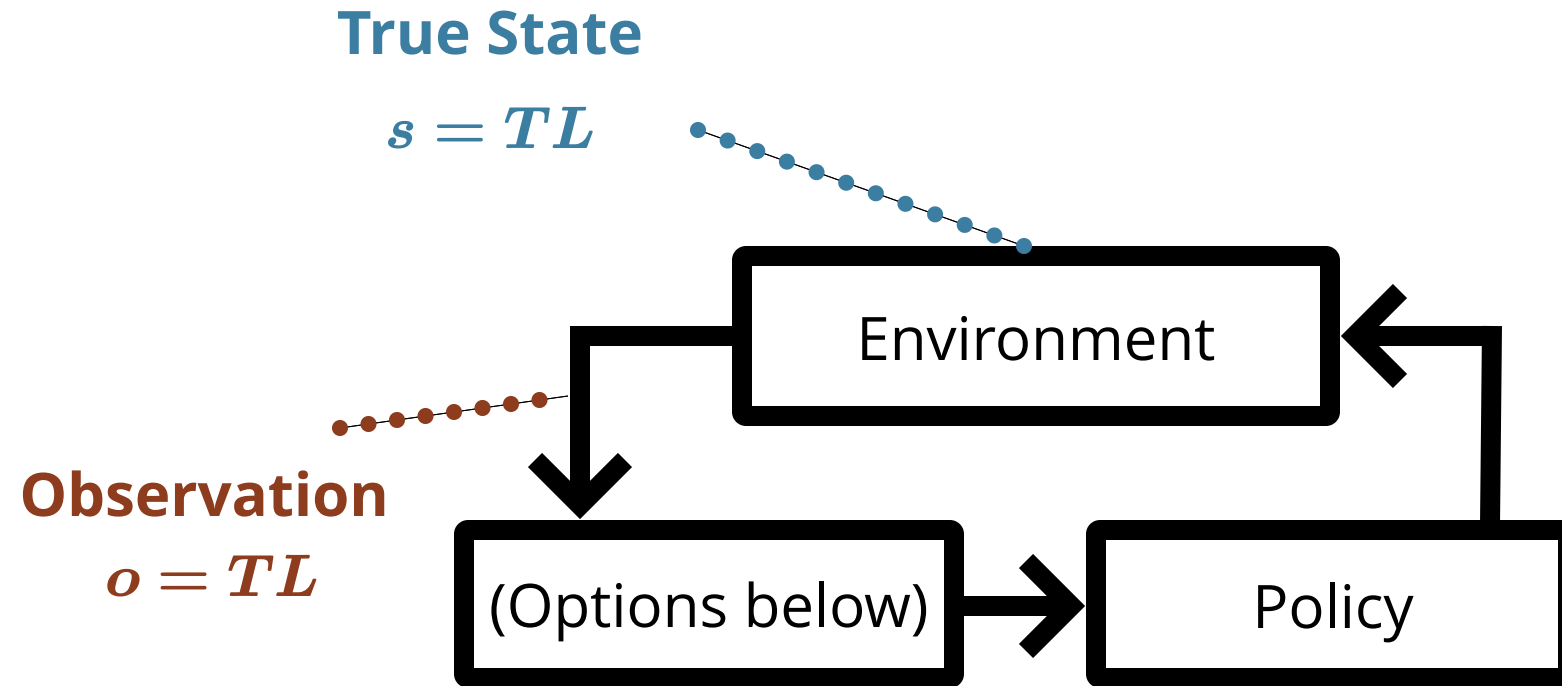
$$s = TL$$



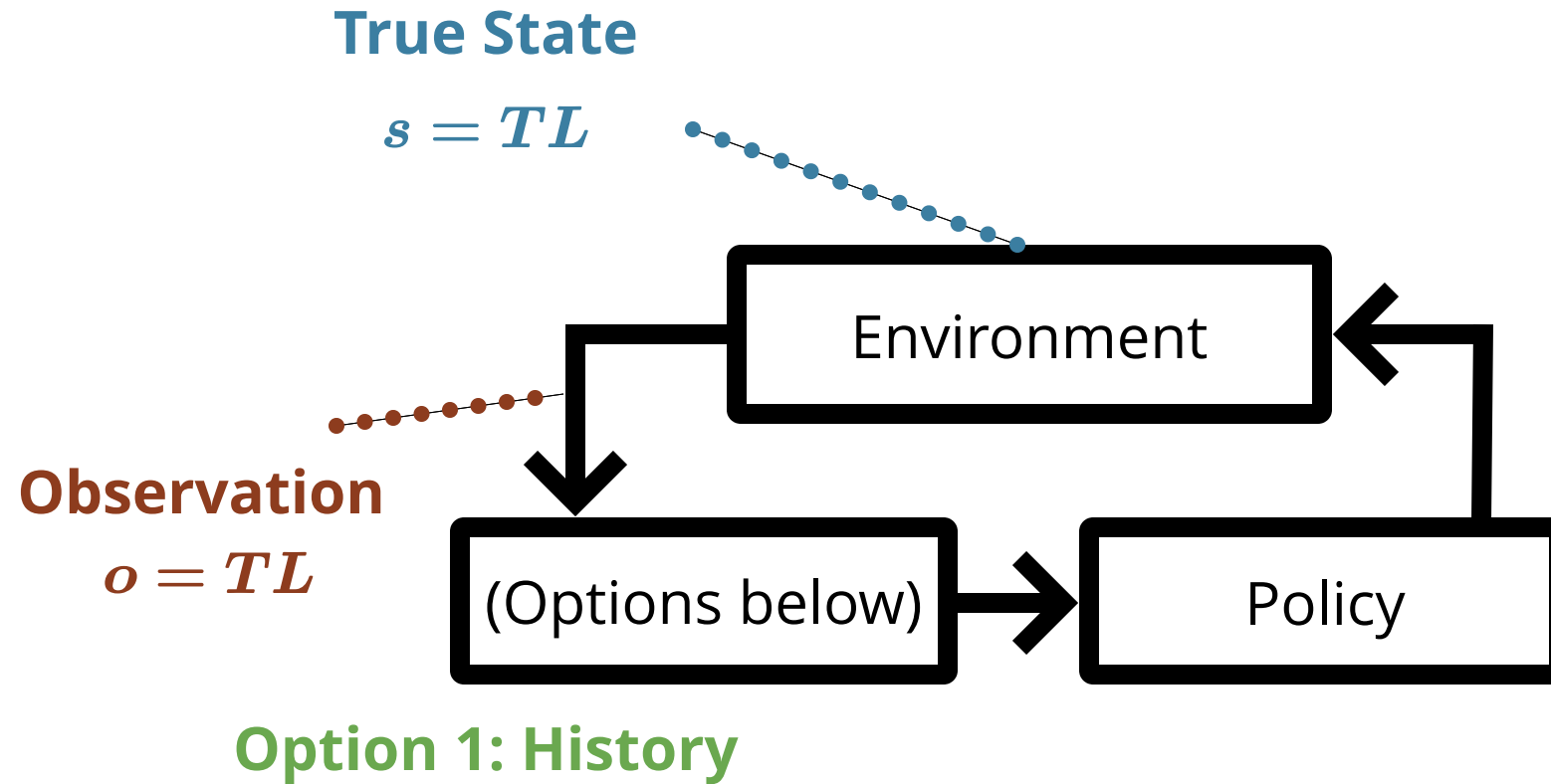
POMDP Sense-Plan-Act Loop



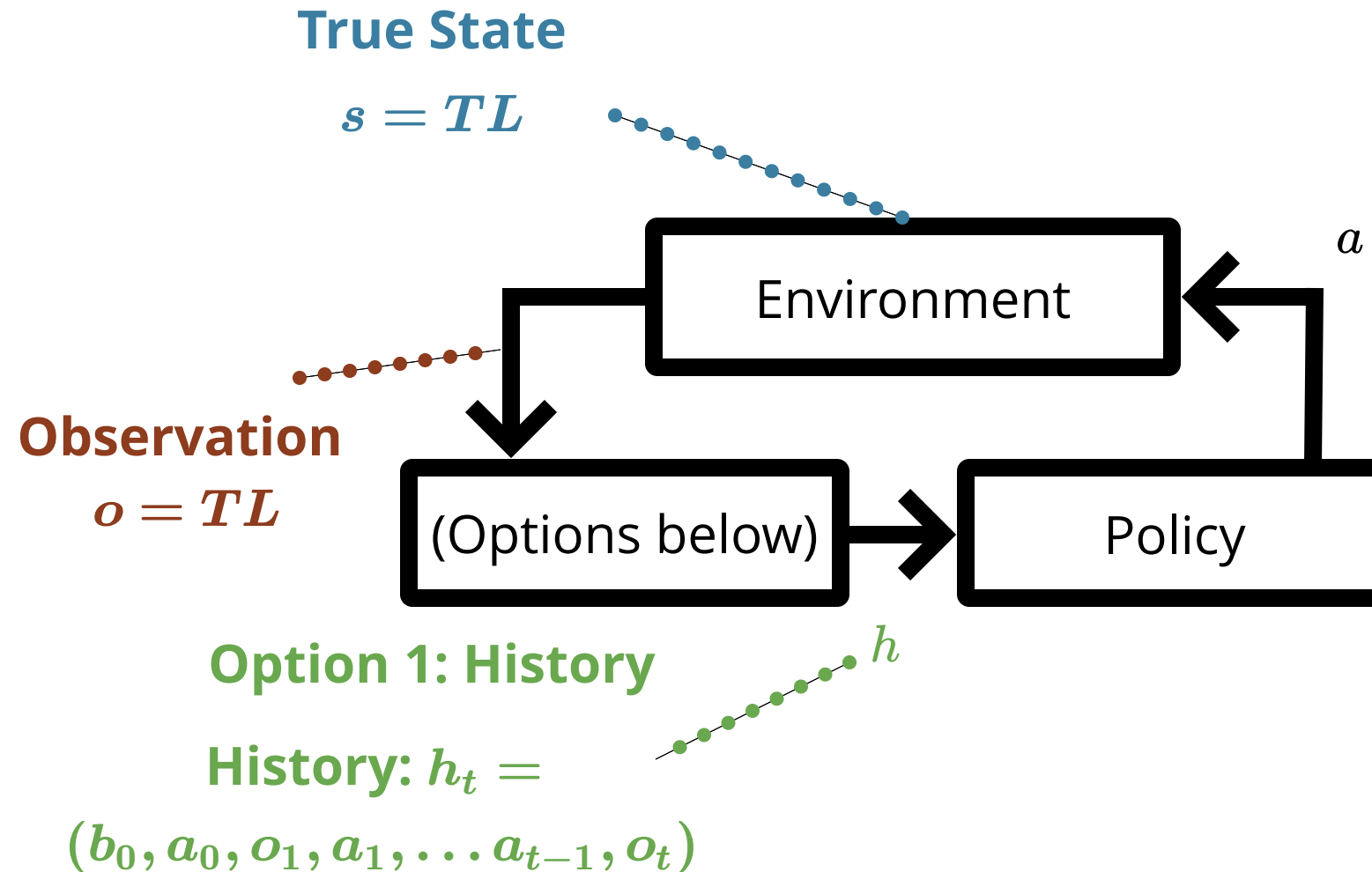
POMDP Sense-Plan-Act Loop



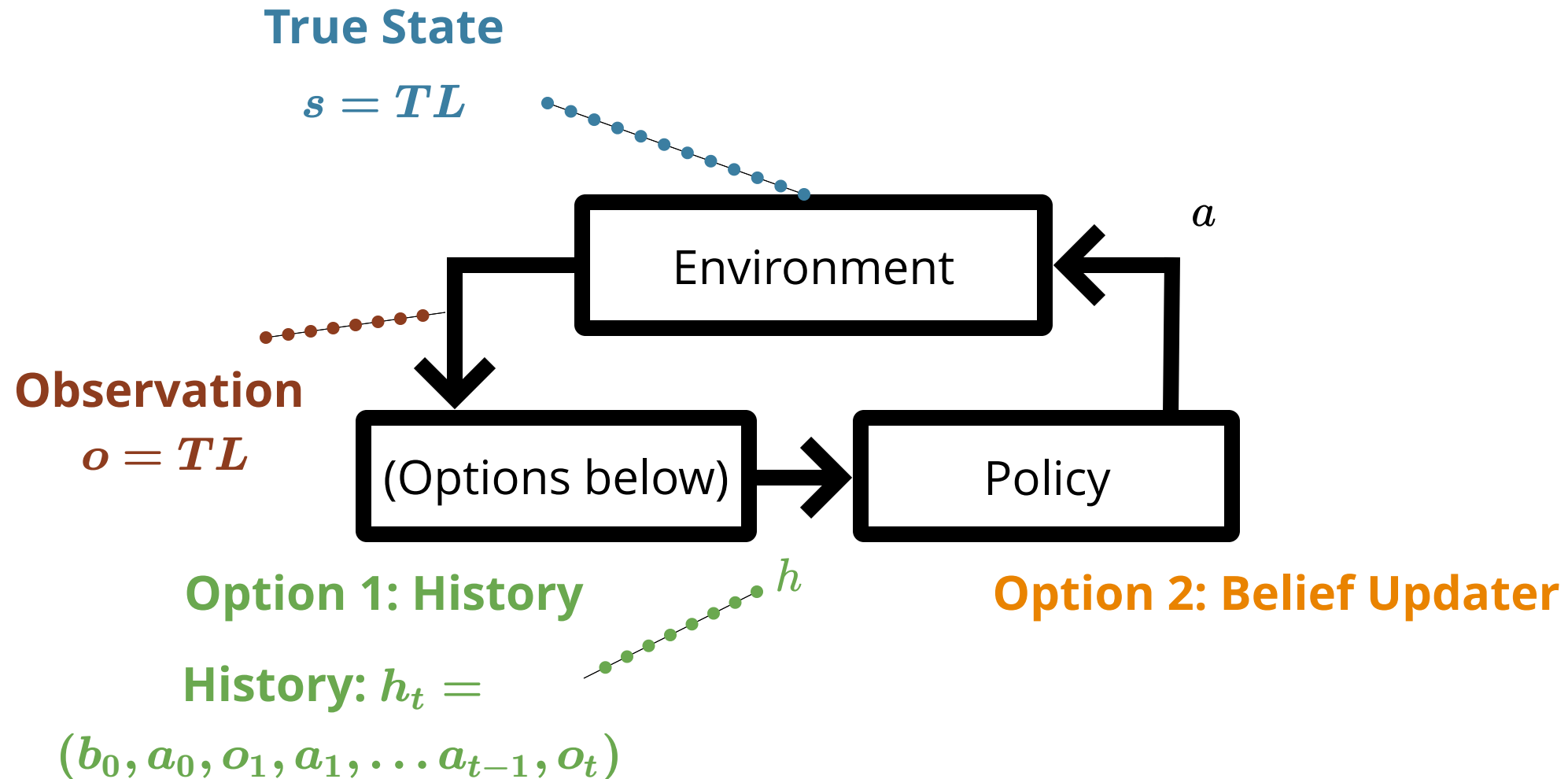
POMDP Sense-Plan-Act Loop



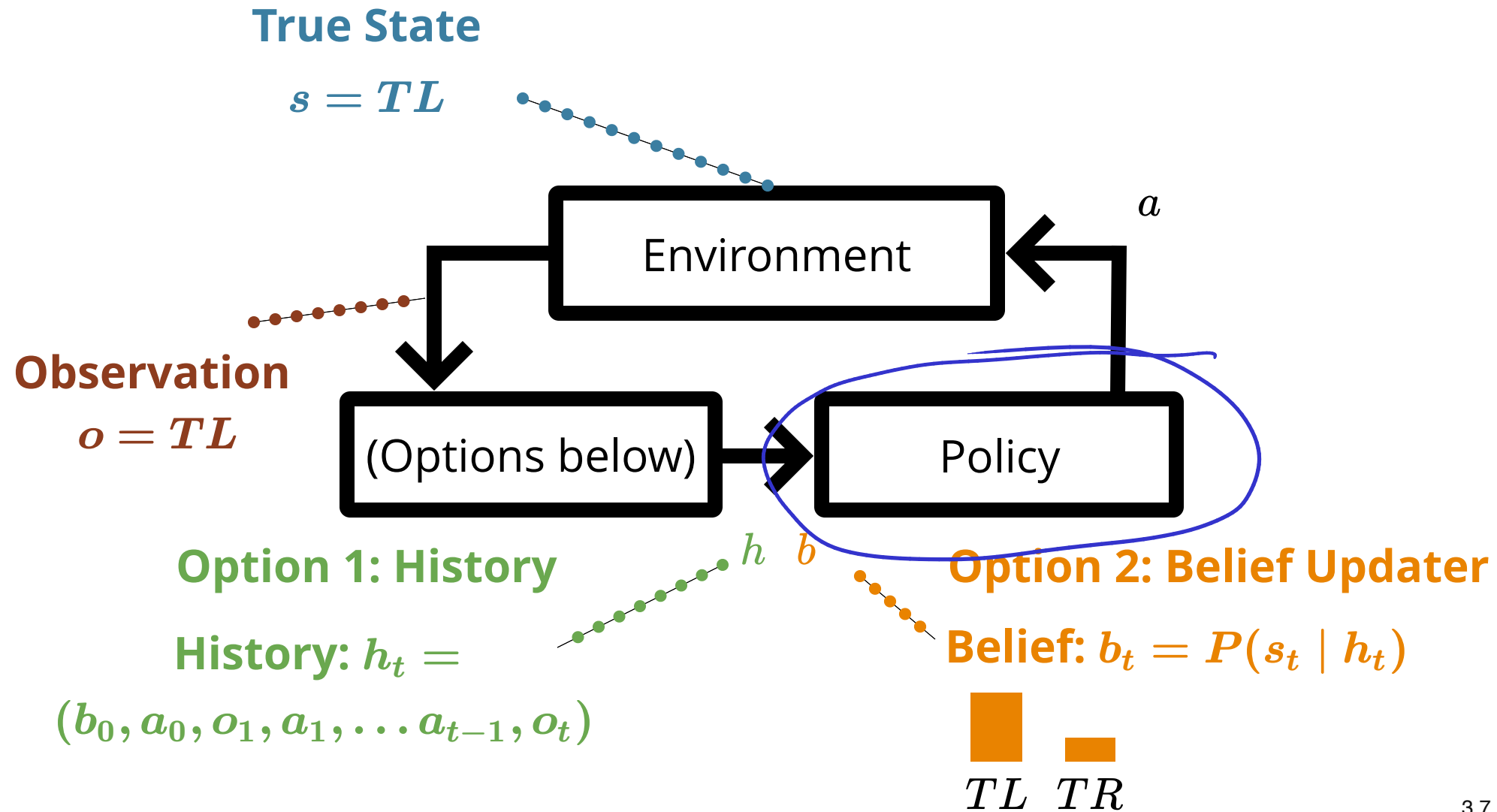
POMDP Sense-Plan-Act Loop



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

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

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

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

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

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

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

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

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

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

Exercise 1: Crying Baby Belief Update

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

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

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

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

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

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

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

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

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

$$\gamma = 0.9$$

Exercise 1: Crying Baby Belief Update

