Recap

Recap

(S, A, T, R, γ)

- POMDP

Recap

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

Recap

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

Recap

- POMDP
- Belief Updates

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

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

Recap

- POMDP
- Belief Updates

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

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

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

Recap

- POMDP
- Belief Updates

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

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

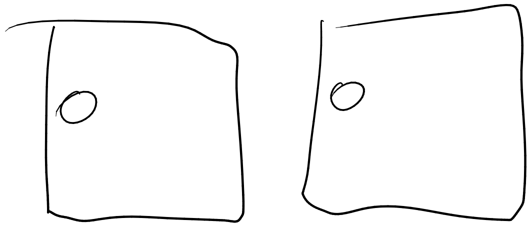
$$b' = \tau(b, a, o)$$

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

Guiding Quesiton

How do we calculate the optimal action in a POMDP?

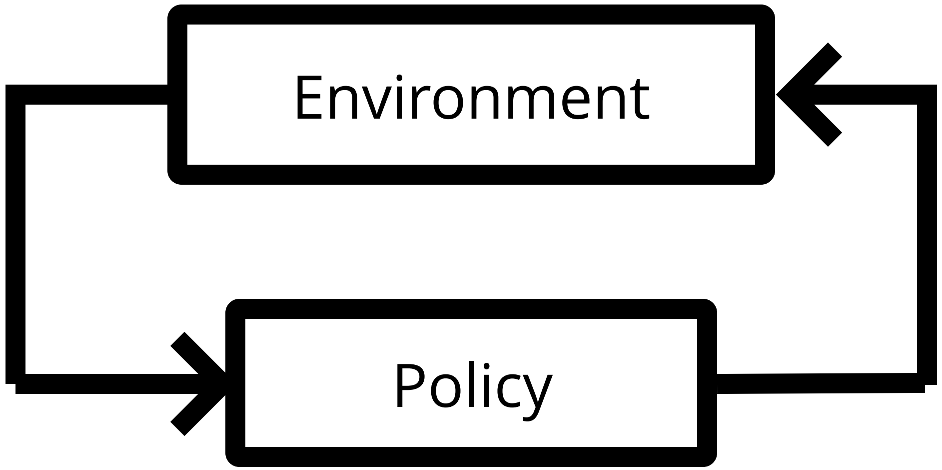
Solving the Tiger POMDP



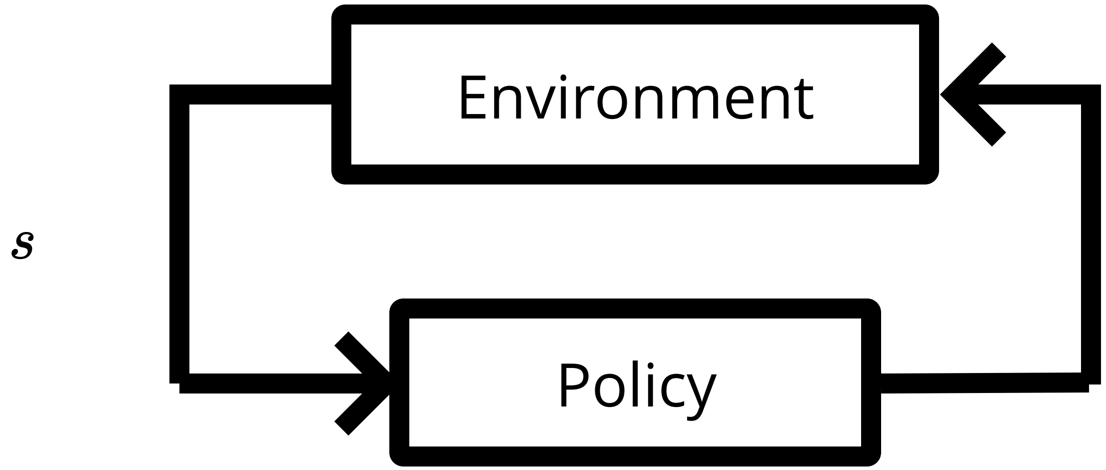
Step 1 = calculate b

Step 2 = take action with highest value