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

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

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

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

Starting at a $b(h) = 0$, calculate

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

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

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

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

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

$$\gamma = 0.9$$

b' with $a = \neg f$ and $o = c$.

$$b'(h) \propto \overset{Z(c|\neg f,h)}{0.8} \left(\overset{T(h|h,\neg f)}{0.1} \overset{b(h)}{1.0} + \overset{T(h|\neg h,\neg f)}{1.0} \overset{b(\neg h)}{0.0} \right) = 0.08$$

$$b'(h) \propto \underset{0.8}{Z(c|\neg f,h)} \left(\underset{0.1}{T(h|\neg h,\neg f)} \overset{1.0}{b(\neg h)} + \underset{1.0}{T(h|h,\neg f)} \overset{0.0}{b(h)} \right)$$

$$b'(h) \propto 0.08$$

$$b'(\neg h) \propto \underset{0.1}{Z(c|\neg f,\neg h)} \left(\underset{0.9}{T(\neg h|\neg h,\neg f)} \overset{1.0}{b(\neg h)} + \underset{1.0}{T(\neg h|h,\neg f)} \overset{0.0}{b(h)} \right)$$

$$= 0.09$$

$b'(h) = 0.08 / (0.08 + 0.09) = 47\%$
 $b'(\neg h) = 0.09 / (0.08 + 0.09) = 53\%$

$b'(h) =$
 $b'(\neg h) =$

Belief Dynamics

Belief Dynamics

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

Belief Dynamics

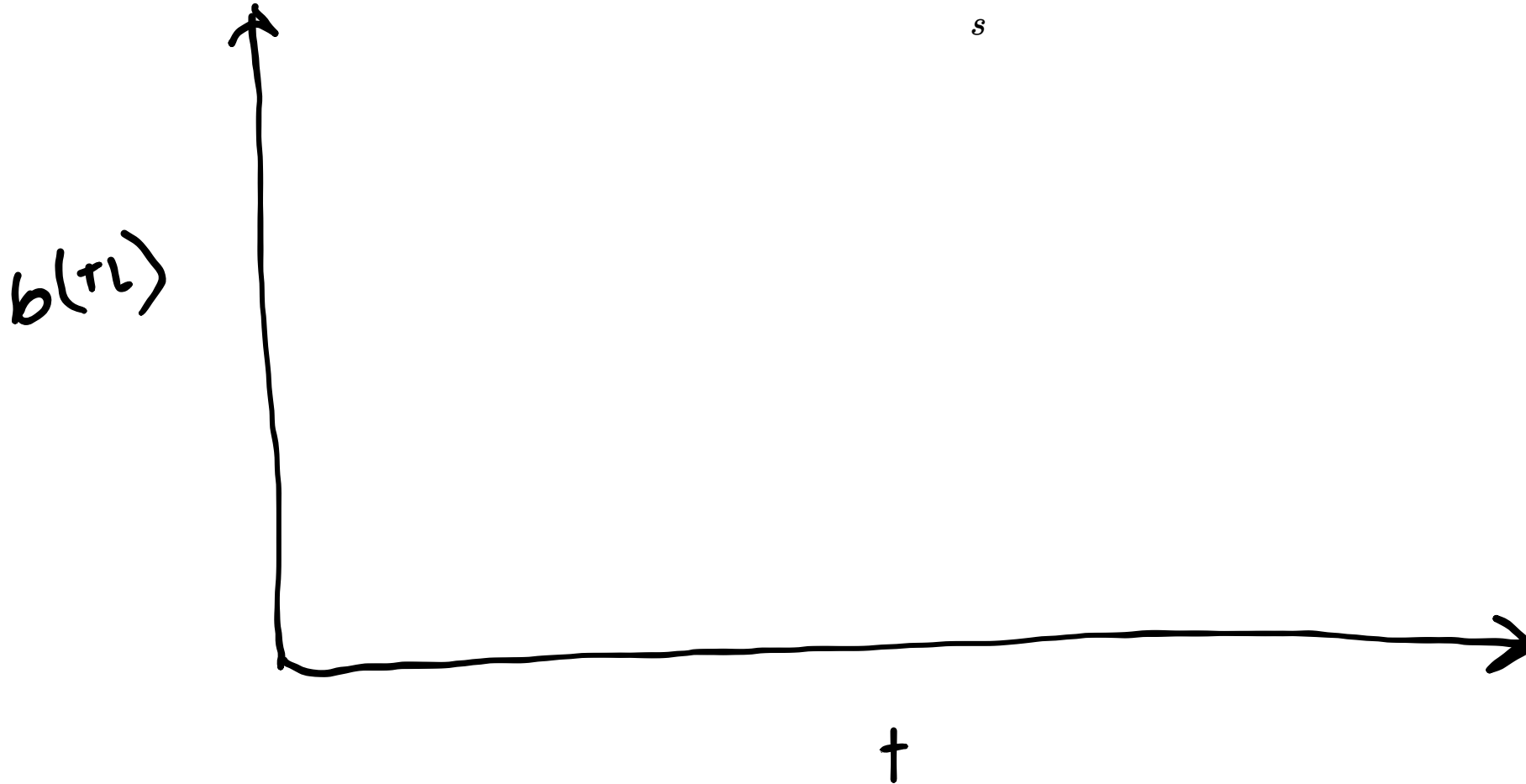
$$b'(s') \propto Z(o \mid a, s') \sum_s T(s' \mid s, a) b(s)$$