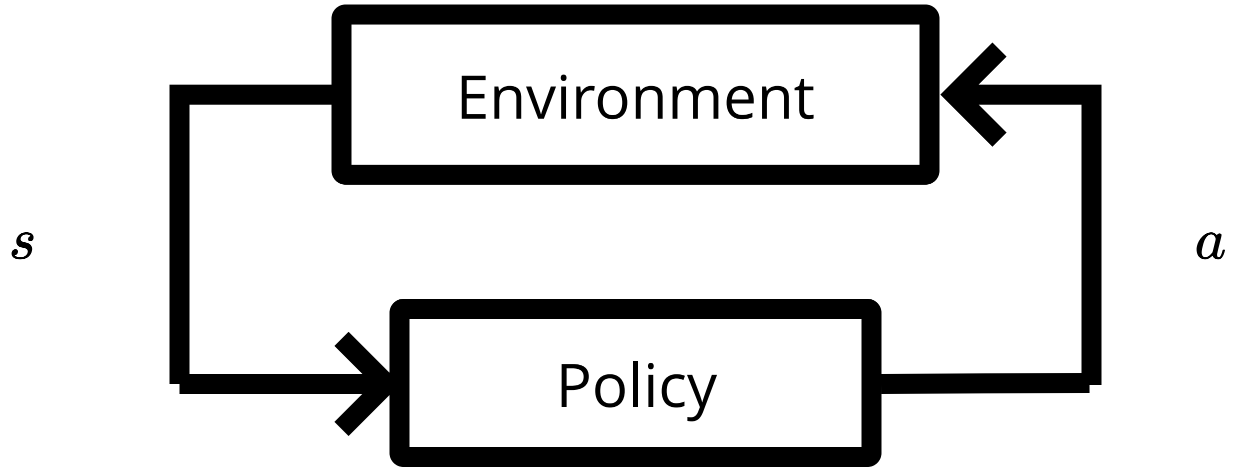
MDP Sense-Plan-Act Loop



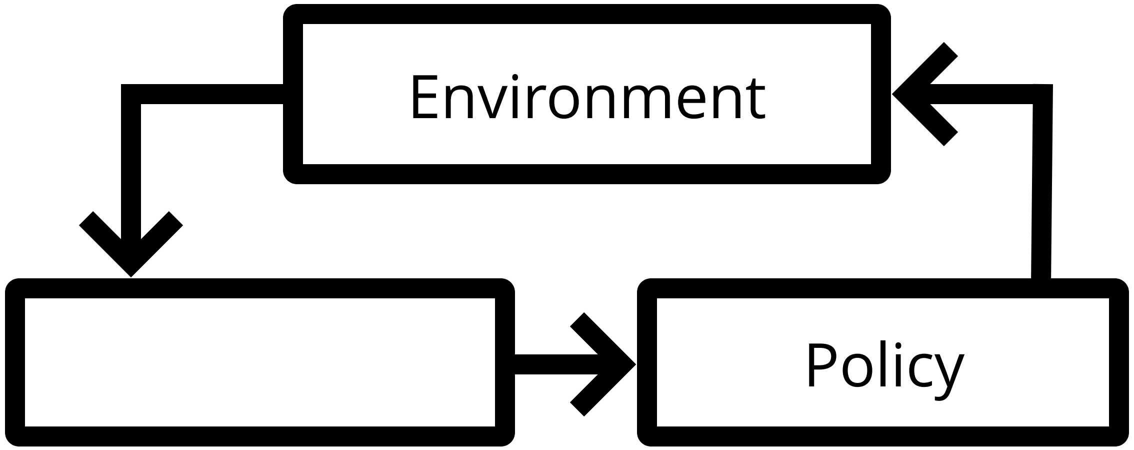
MDP Sense-Plan-Act Loop



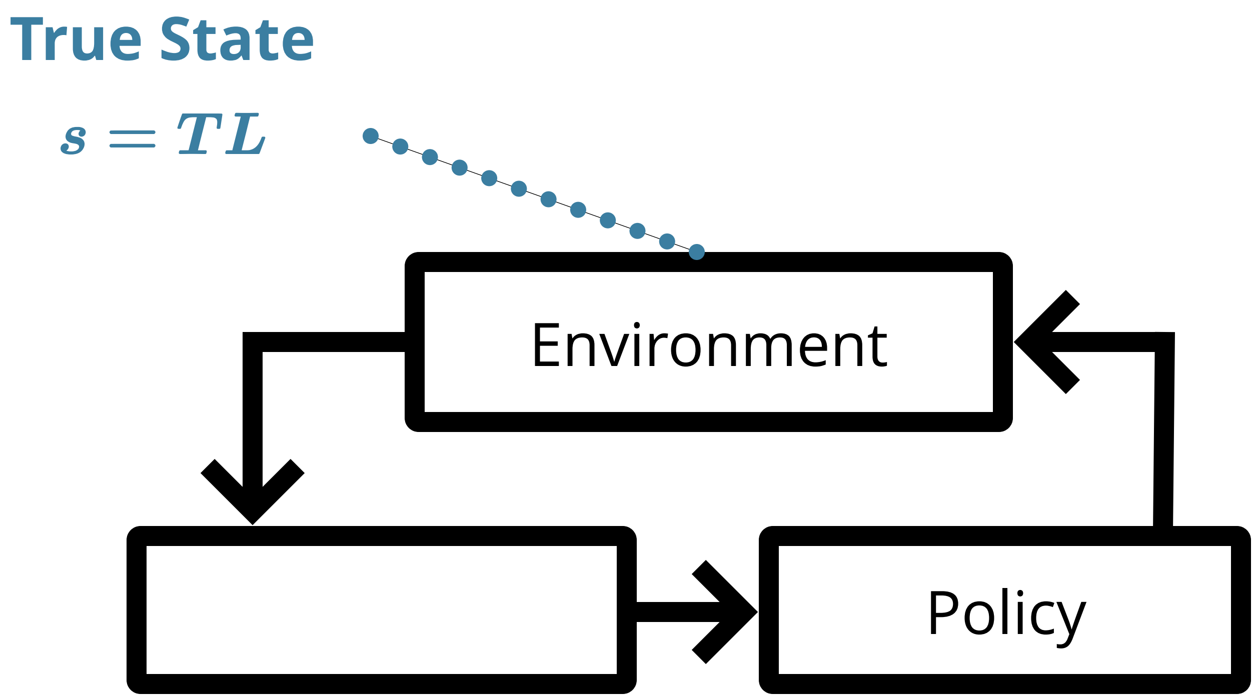
MDP Sense-Plan-Act Loop



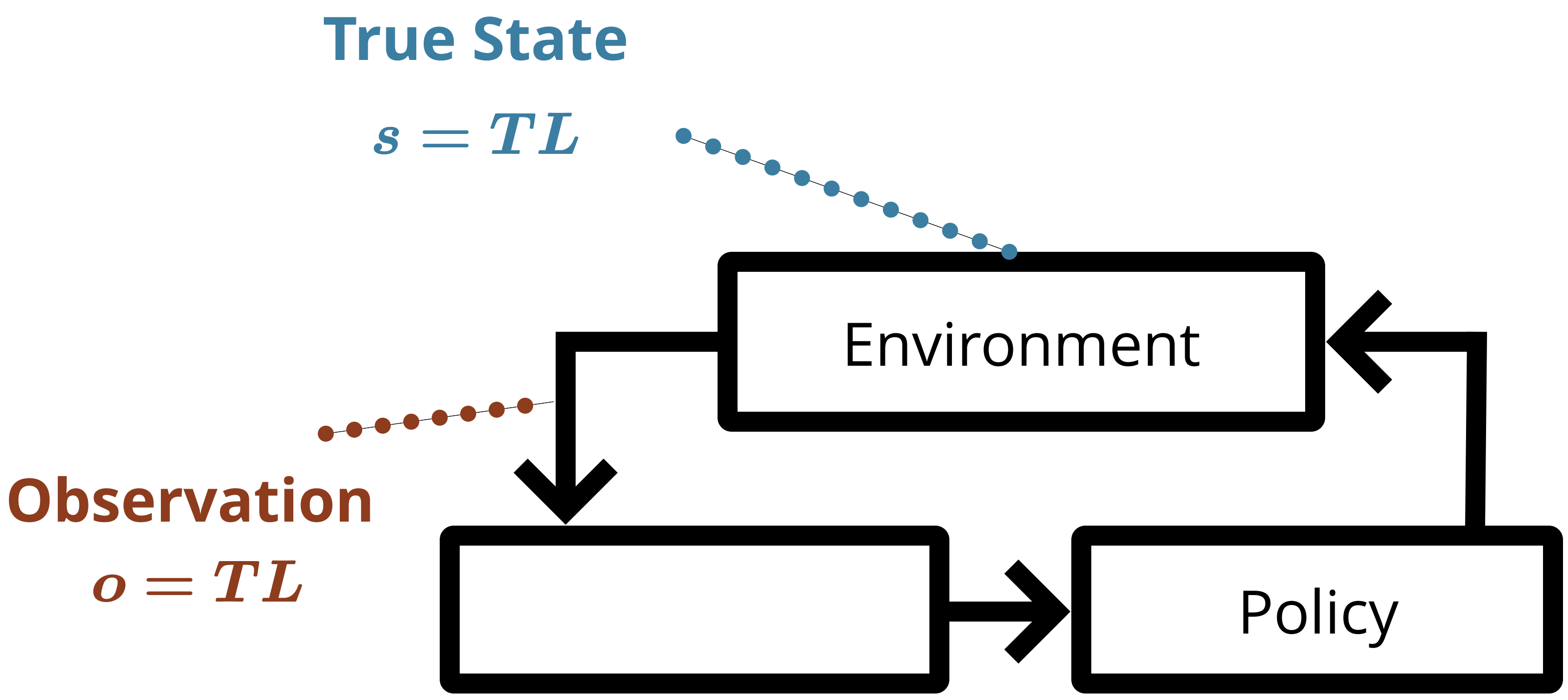
POMDP Sense-Plan-Act Loop



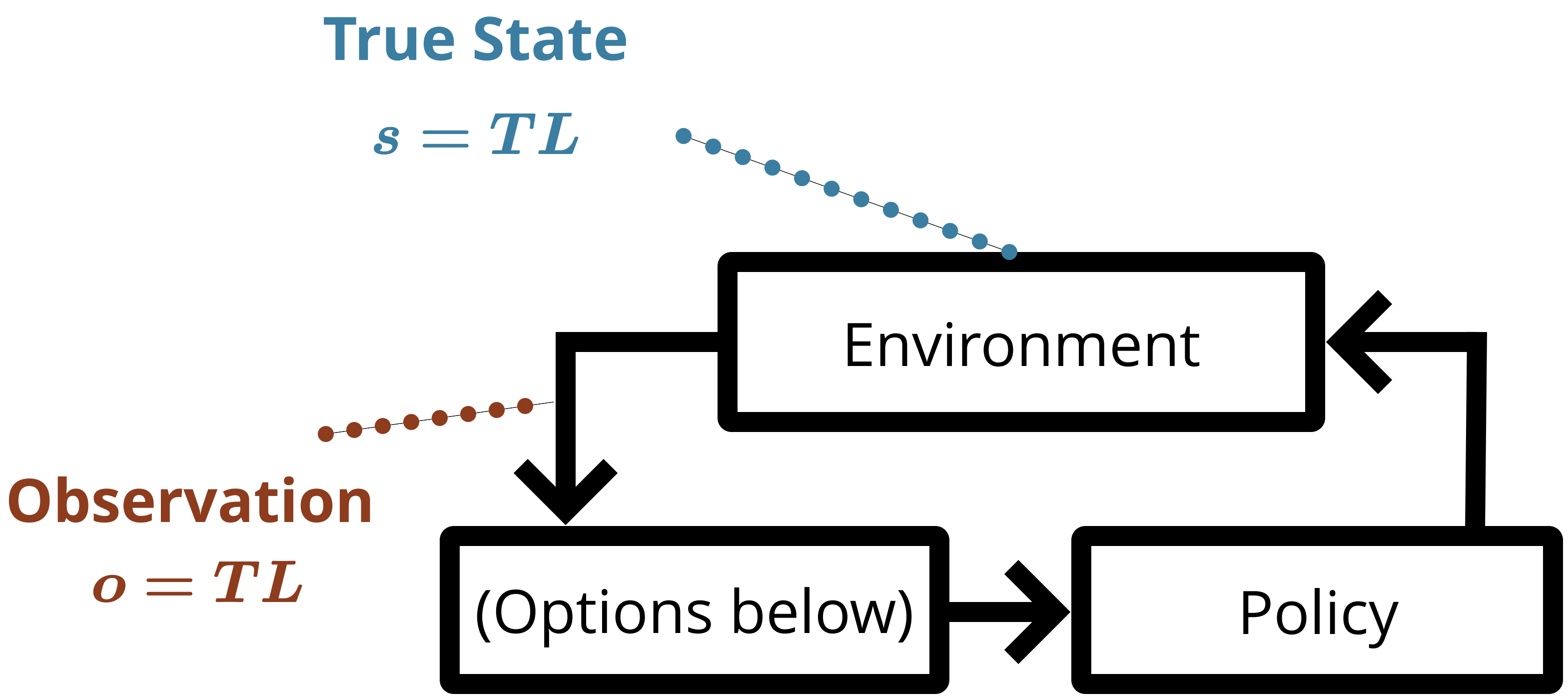
POMDP Sense-Plan-Act Loop



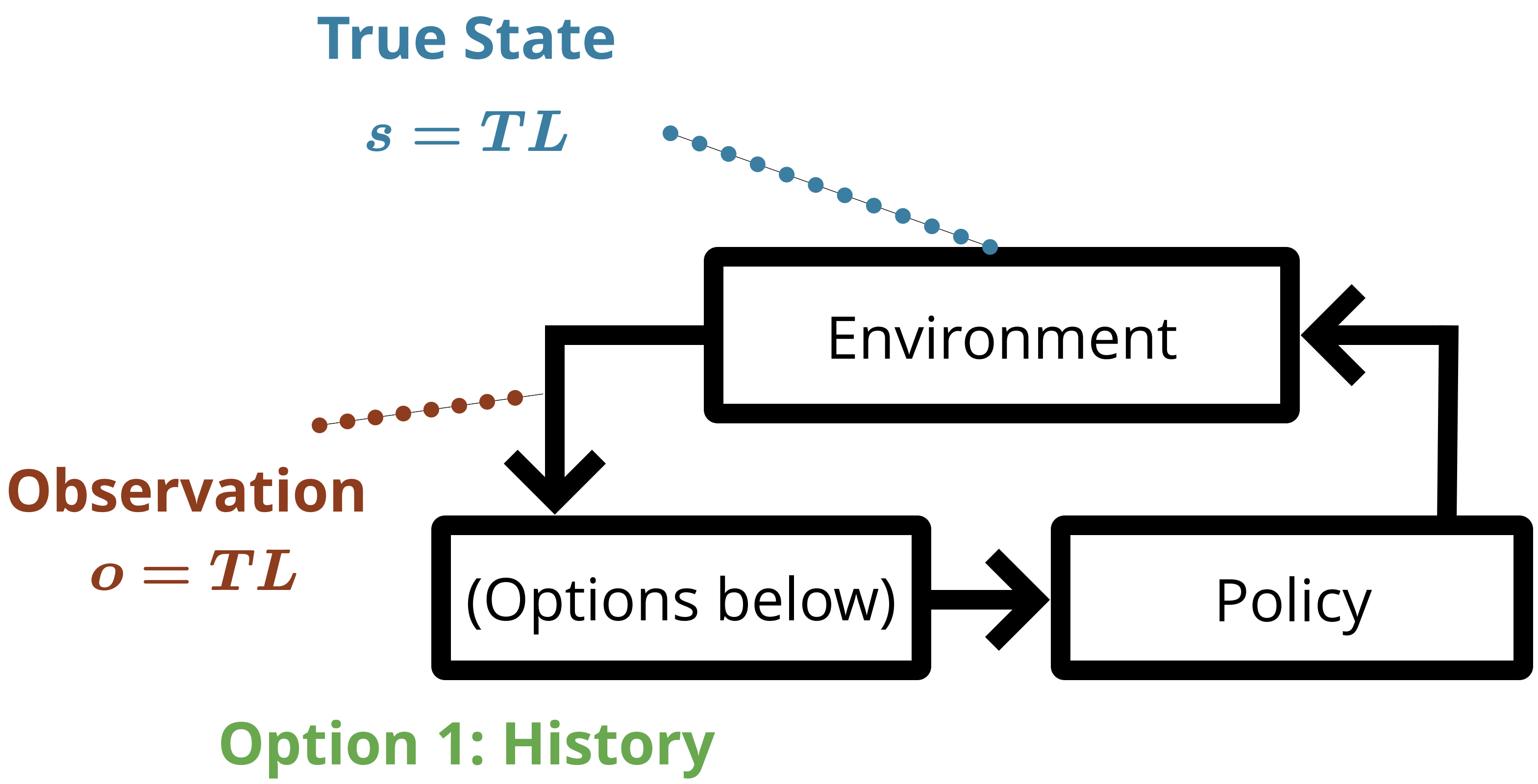
POMDP Sense-Plan-Act Loop



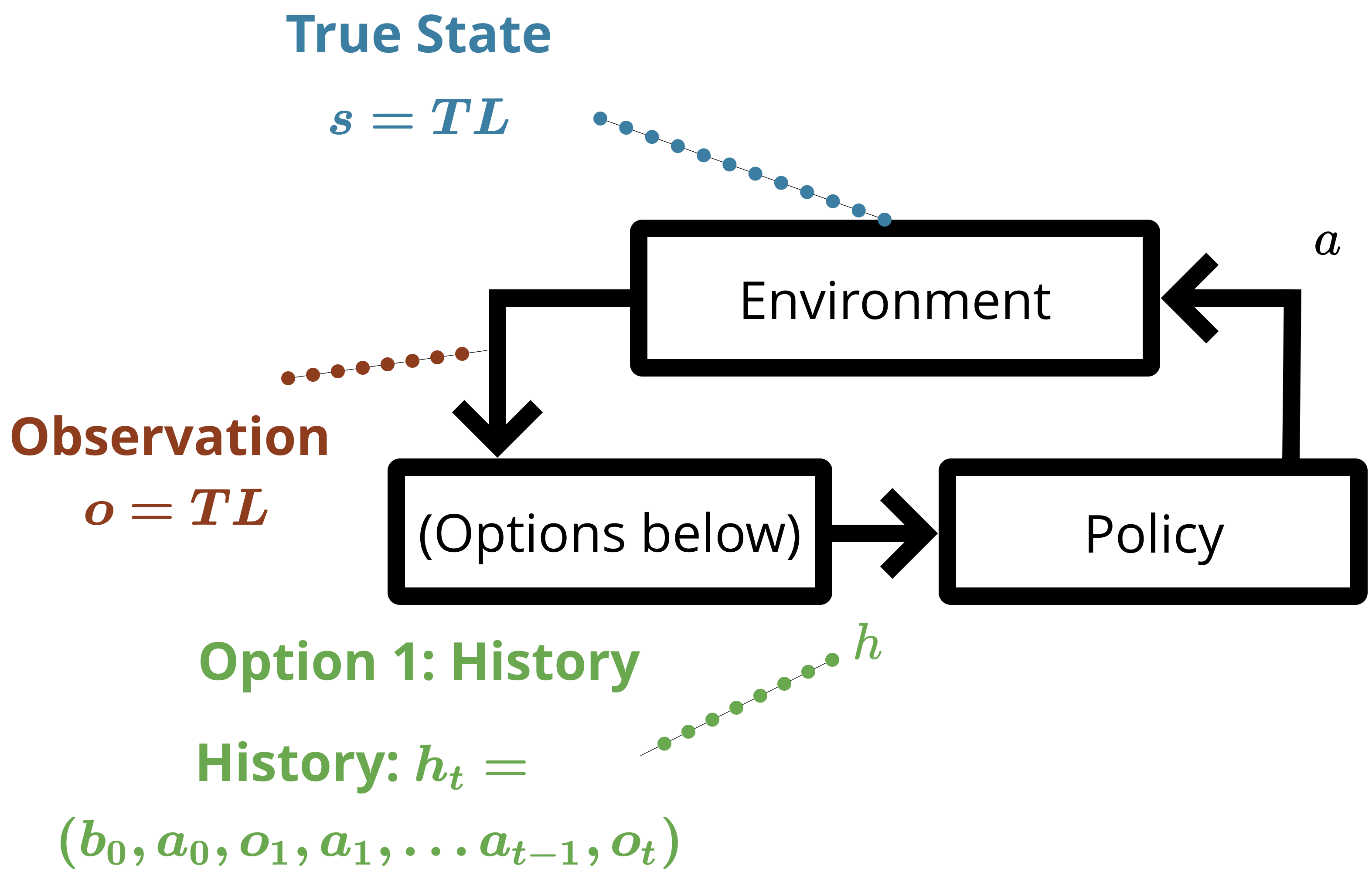
POMDP Sense-Plan-Act Loop



POMDP Sense-Plan-Act Loop



POMDP Sense-Plan-Act Loop