$b(\tau_L)$

†

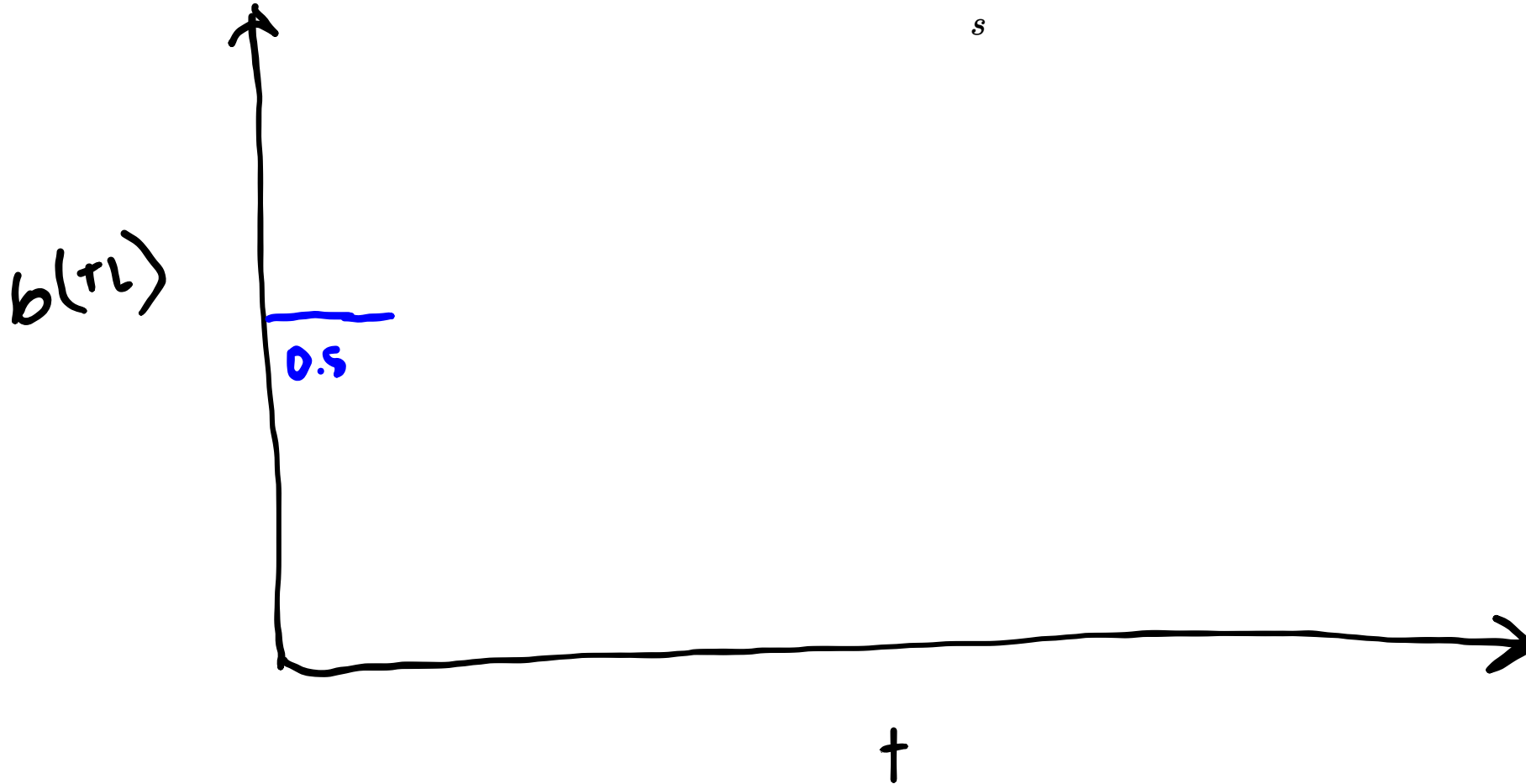
Belief Dynamics

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



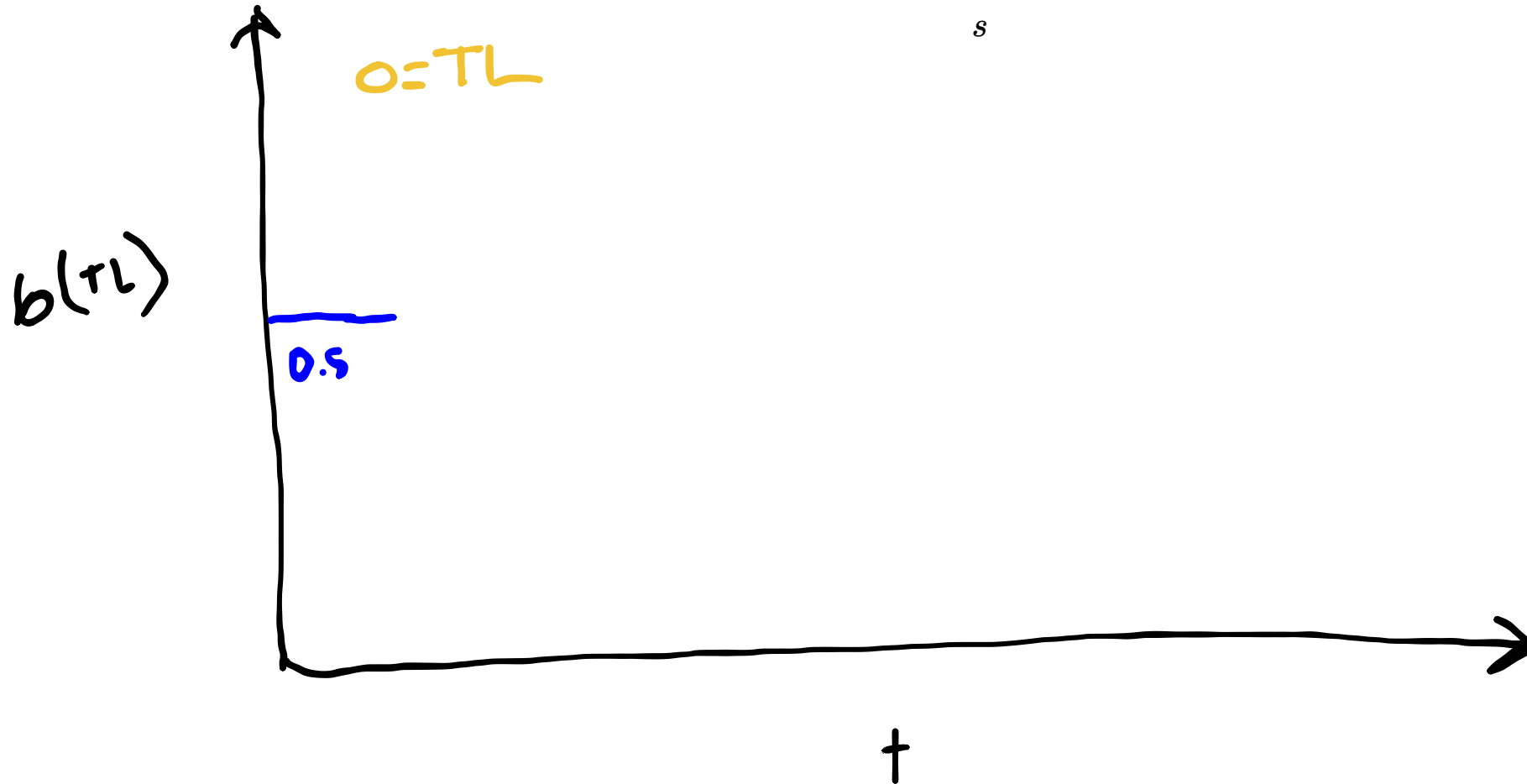
Belief Dynamics

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



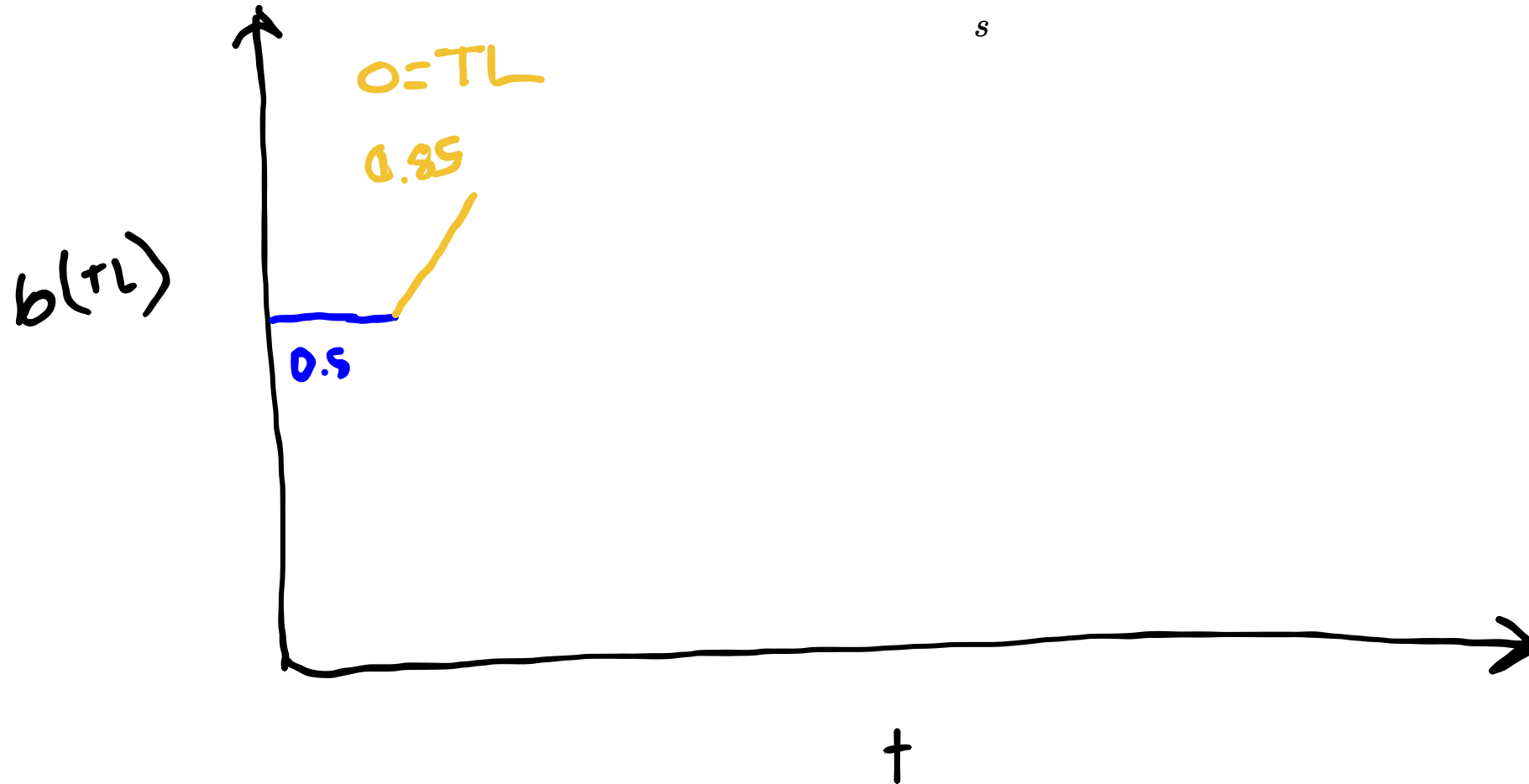
Belief Dynamics

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



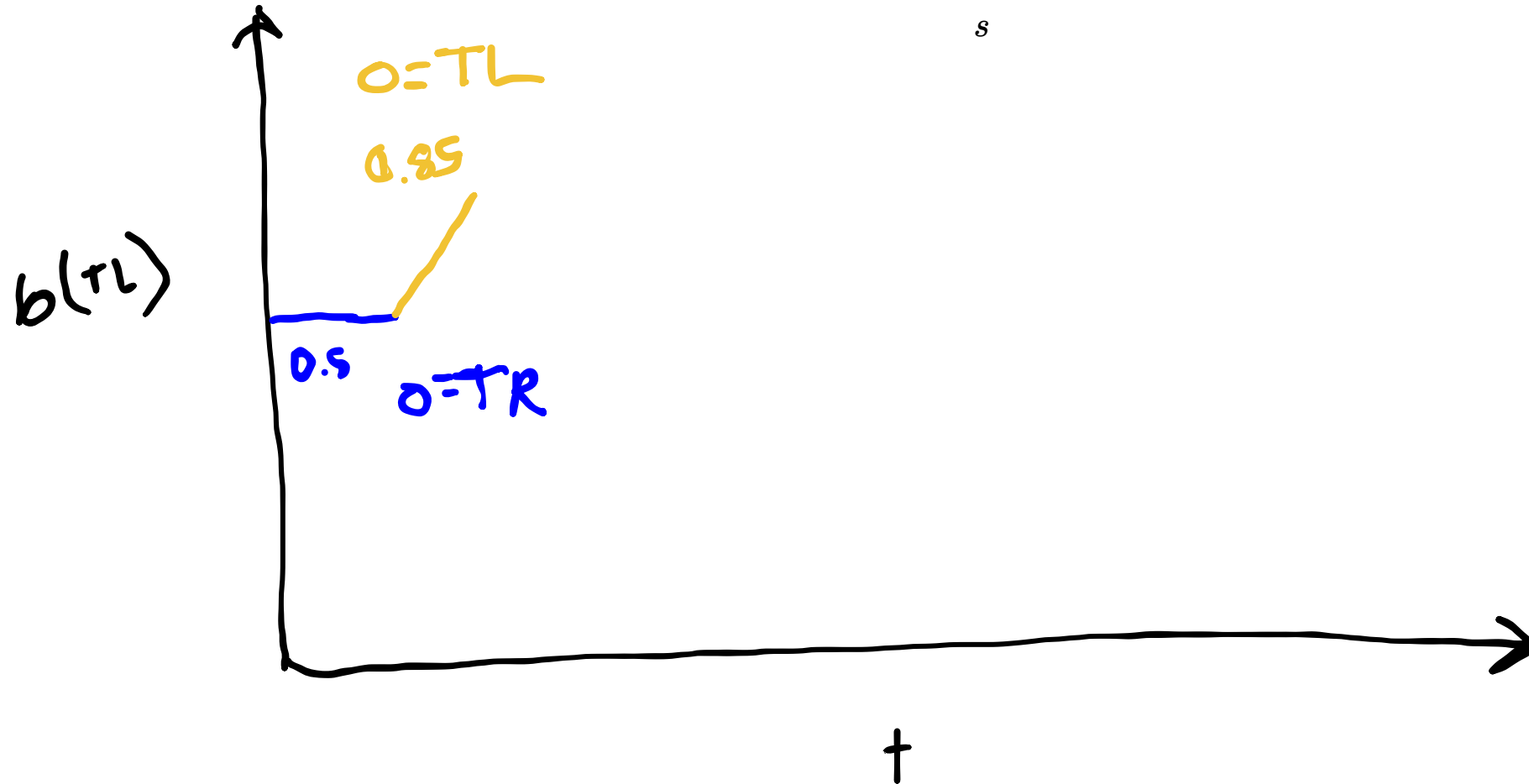
Belief Dynamics

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



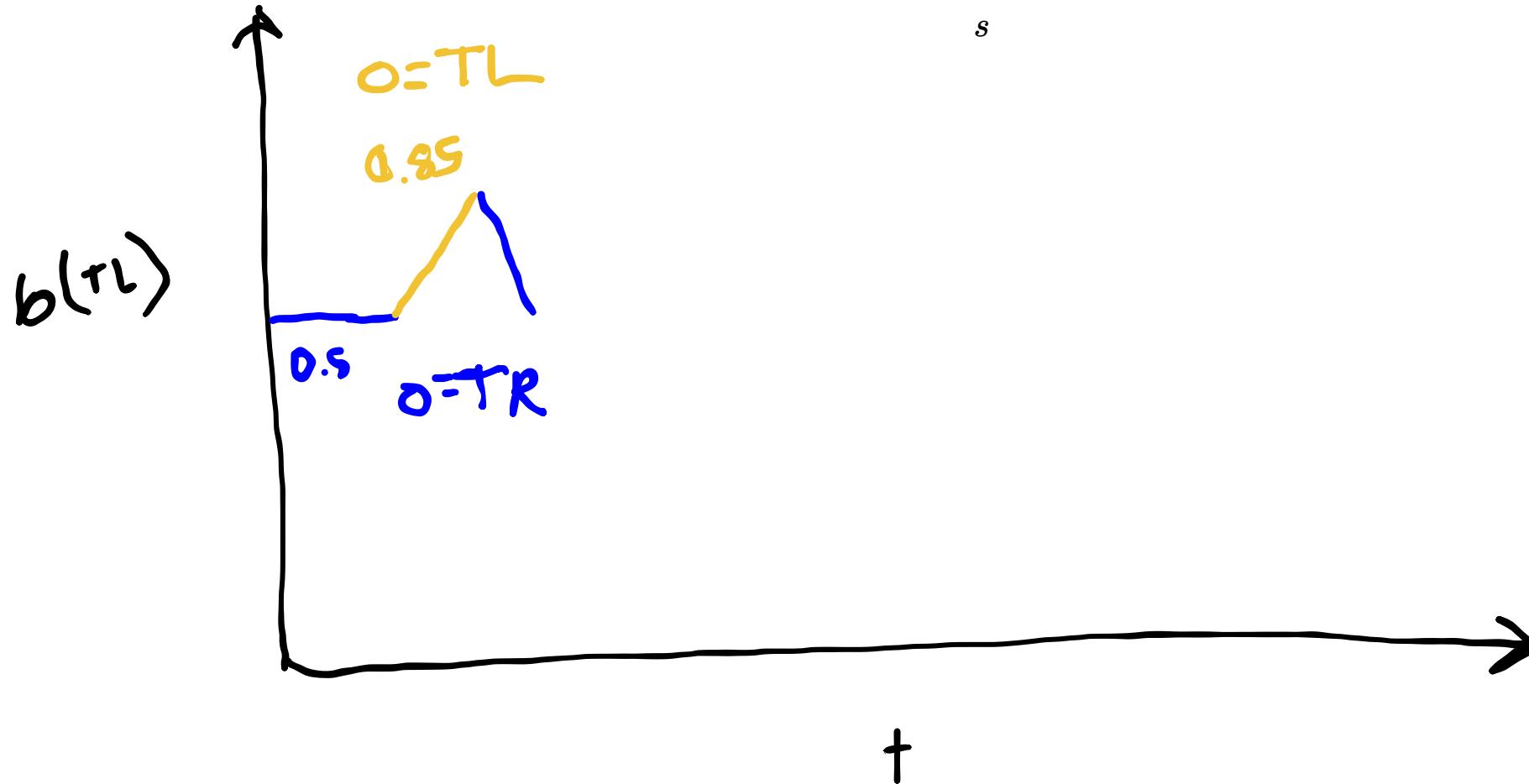
Belief Dynamics

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



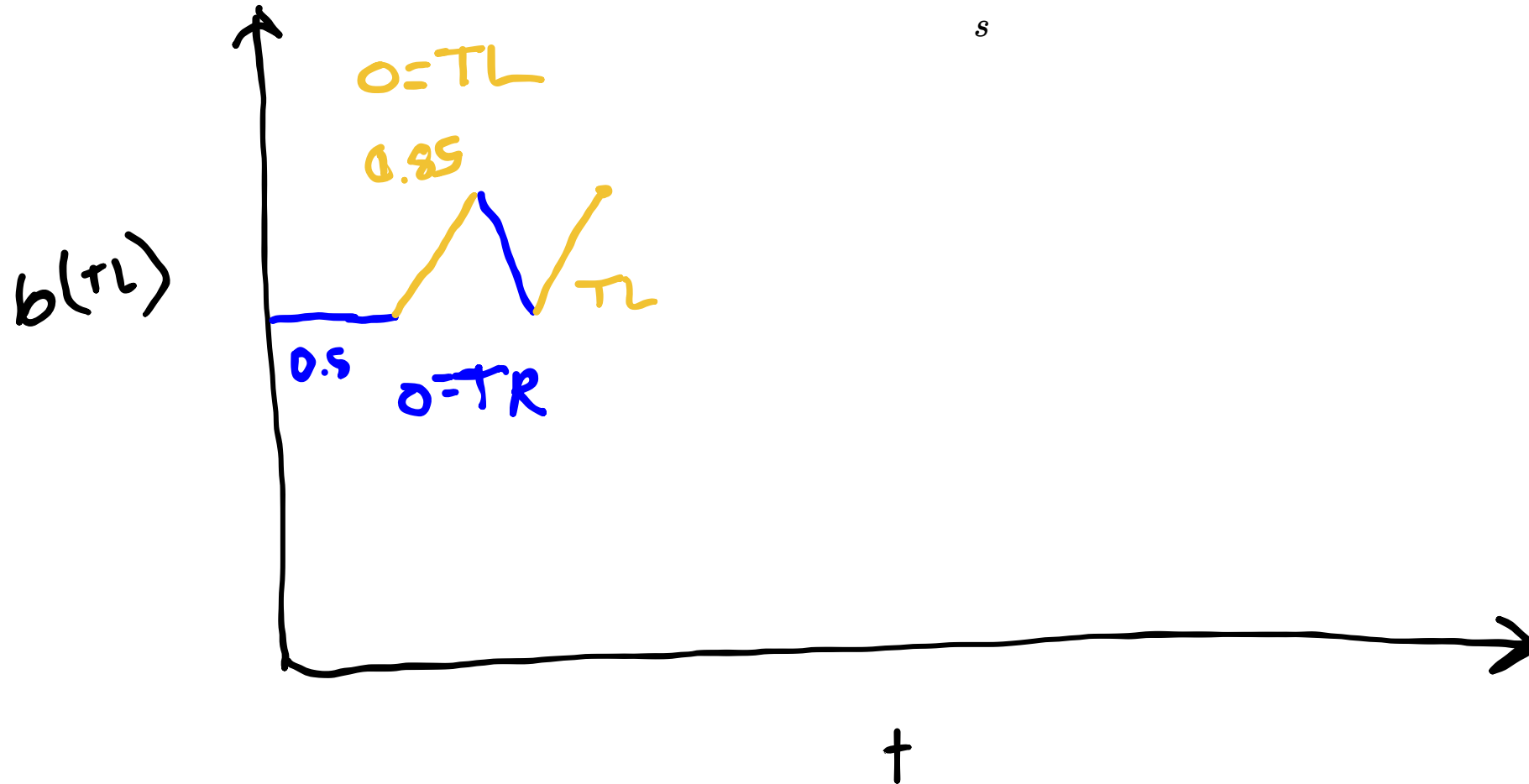
Belief Dynamics

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



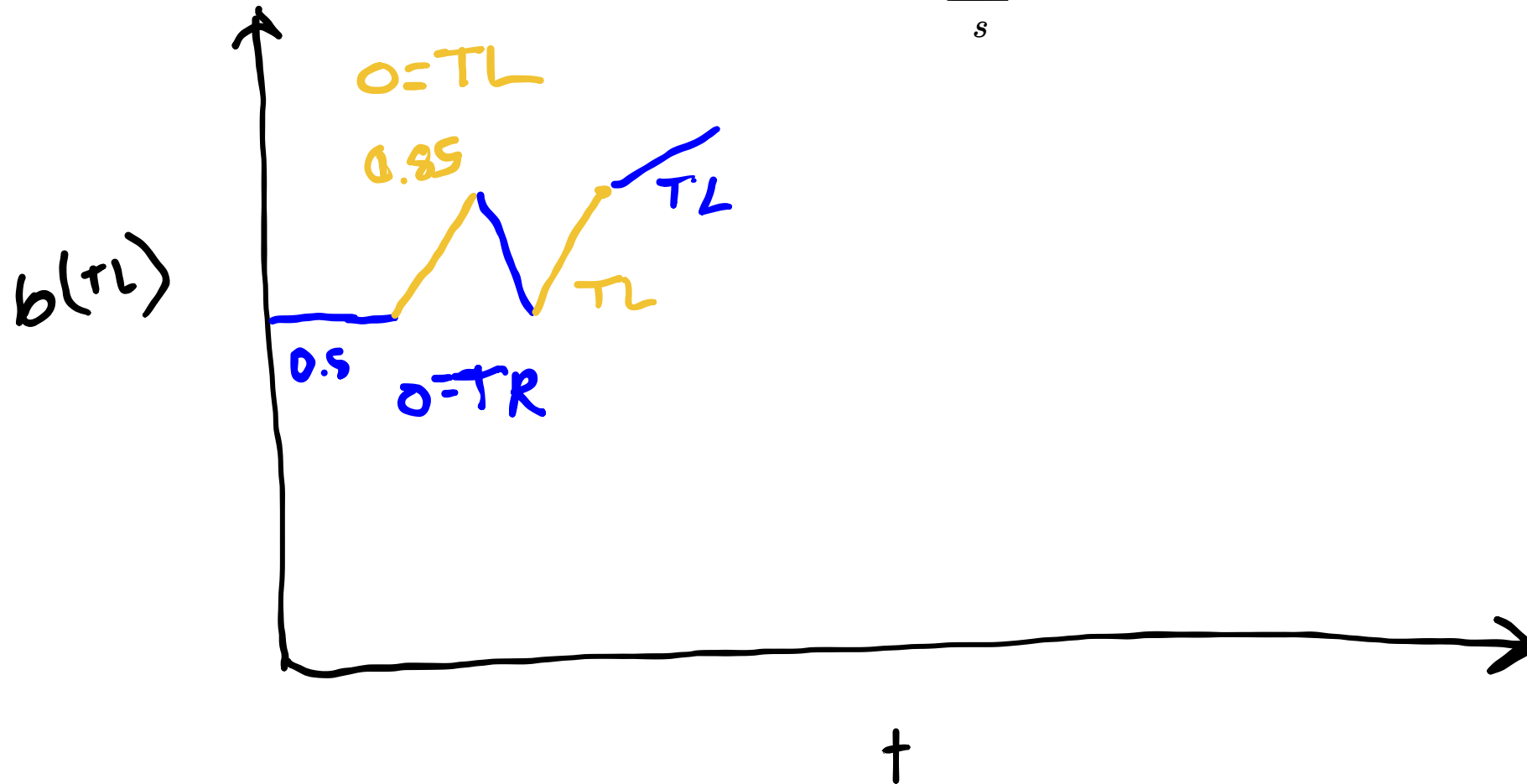
Belief Dynamics

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



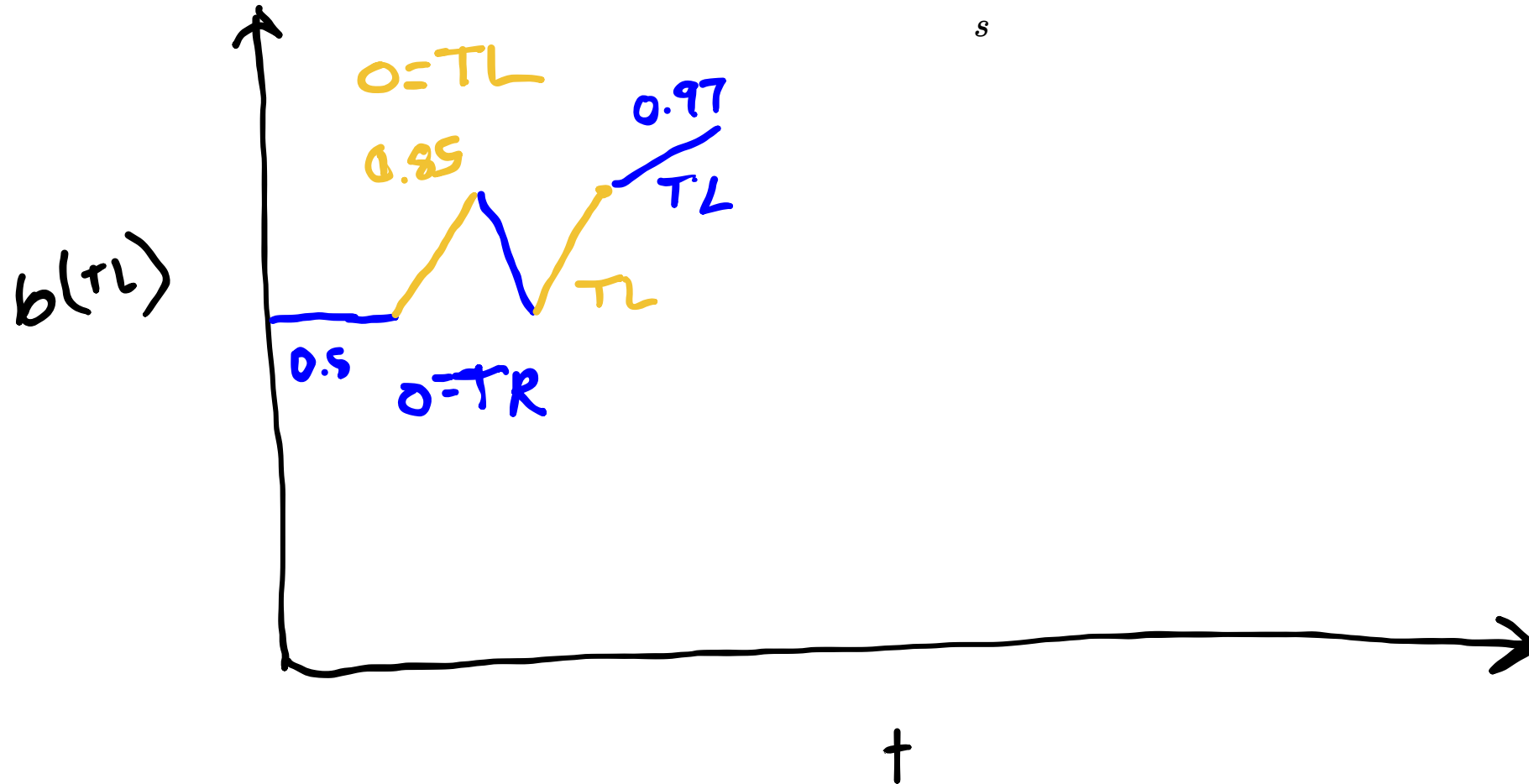
Belief Dynamics

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



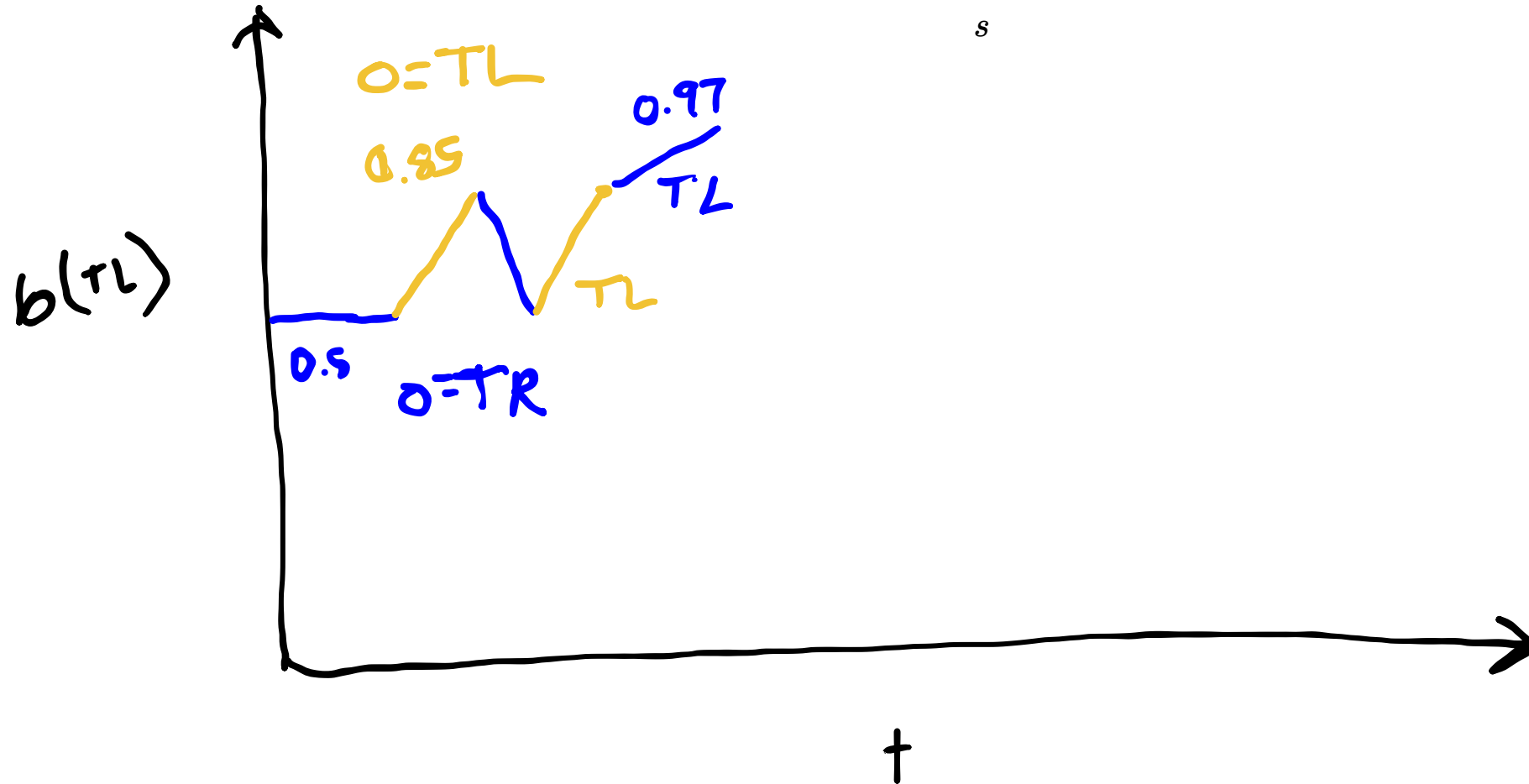
Belief Dynamics

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



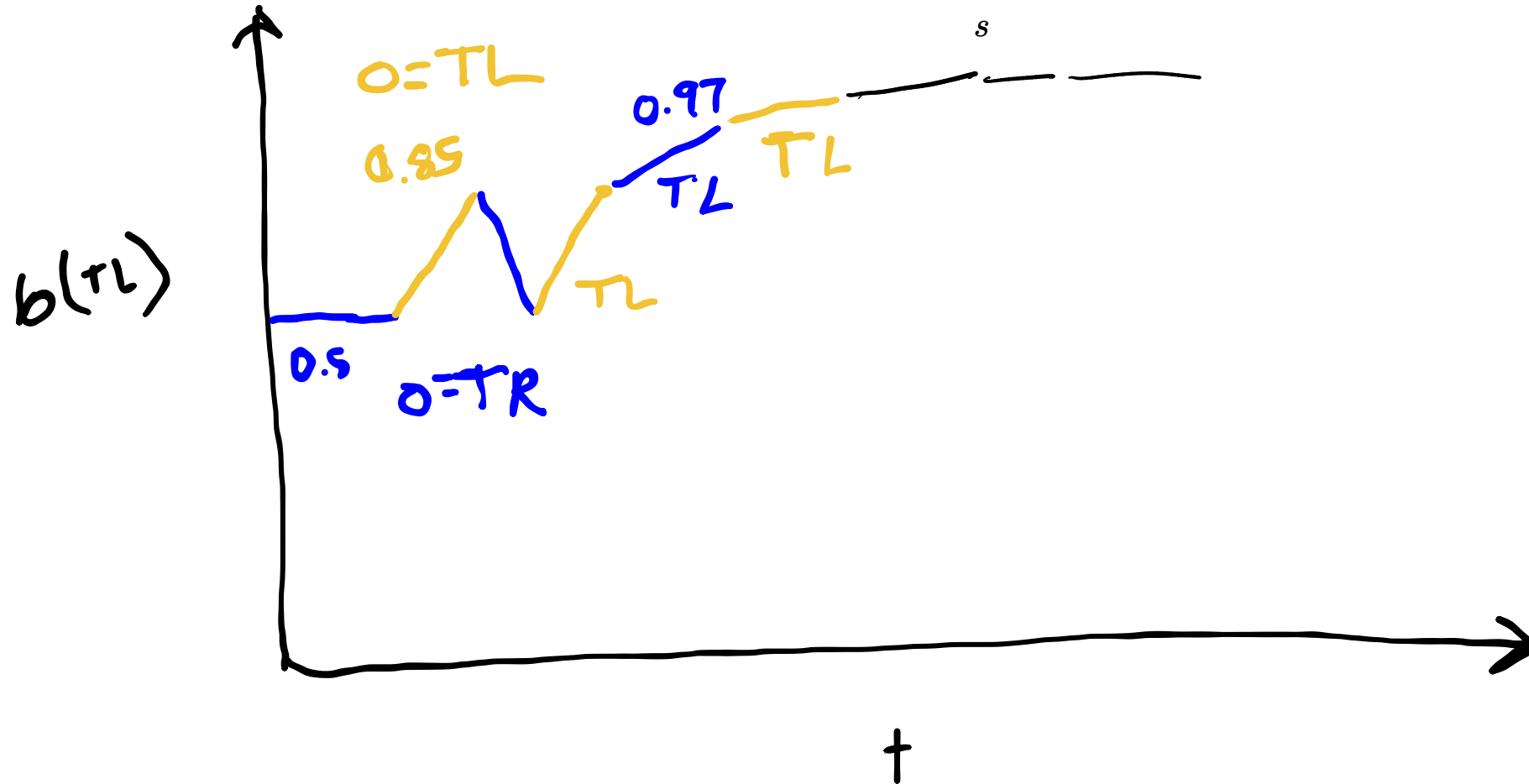
Belief Dynamics

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

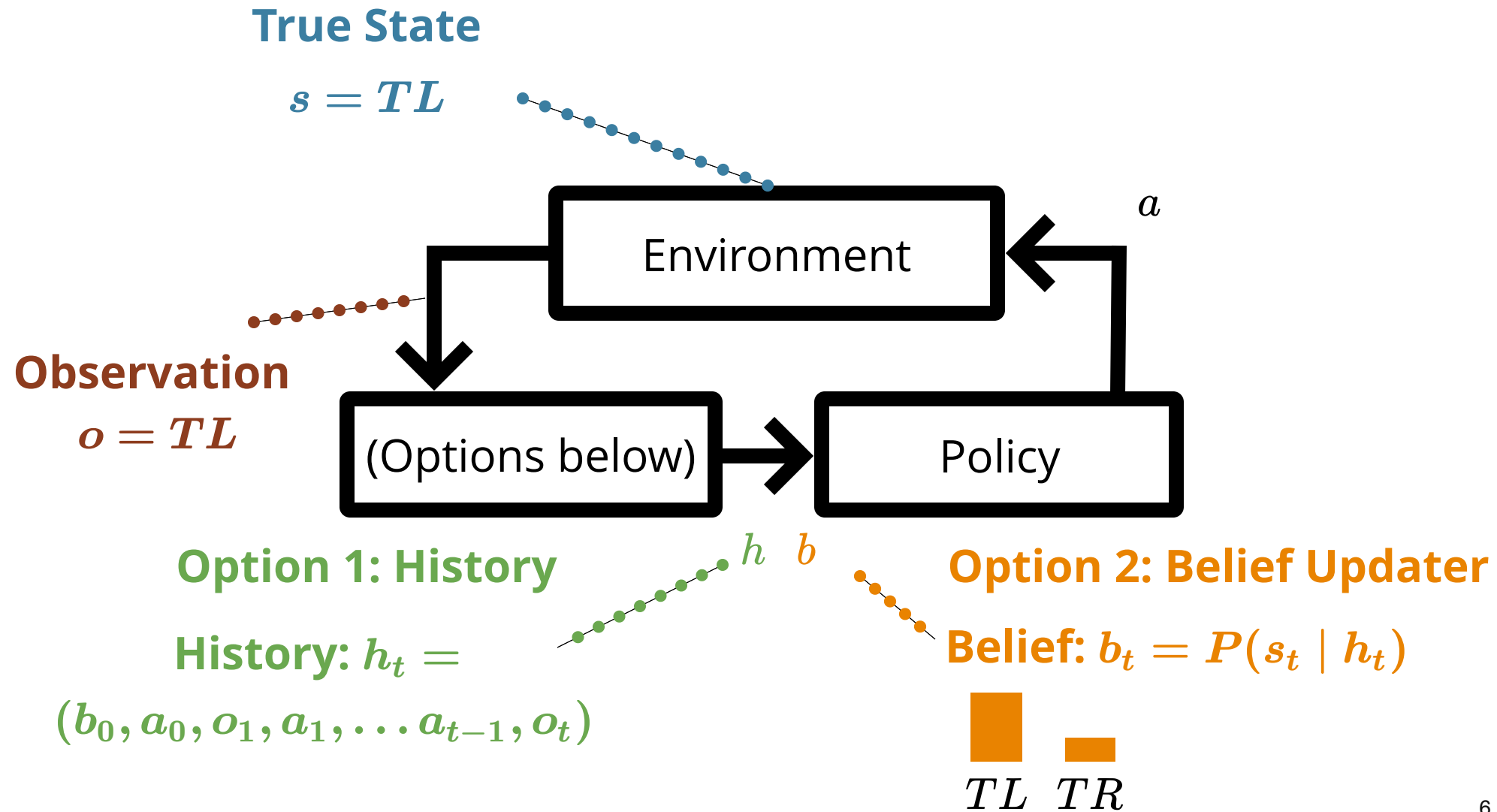


Belief Dynamics

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

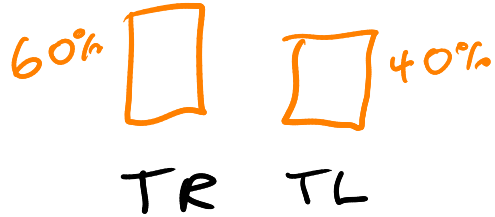


POMDP Sense-Plan-Act Loop



Guiding Quesiton

How do we calculate the optimal action in a POMDP?



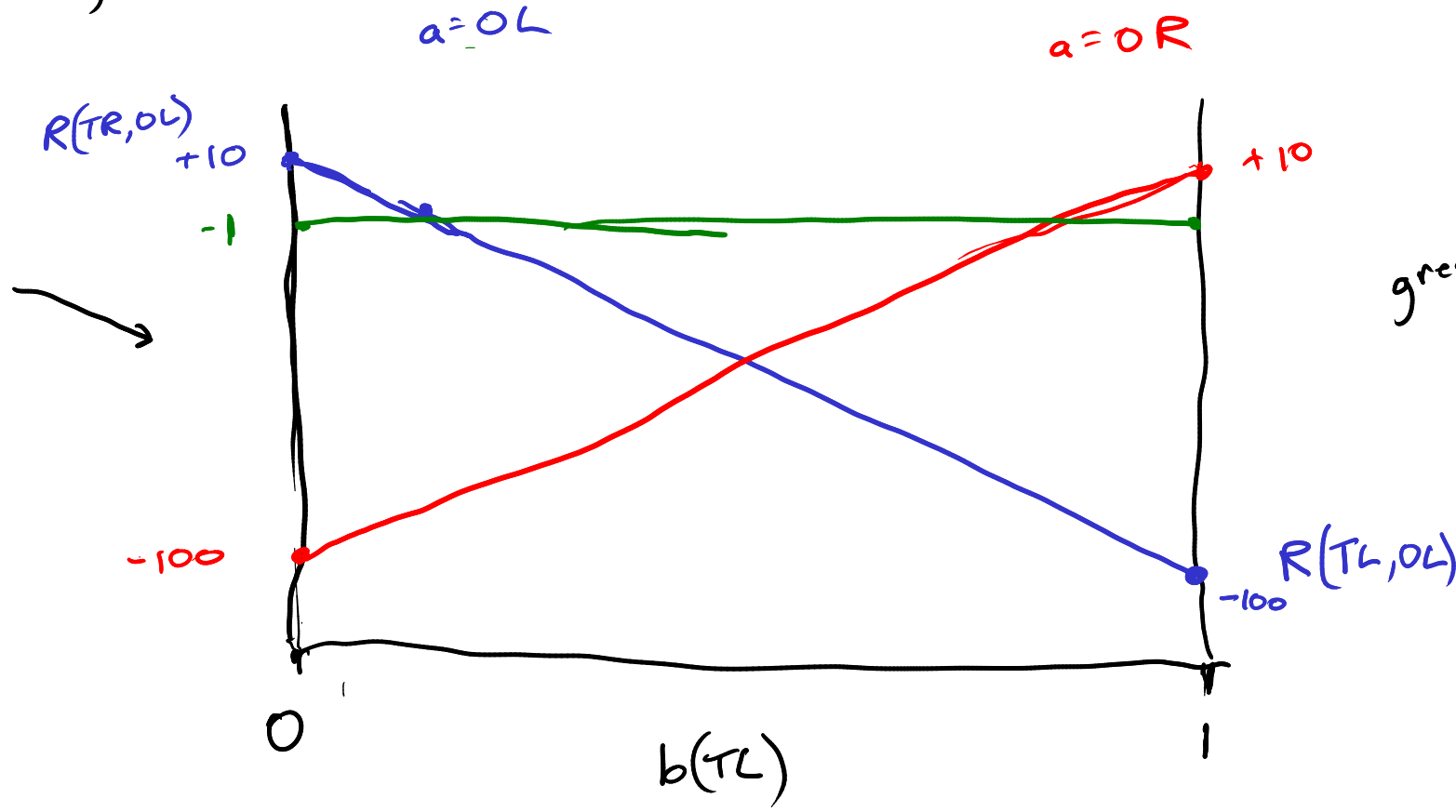
One-step utility

Reward +10 empty
 -1 Listen
 -100 tiger

$$R(b, a)$$

$$= E[R(s, a)]$$

$s \sim b$



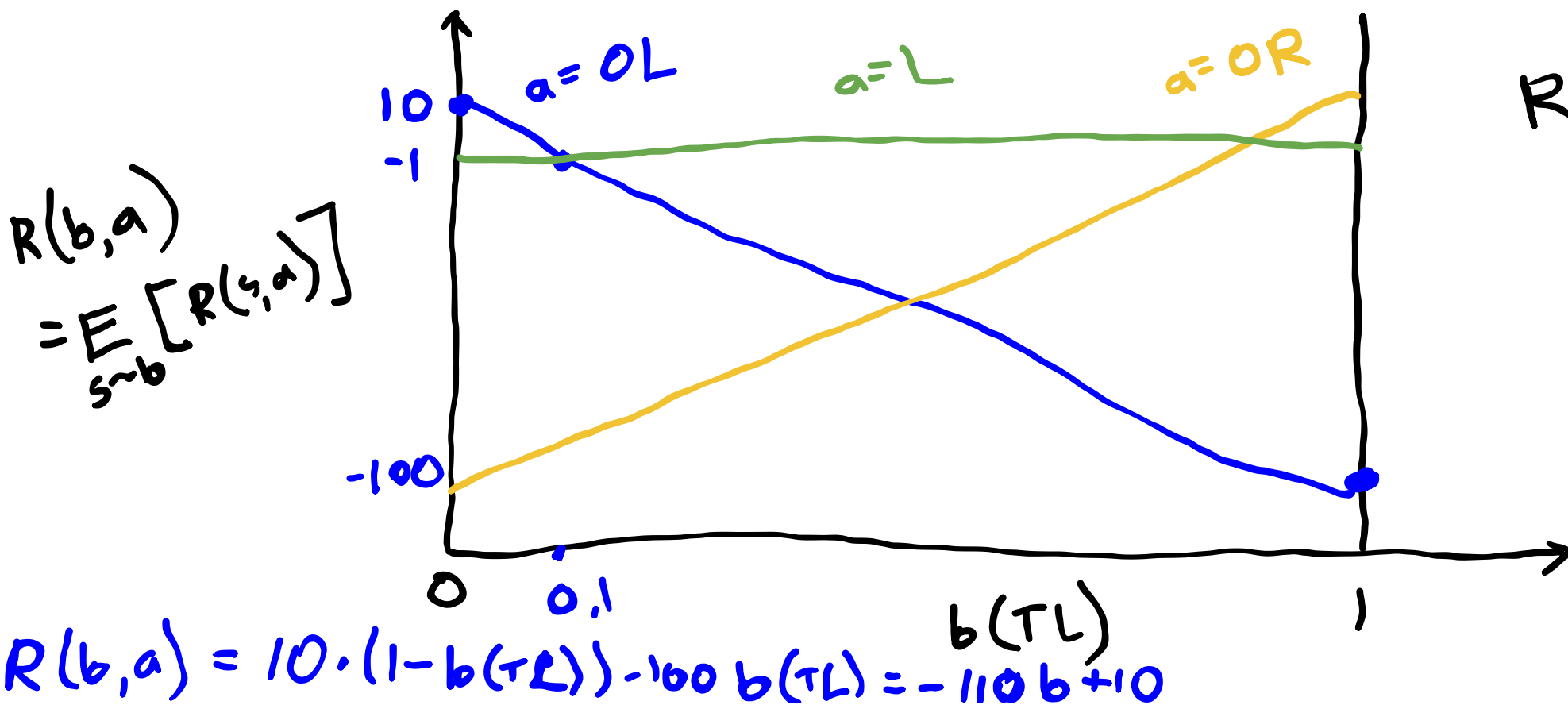
$R(b, a) = \bar{r}^a \cdot b$
 greedy action
 $\arg \max_a \bar{r}^a \cdot b$
 1-step
 α -vector

$$R(b, OL) = +10(1 - b(TL)) + (-100)b(TL)$$

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

One-step utility