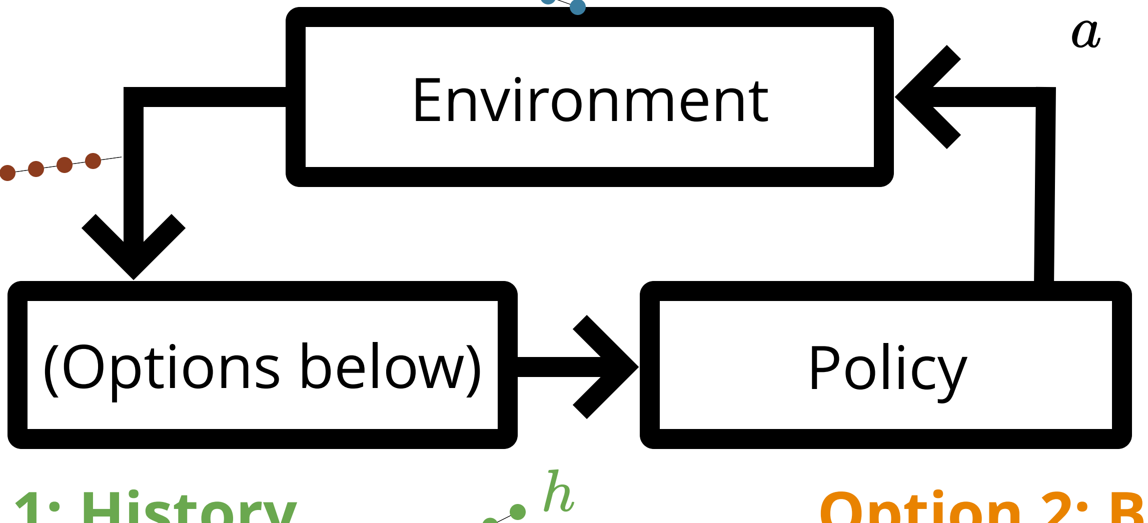
POMDP Sense-Plan-Act Loop

True State

$s = TL$

Observation

$o = TL$

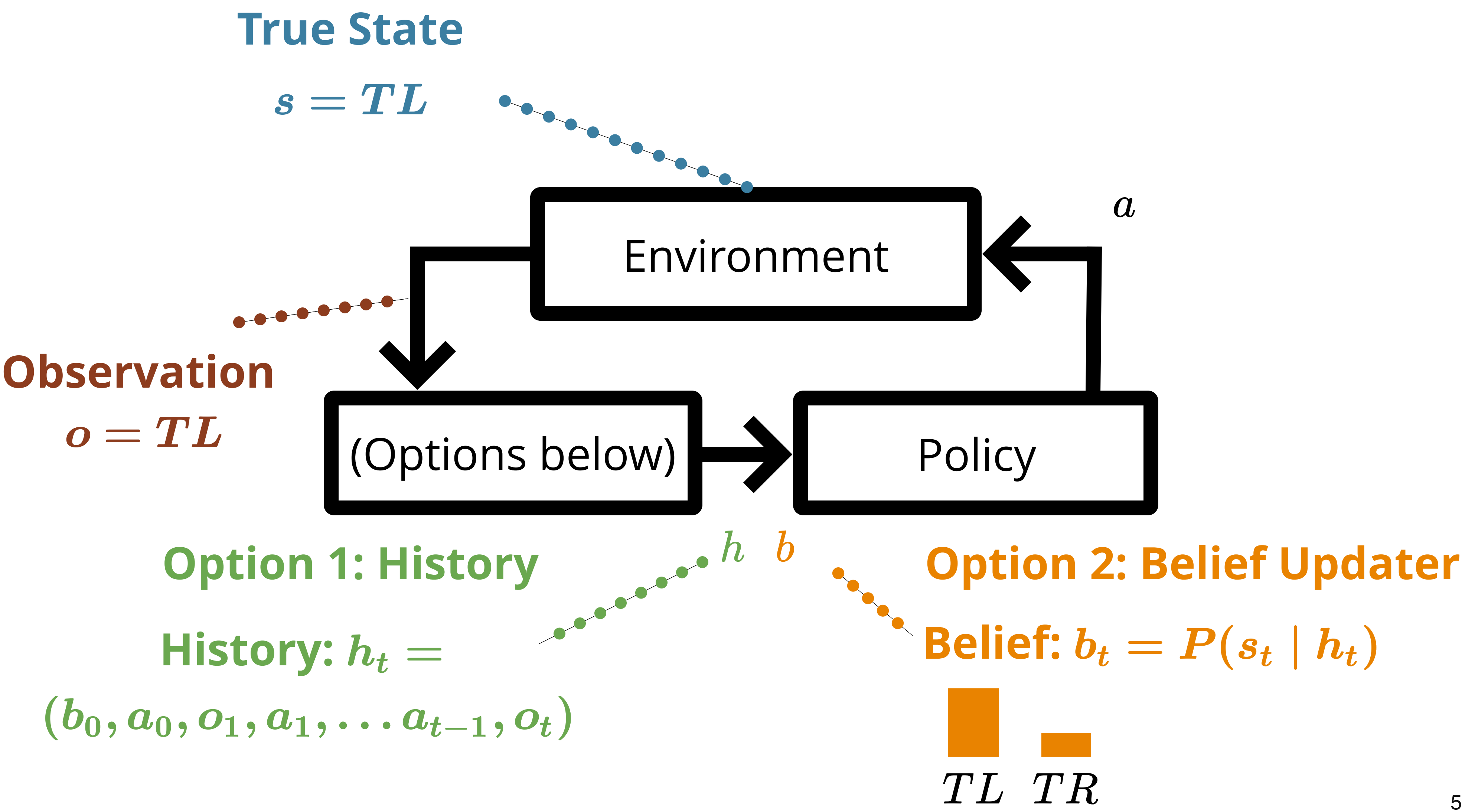


Option 1: History

History: $h_t = (b_0, a_0, o_1, a_1, \dots a_{t-1}, o_t)$

Option 2: Belief Updater

POMDP Sense-Plan-Act Loop



Belief Dynamics

Belief Dynamics

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

Belief Dynamics

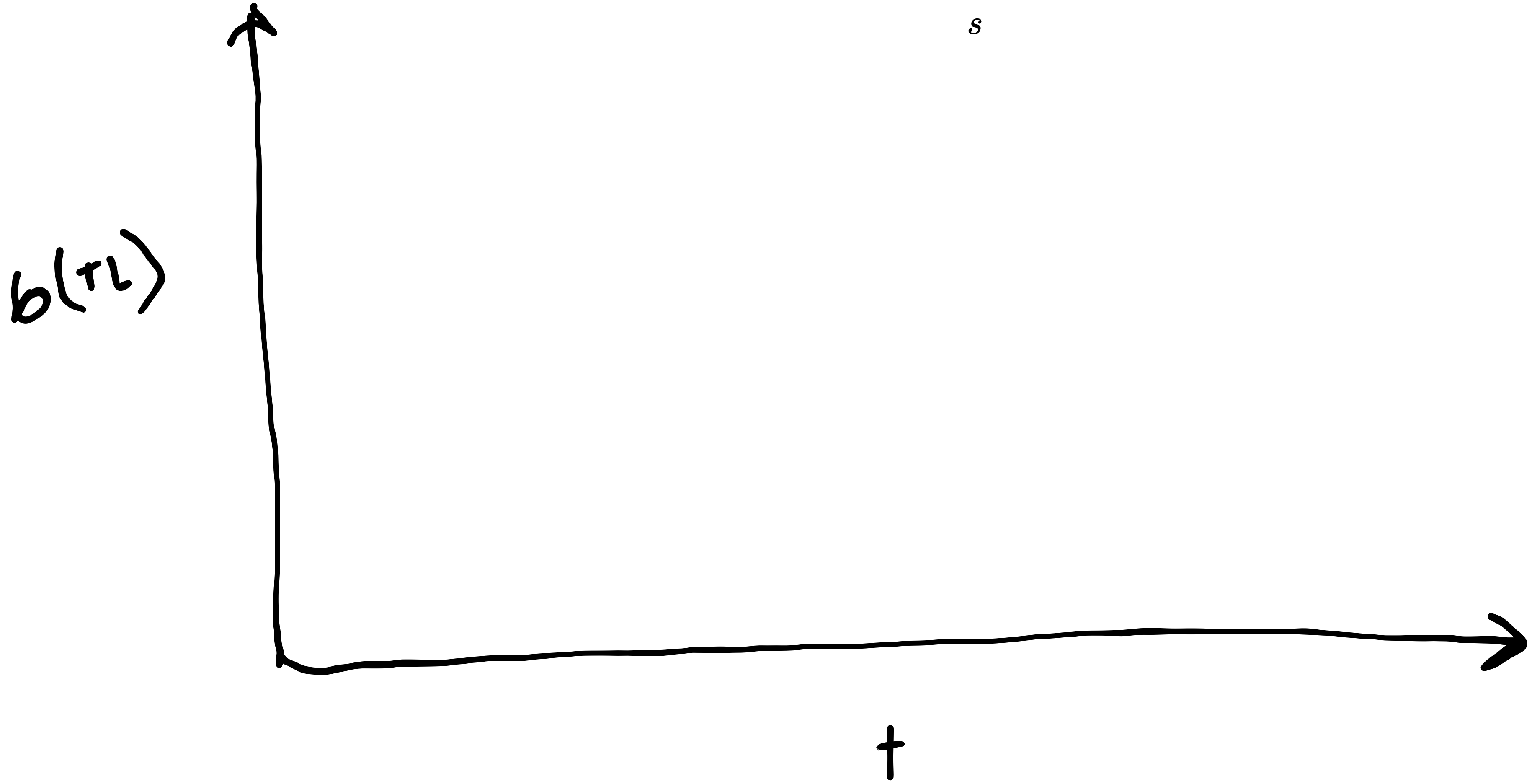
$$b'(s') \propto Z(o \mid a, s') \sum_s T(s' \mid s, a)\, b(s)$$

b(TL)

t

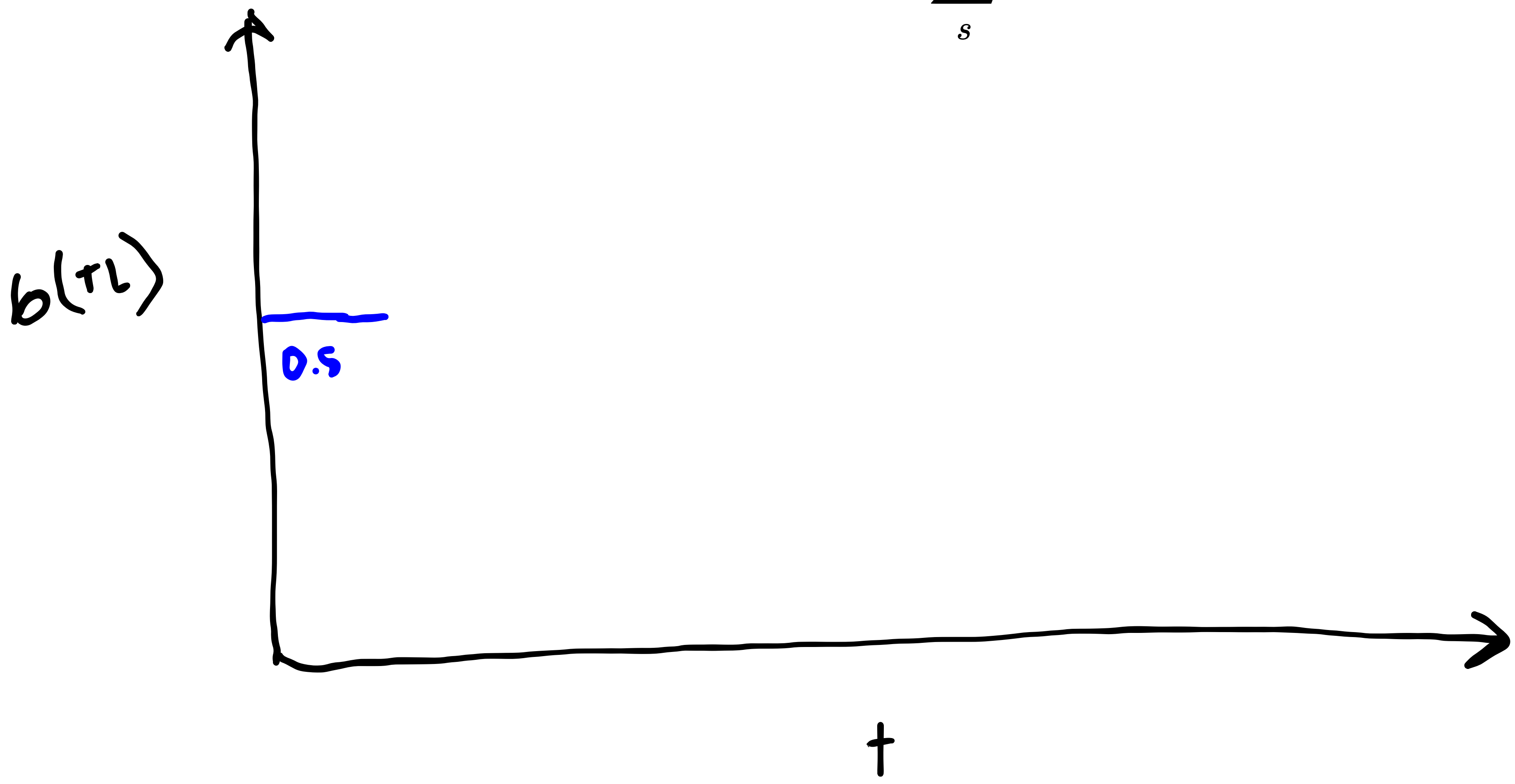
Belief Dynamics

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



Belief Dynamics

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



Belief Dynamics

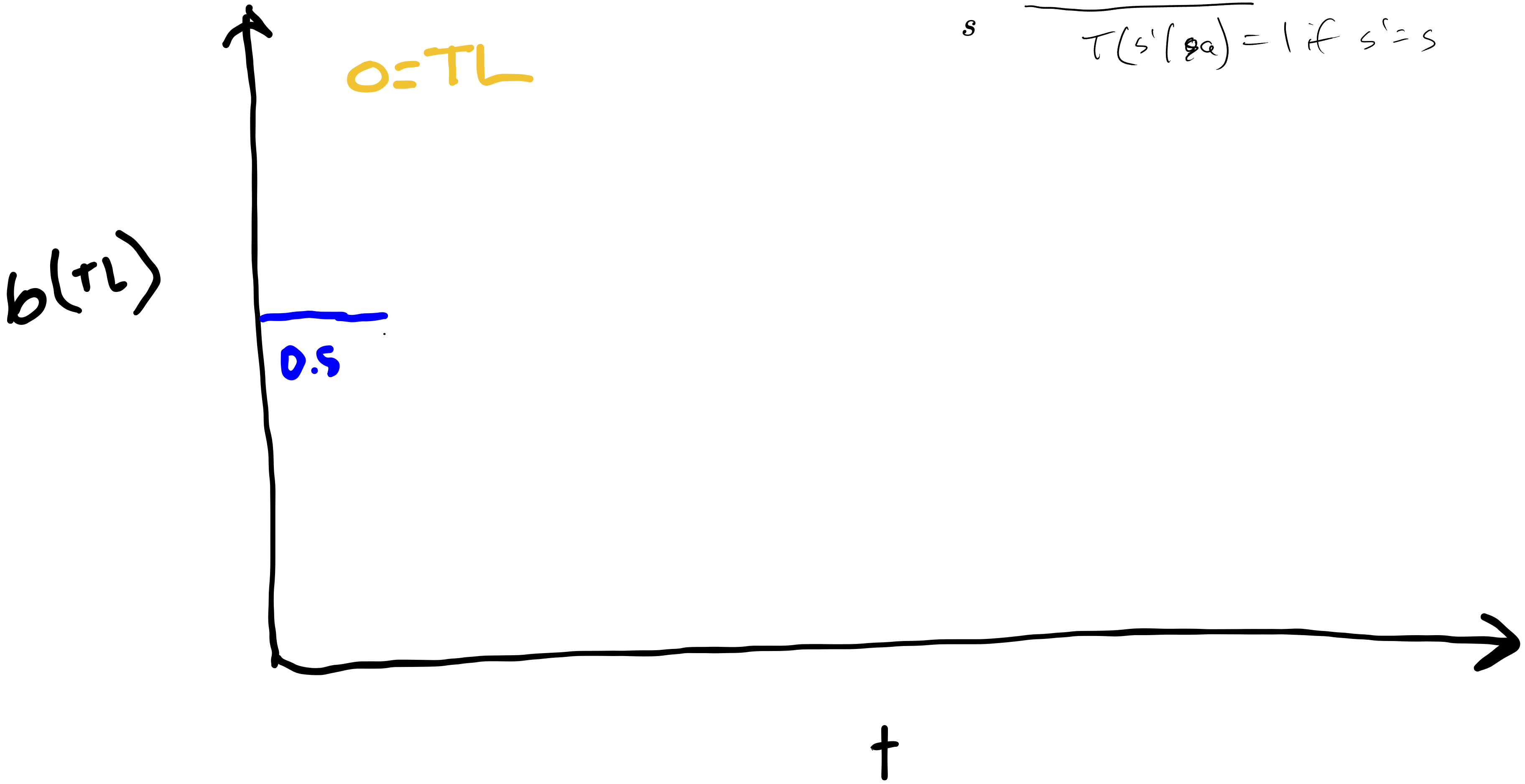
$$b'(s') \propto Z(o \mid a, s') \sum_s \underbrace{T(s' \mid s, a)\, b(s)}$$