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



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

↑
 α -vector

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

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

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

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

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

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

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

Exercise 2: Crying Baby 1-Step Utility

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

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

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

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

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

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

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

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

Exercise 2: Crying Baby 1-Step Utility

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

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

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

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

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

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

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

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

$$\gamma = 0.9$$

Exercise 2: Crying Baby 1-Step Utility

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

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

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

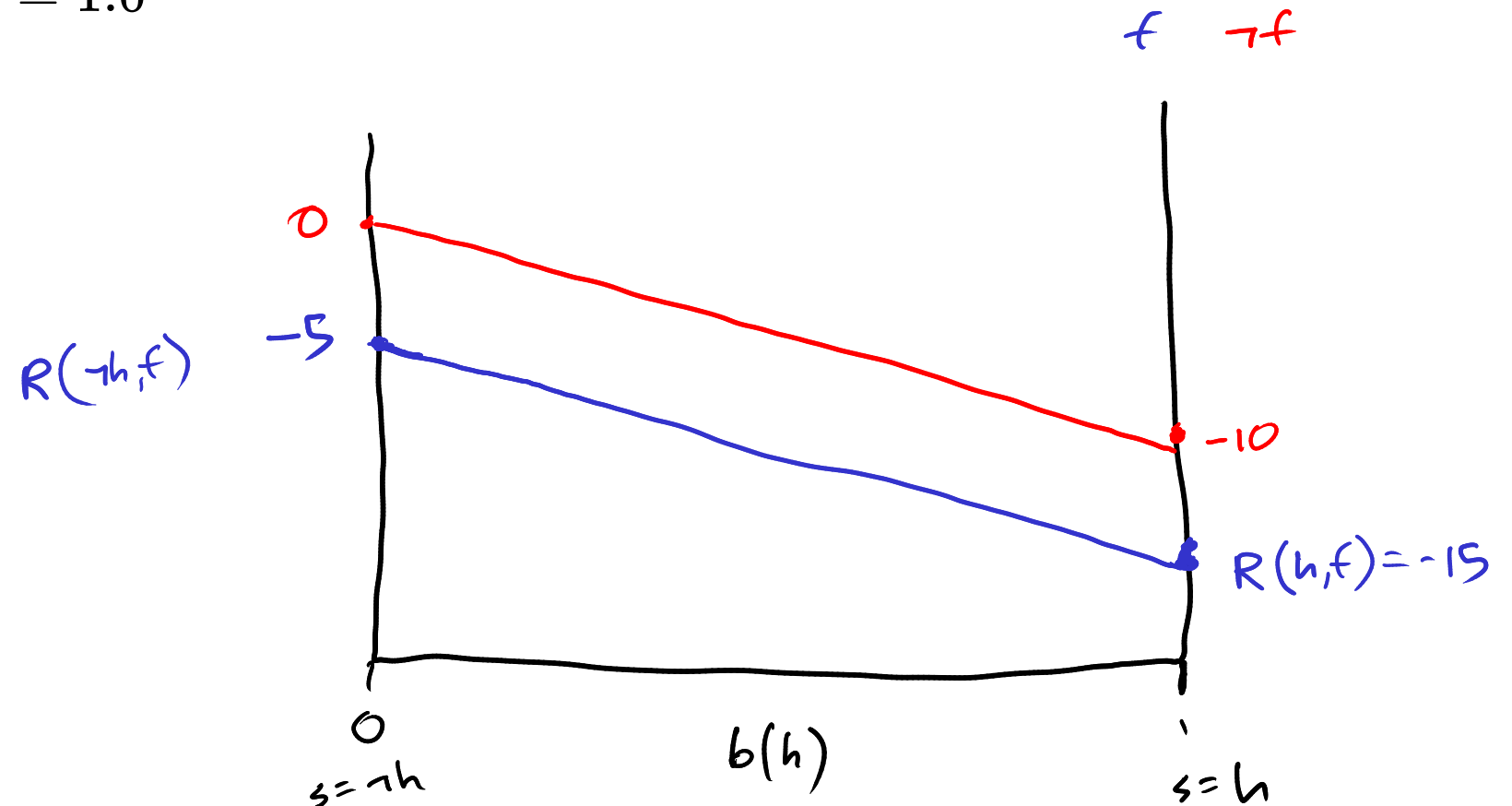
$$R(a) = \begin{cases} -5 & \text{if } a = f \\ 0 & \text{otherwise} \end{cases}$$