$T(s' \mid s, a) = 1$ if $s' = s$

$b'(TL) \propto 0.85 \cdot 0.5$

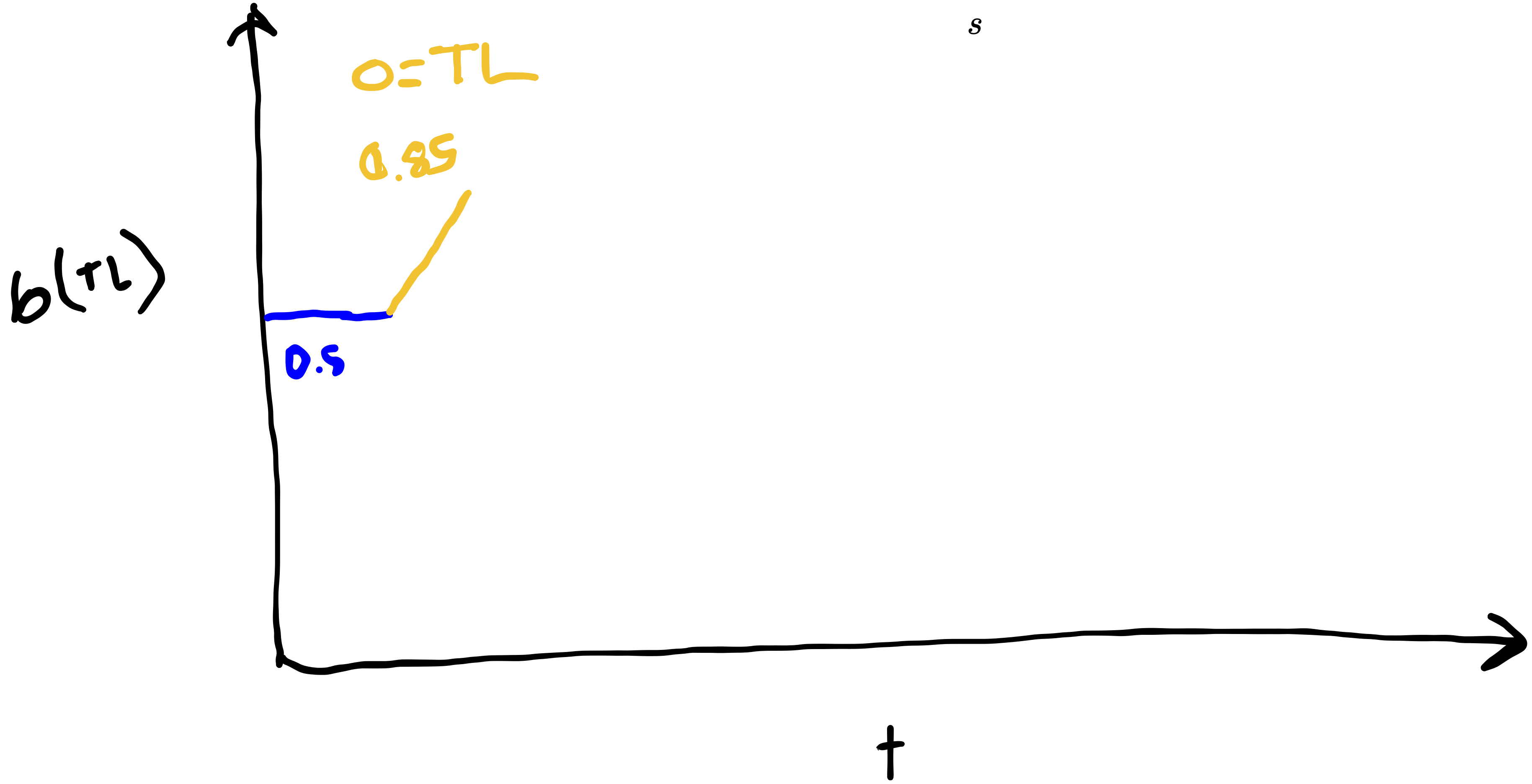
$b'(TR) \propto 0.15 \cdot 0.5$

o



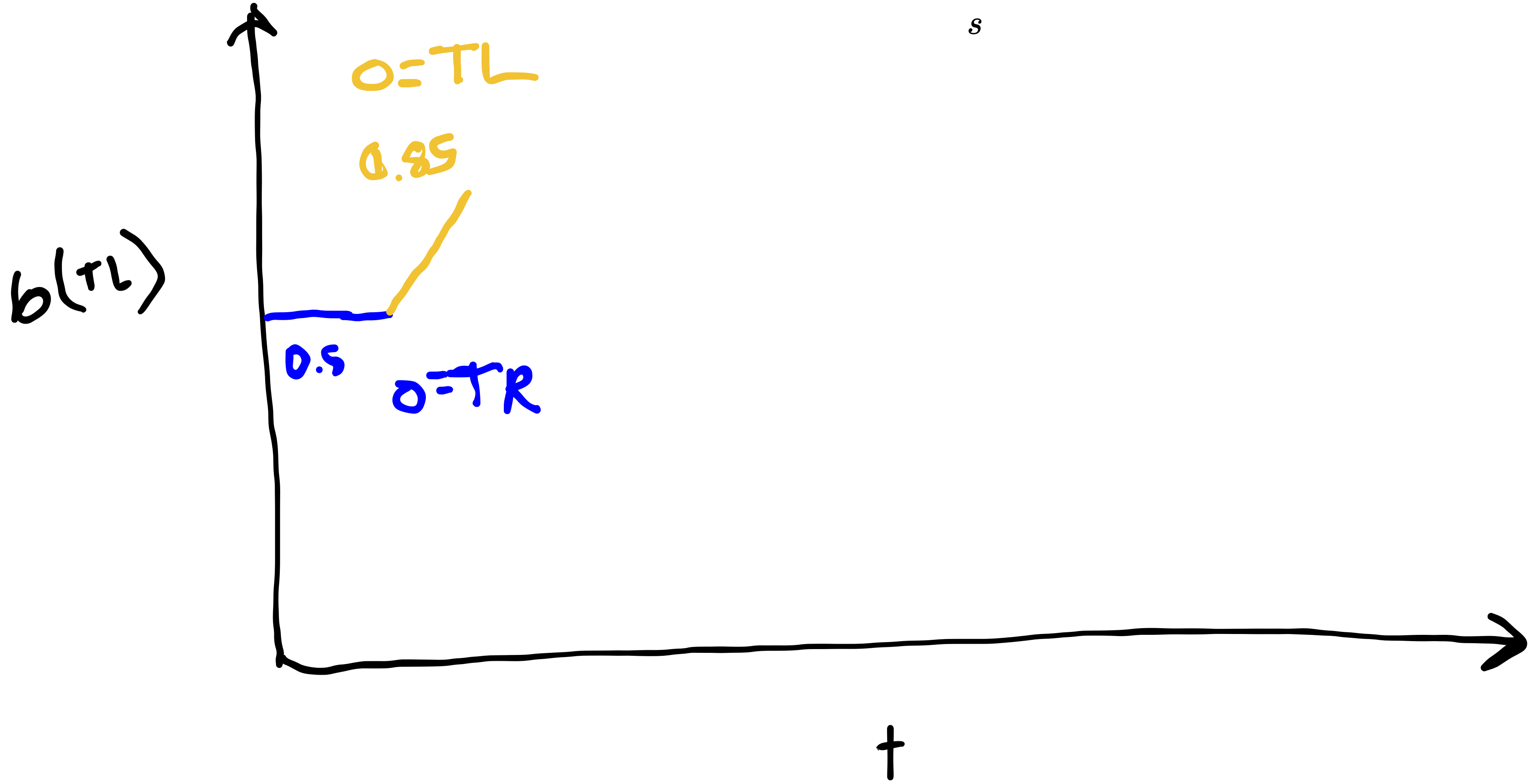
Belief Dynamics

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



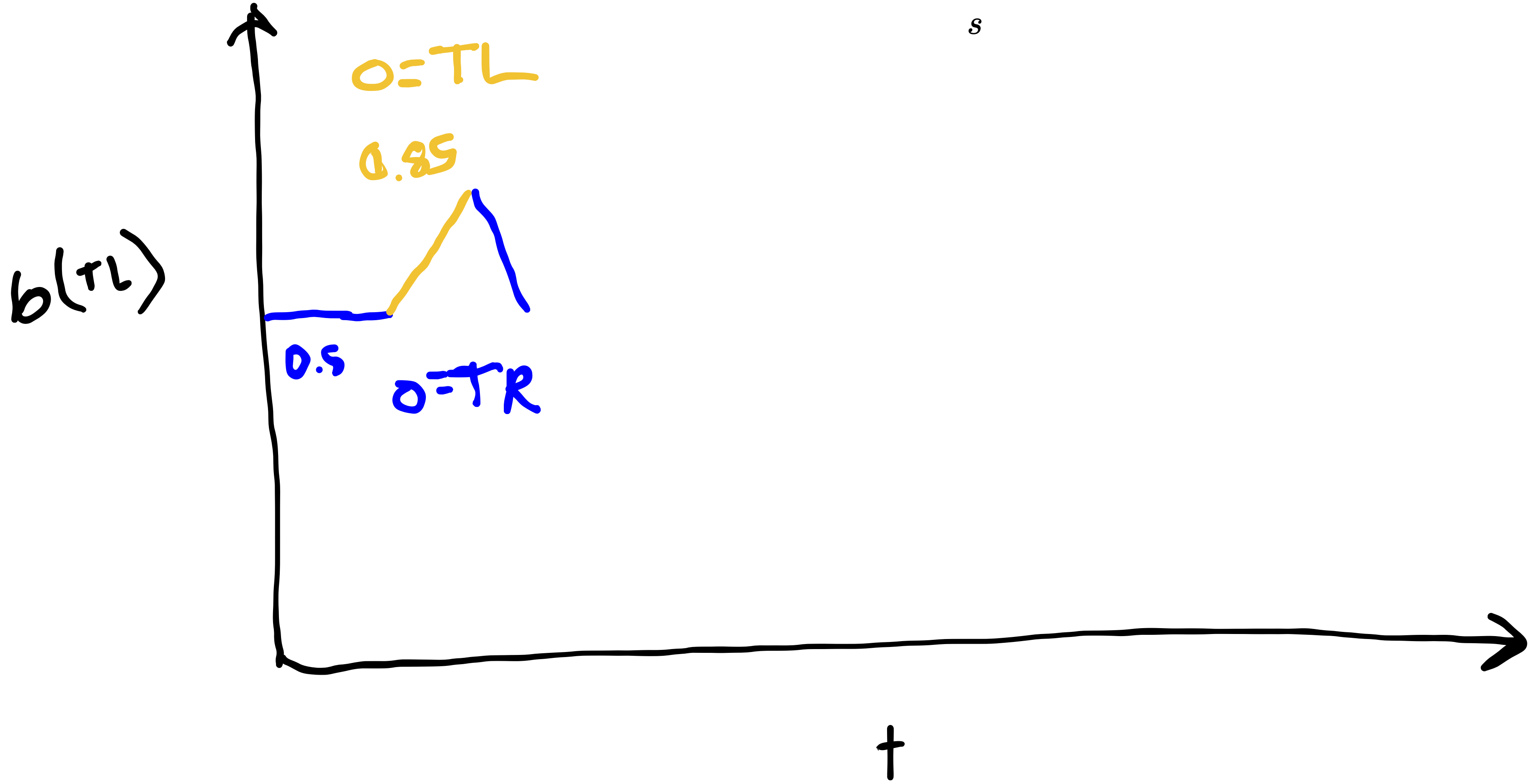
Belief Dynamics

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



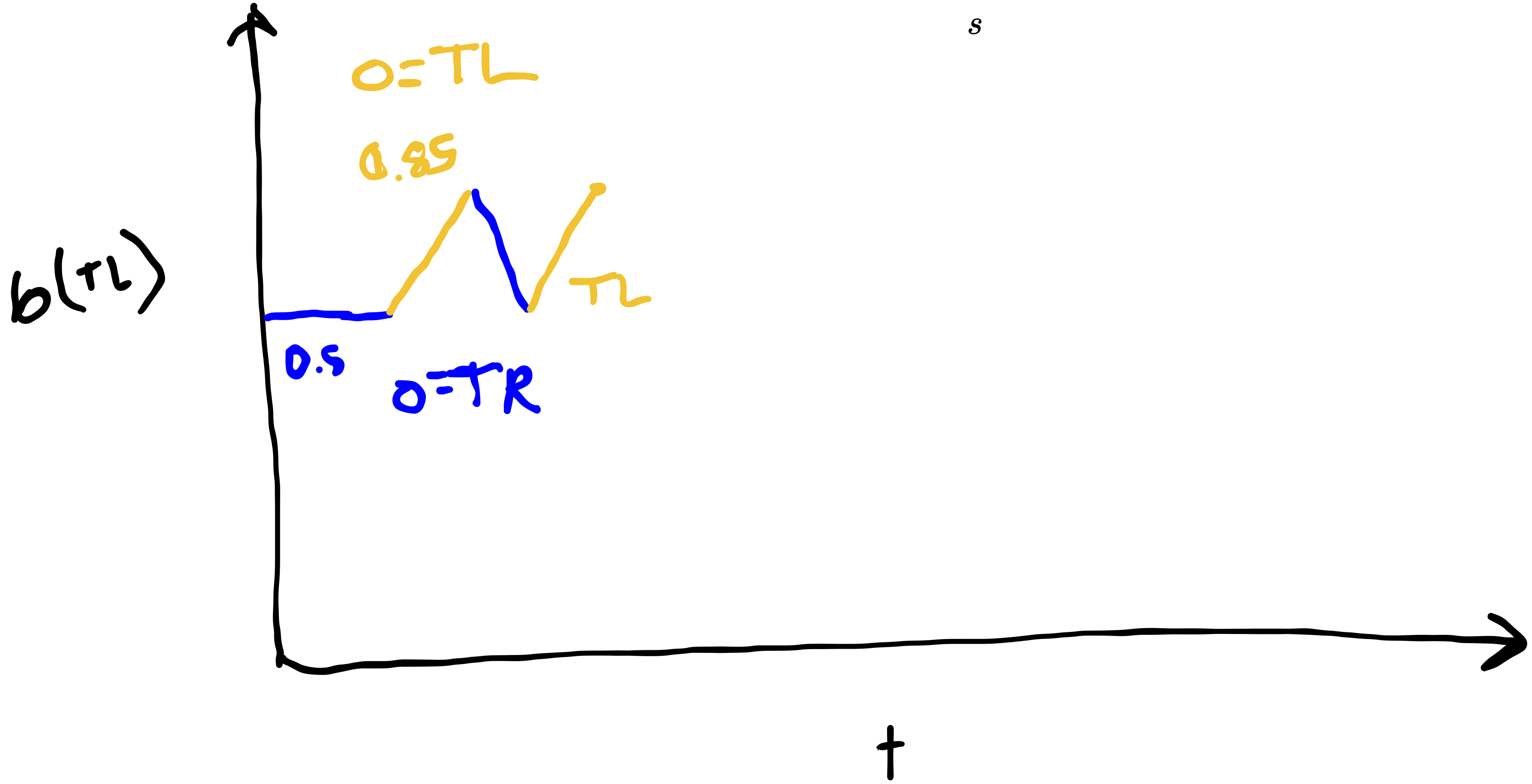
Belief Dynamics

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



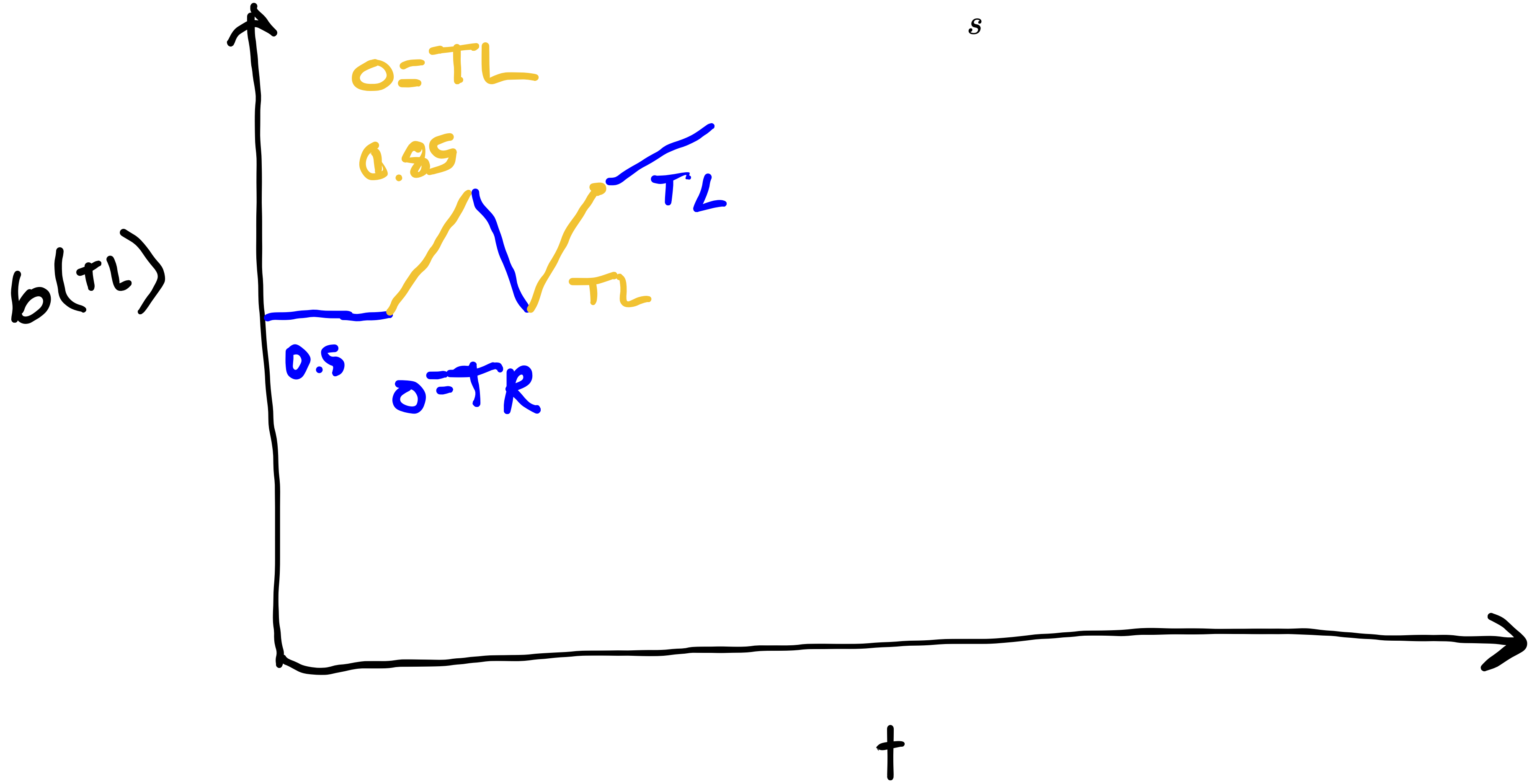
Belief Dynamics

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



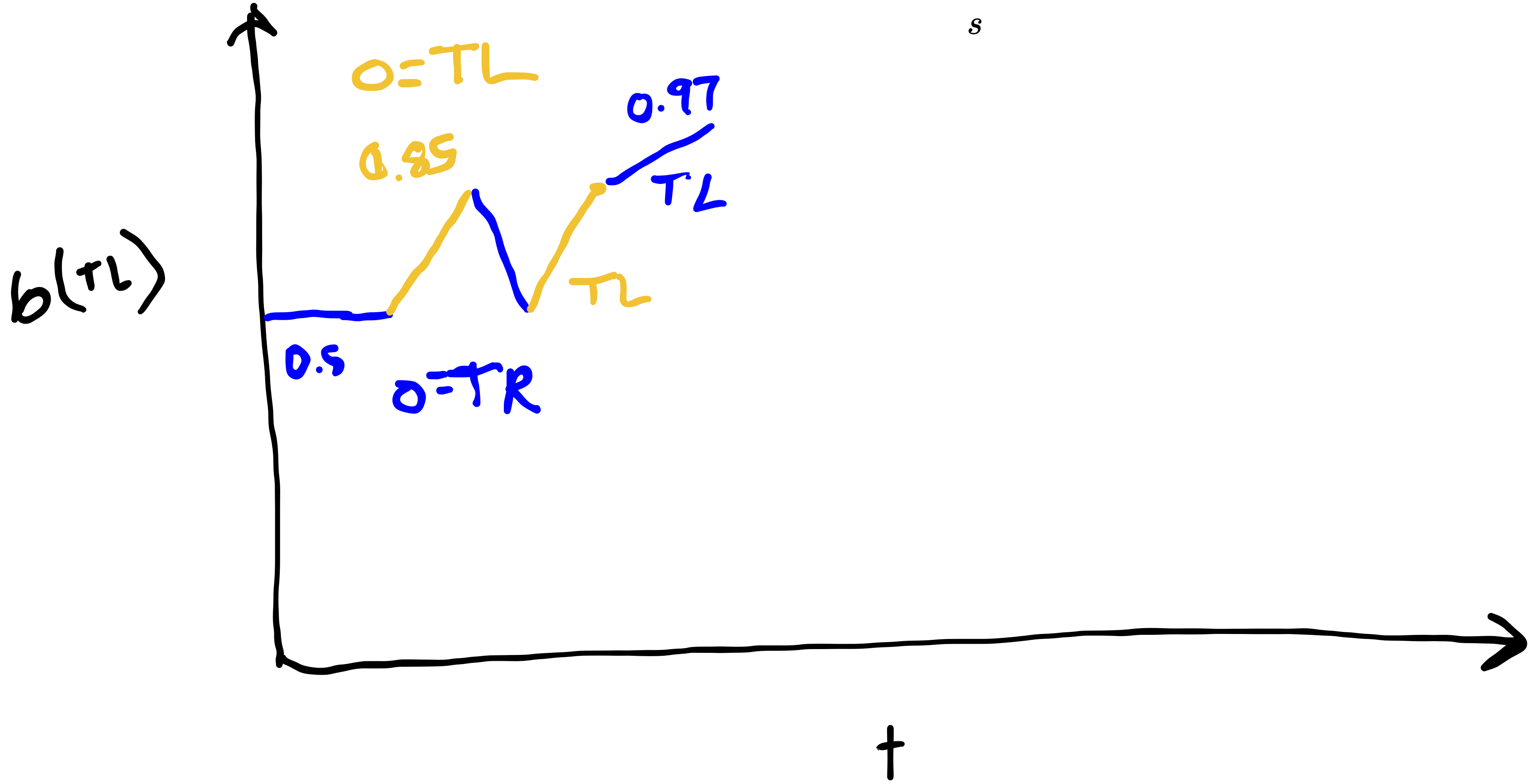
Belief Dynamics

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