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

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

$$\gamma = 0.9$$

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

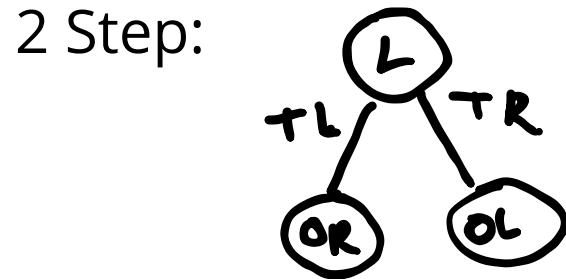
1 Step: (L) (OL) (OR)

2 Step:

Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

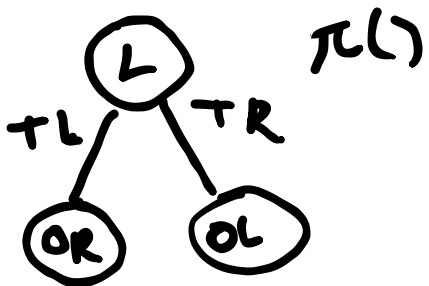
1 Step: (L) (OL) (OR)



Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

1 Step: 

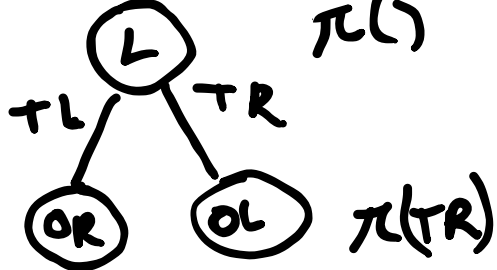
2 Step: 

Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

1 Step: (L) (OL) (OR)

2 Step:



```
graph TD; L((L)) -- TL --> OR1((OR)); L -- TR --> OL((OL));
```

$\pi(L)$

$\pi(TR)$

Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

1 Step: (L) (OL) (OR)

2 Step:

The first 2-step plan is a tree with root node (L) . It has two children: (OR) reached via edge TL , and (OL) reached via edge TR . The label $\pi(L)$ is placed to the right of the root, and $\pi(TR)$ is placed to the right of the (OL) node.

The second 2-step plan is a tree with root node (L) . It has two children: (OR) reached via edge TL , and (L) reached via edge TR .

Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

1 Step: (L) (OL) (OR)

2 Step:

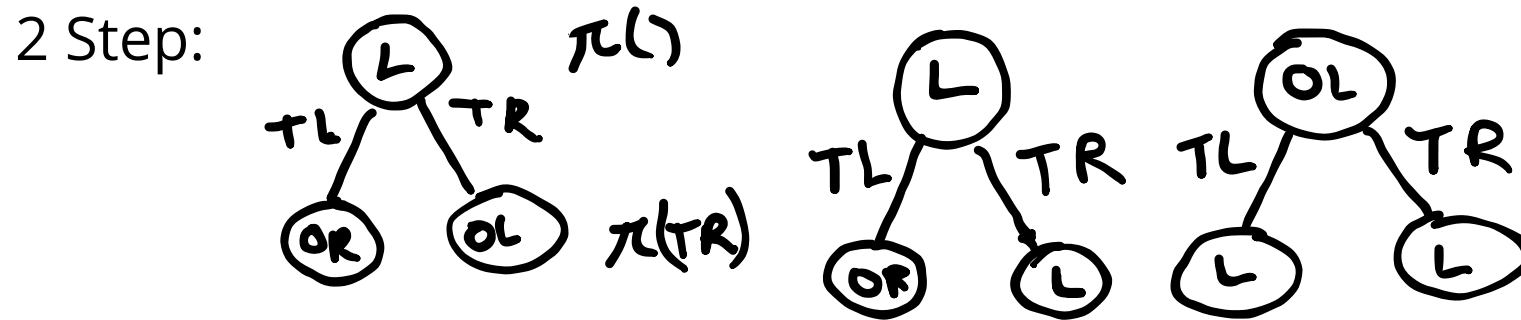
The diagrams show three conditional plans for 2 steps:

- Plan 1: Root node (L) with children (OR) (labeled TL) and (OL) (labeled TR). The label $\pi(L)$ is next to the root, and $\pi(TR)$ is next to the (OL) node.
- Plan 2: Root node (L) with children (OR) (labeled TL) and (L) (labeled TR).
- Plan 3: Root node (OL) with children (L) (labeled TL) and (L) (labeled TR).

Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

1 Step: (L) (OL) (OR)

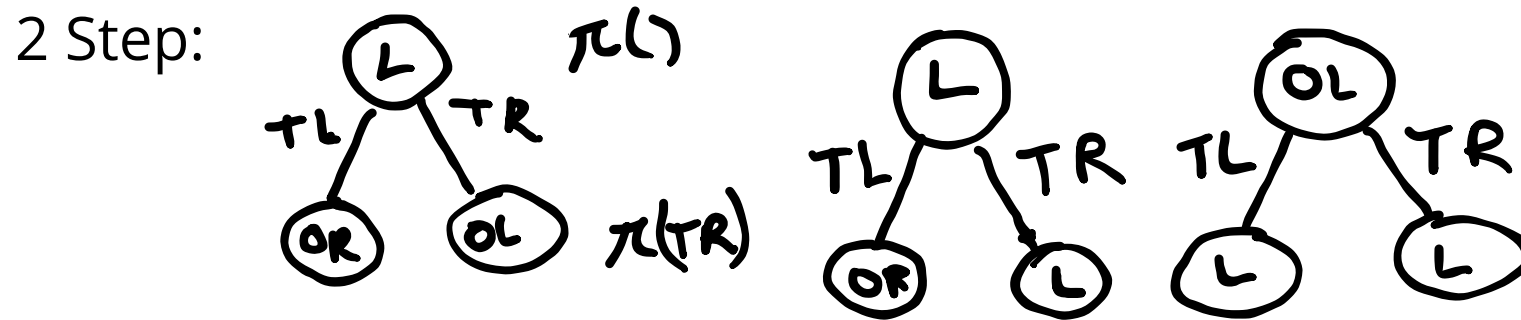


$$|A| \frac{(|O|^h - 1)}{(|O| - 1)}$$

Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

1 Step: (L) (OL) (OR)



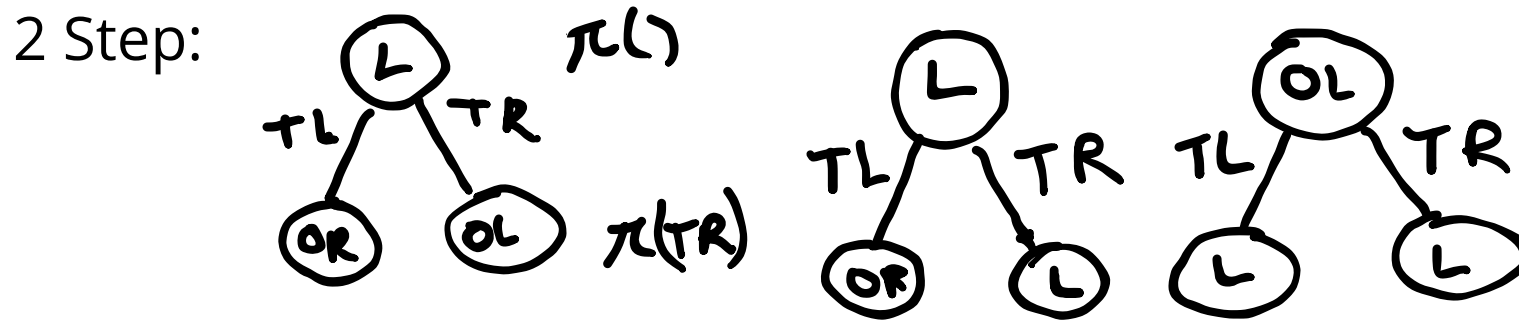
$$|A| \frac{(|O|^h - 1)}{(|O| - 1)}$$

27 two step plans!

Alpha Vectors for Conditional Plans

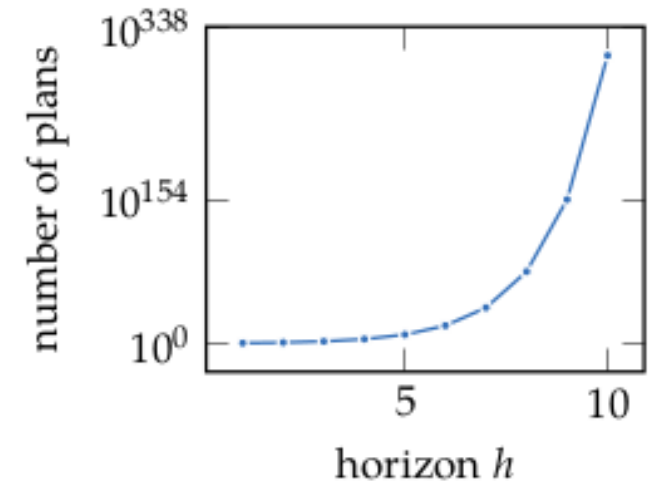
Conditional Plans: fixed-depth history-based policies

1 Step: (L) (OL) (OR)



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

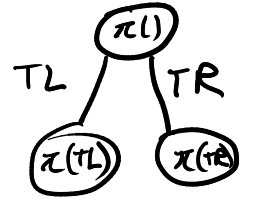
27 two step plans!



Alpha Vectors for Conditional Plans

Alpha Vectors for Conditional Plans

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

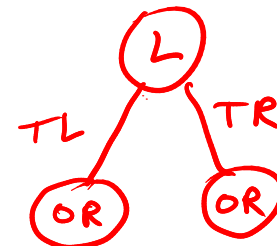
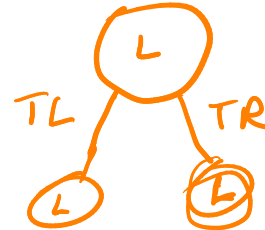
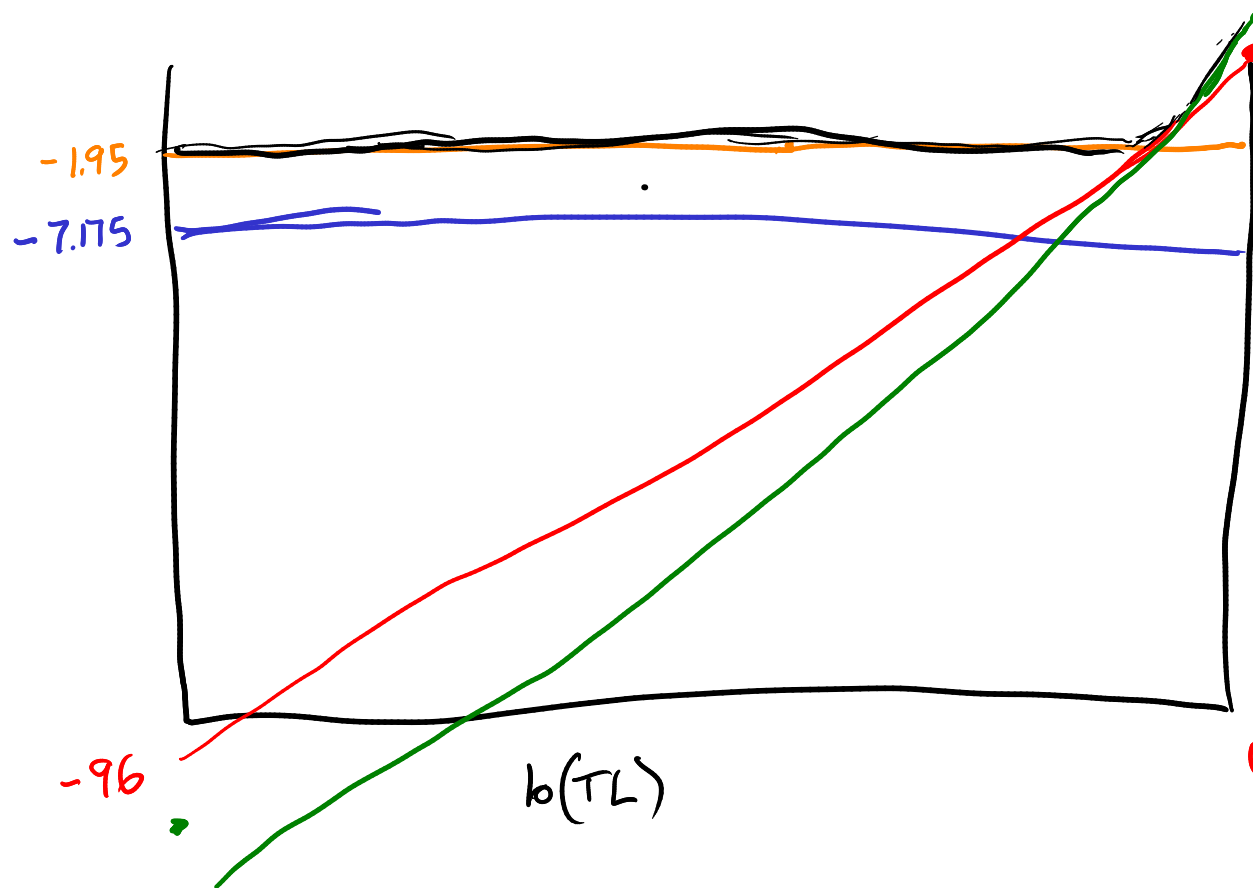


Alpha Vectors for Conditional Plans

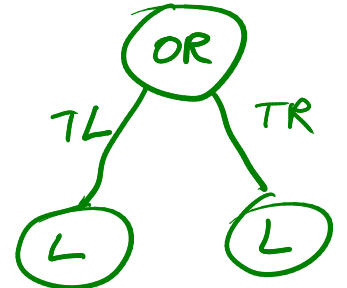
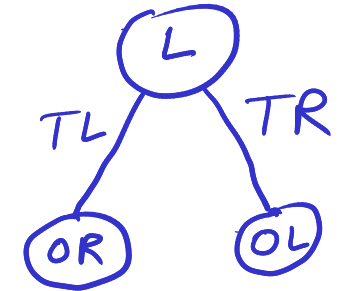
$$U^\pi(s) = R(s, \pi()) + \gamma \left[\sum_{s'} T(s' | s, \pi()) \sum_o \cancel{P}(o | \pi(), s') \underline{U^{\pi(o)}(s')} \right]$$

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

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



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



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

POMDP Value Functions

POMDP Value Functions

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

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

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

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

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

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

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

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

Exercise: 2-Step Crying Baby α Vectors

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

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

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

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

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

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

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

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

Exercise: 2-Step Crying Baby α Vectors

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

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

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

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

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

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

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

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

$$\gamma = 0.9$$

Exercise: 2-Step Crying Baby α Vectors

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

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

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

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

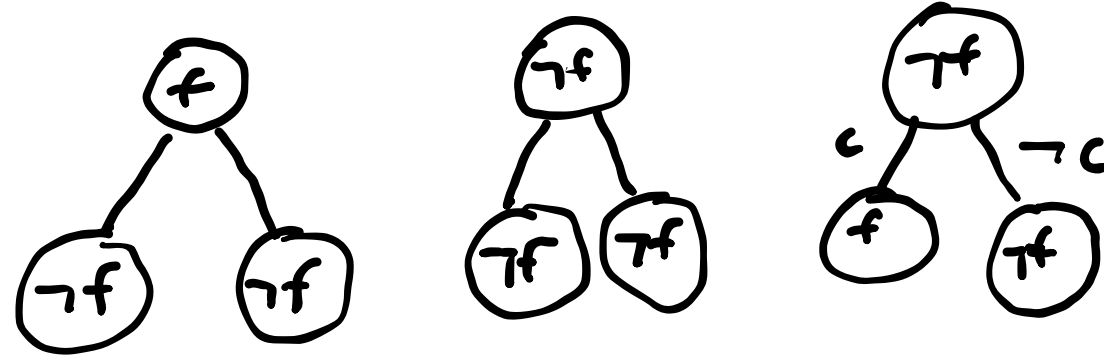
$$R(s) = \begin{cases} -10 & \text{if } s = h \\ 0 & \text{otherwise} \end{cases}$$

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

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

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

$$\gamma = 0.9$$



Exercise: 2-Step Crying Baby α Vectors

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

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

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

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

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

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

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

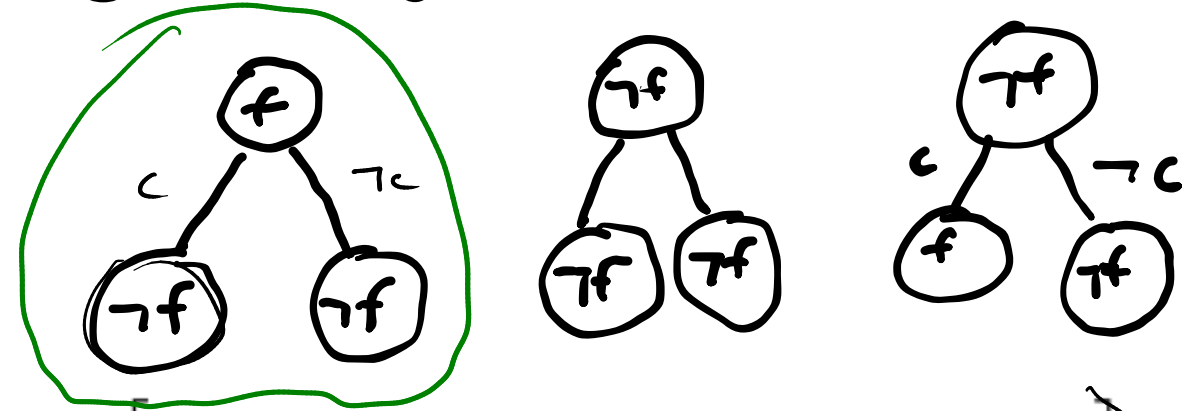
$$R(s) = \begin{cases} -10 & \text{if } s = h \\ 0 & \text{otherwise} \end{cases}$$

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

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

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

$$\gamma = 0.9$$



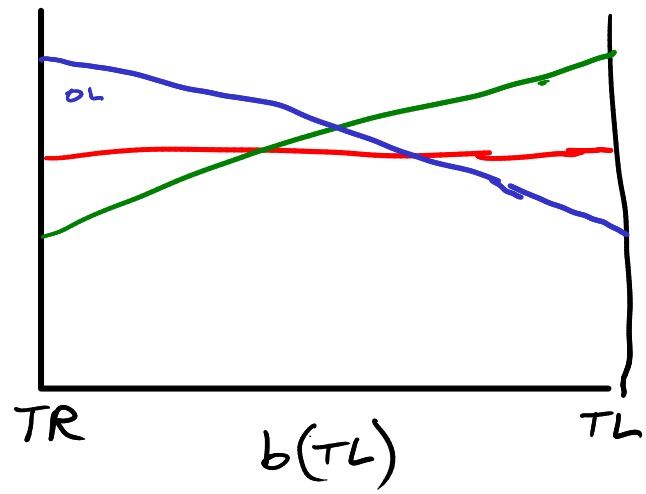
$$U^\pi(s) = R(s, \pi()) + \gamma \left[\sum_{s'} T(s' \mid s, \pi()) \sum_o O(o \mid \pi(), s') \underbrace{U^{\pi(o)}(s')}_{U^{\pi(c)}(h), U^{\pi(\neg c)}(h)} \right]$$

$$U^\pi(h) = R(h, f)$$

$$+ \gamma \left(\cancel{T(h \mid h, f)} (Z(c \mid f, h) U^{\pi(c)}(h) + Z(\neg c \mid f, h) U^{\pi(\neg c)}(h)) \right.$$

$$\left. + T(\neg h \mid h, f) (Z(c \mid f, h) U^{\pi(c)}(\neg h) + Z(\neg c \mid f, h) U^{\pi(\neg c)}(\neg h)) \right)$$

$$-15 + 0.9 (1.0 (0.8 \cdot 0 + 0.2 \cdot 0)) = -15$$

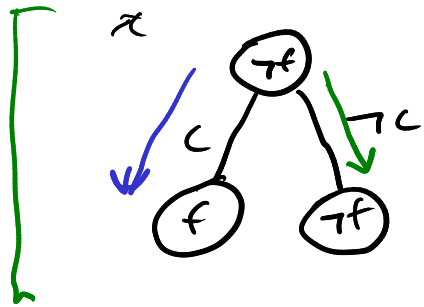


$$\alpha[s] = U^{\pi}(s)$$

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

$$\begin{aligned} U^f(h) &= -15 \\ U^f(\neg h) &= -5 \end{aligned}$$

$$\begin{aligned} U^{\neg f}(h) &= -10 \\ U^{\neg f}(\neg h) &= 0 \end{aligned}$$



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

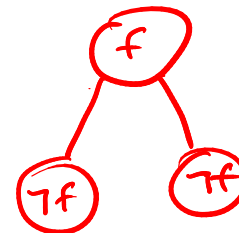
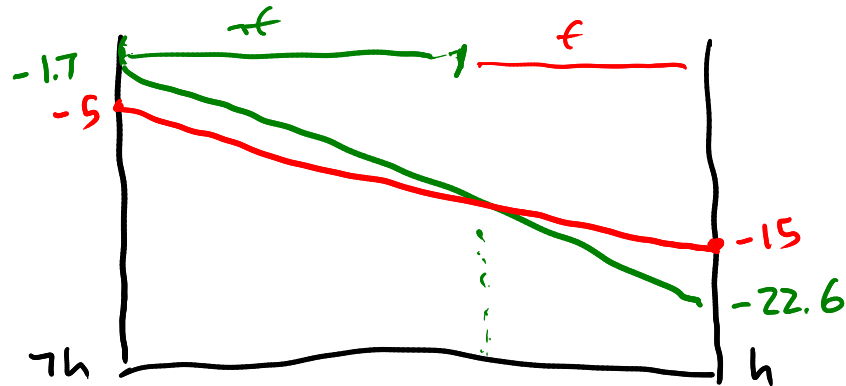
$$U^\pi(s) = R(s, \neg f) + \gamma \left(T(h|s, \neg f) \left(\underbrace{Z(c|\neg f, h)}_{0.8} U^f(h) + \underbrace{Z(\neg c|\neg f, h)}_{0.2} U^{\neg f}(h) \right) + T(\neg h|s, \neg f) \left(\underbrace{Z(c|\neg f, \neg h)}_{0.1} U^f(\neg h) + \underbrace{Z(\neg c|\neg f, \neg h)}_{0.9} U^{\neg f}(\neg h) \right) \right)$$

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

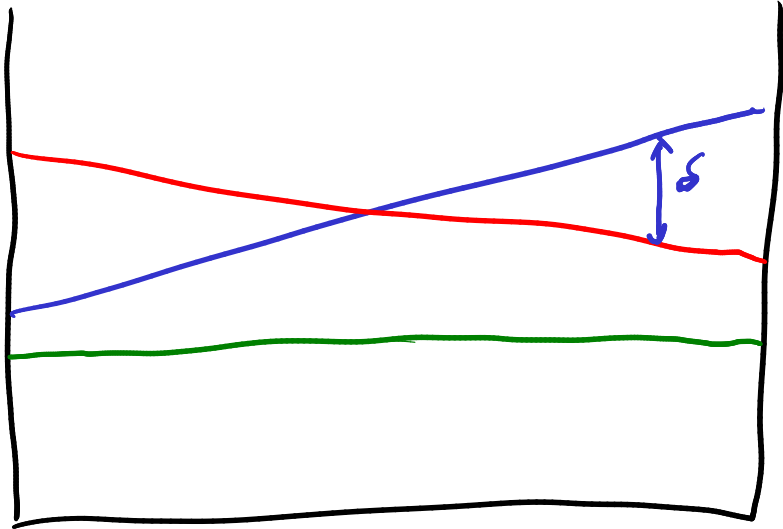
$$U^\pi(h) = -10 + 0.9 (1.0 \cdot -14 + 0 \cdot -0.5) = -22.6$$

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

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



α -Vector Pruning



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

- 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 $n \in 1 \dots d$

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

Prune Γ^n

$$\Gamma^0 = \{\}$$

$\Gamma^1 = \alpha$ vectors
for 1 step conditional
plans

$$\{\oplus, \boxed{\neg \oplus}\}$$

$\Gamma^2 = \alpha$ vectors for
2-step conditional
plans

$$\left\{ \begin{array}{c} \oplus \\ \swarrow \quad \searrow \\ \oplus \quad \oplus \end{array} , \begin{array}{c} \oplus \\ \swarrow \quad \searrow \\ \neg \oplus \quad \oplus \end{array} , \dots \right\} \quad \Gamma^2$$

$$\Gamma^3$$

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