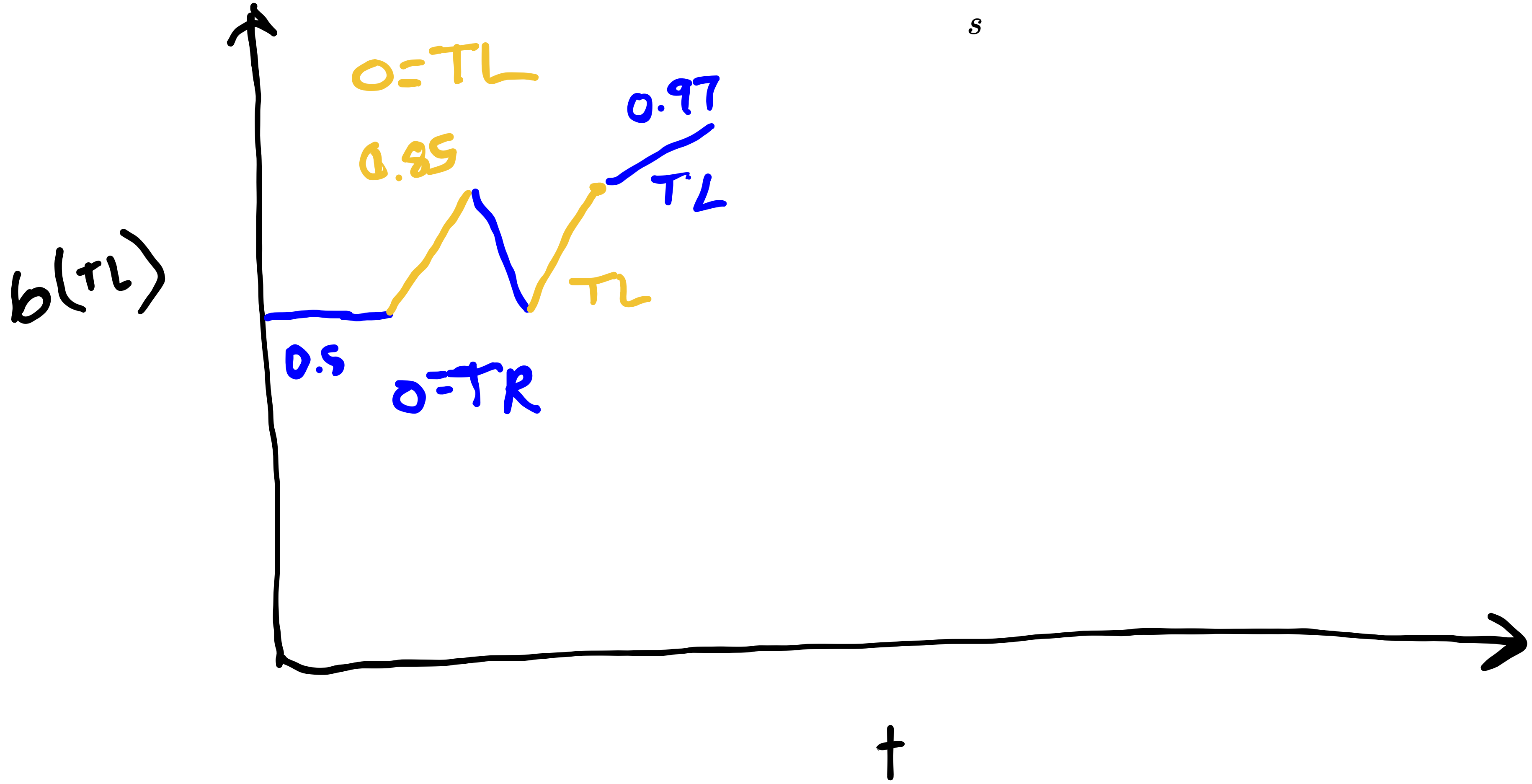
Belief Dynamics

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



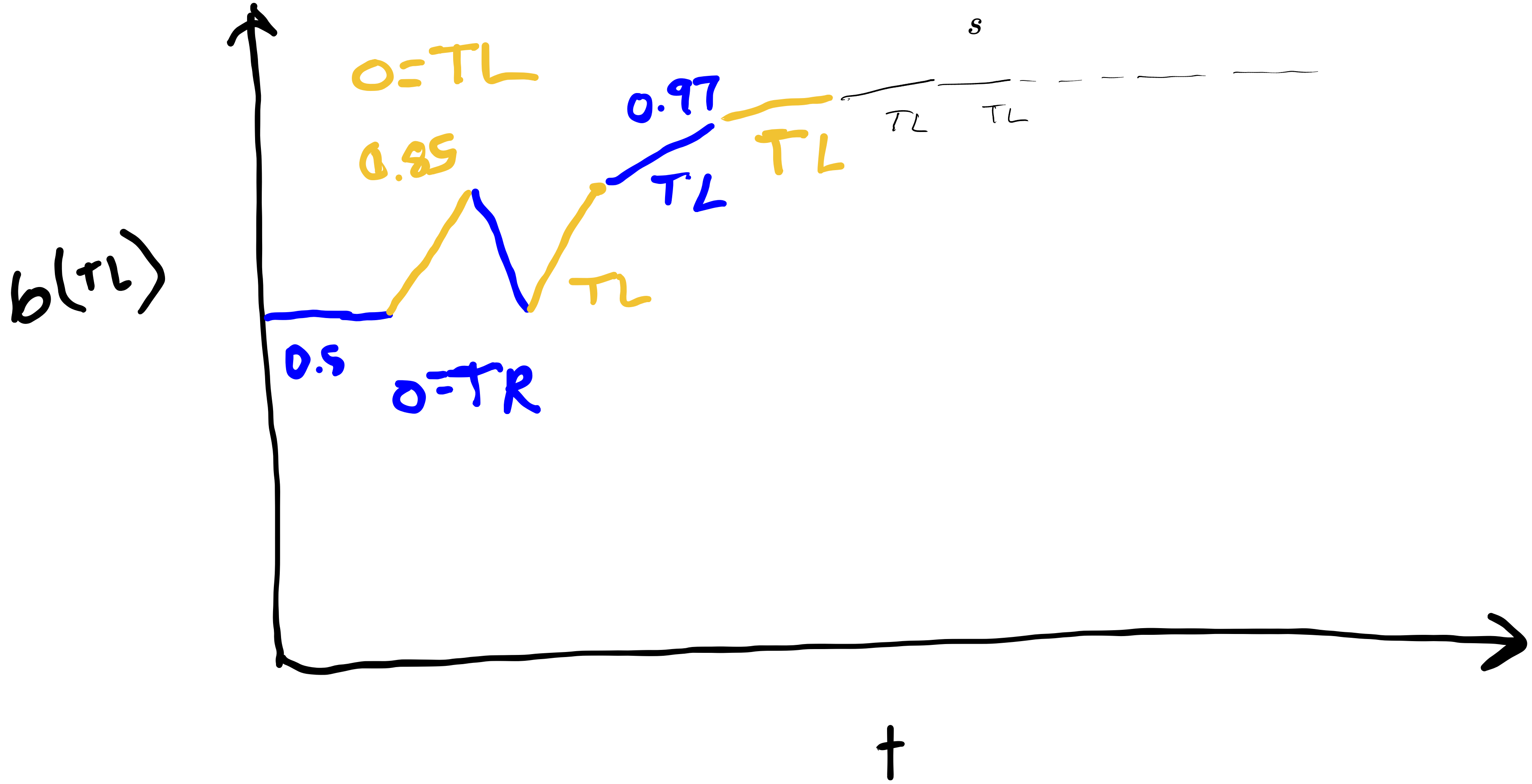
Belief Dynamics

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

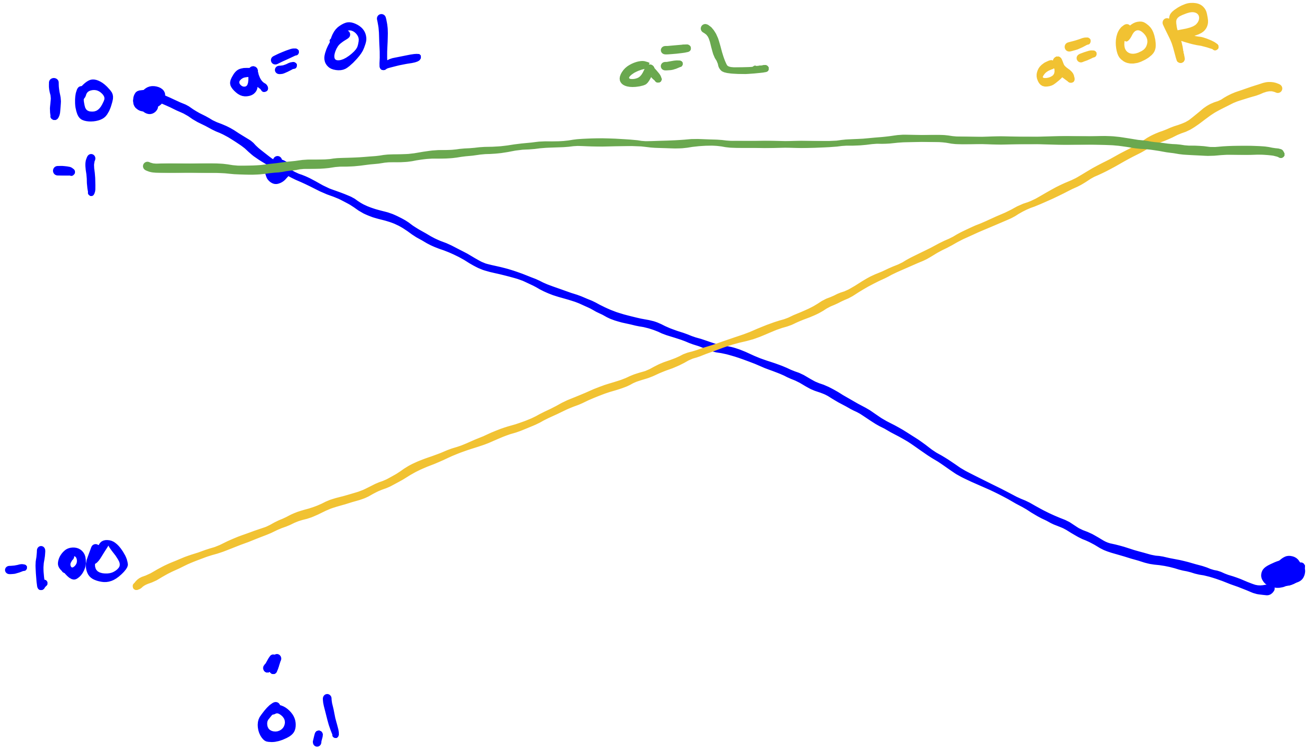


Belief Dynamics

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



One-step utility



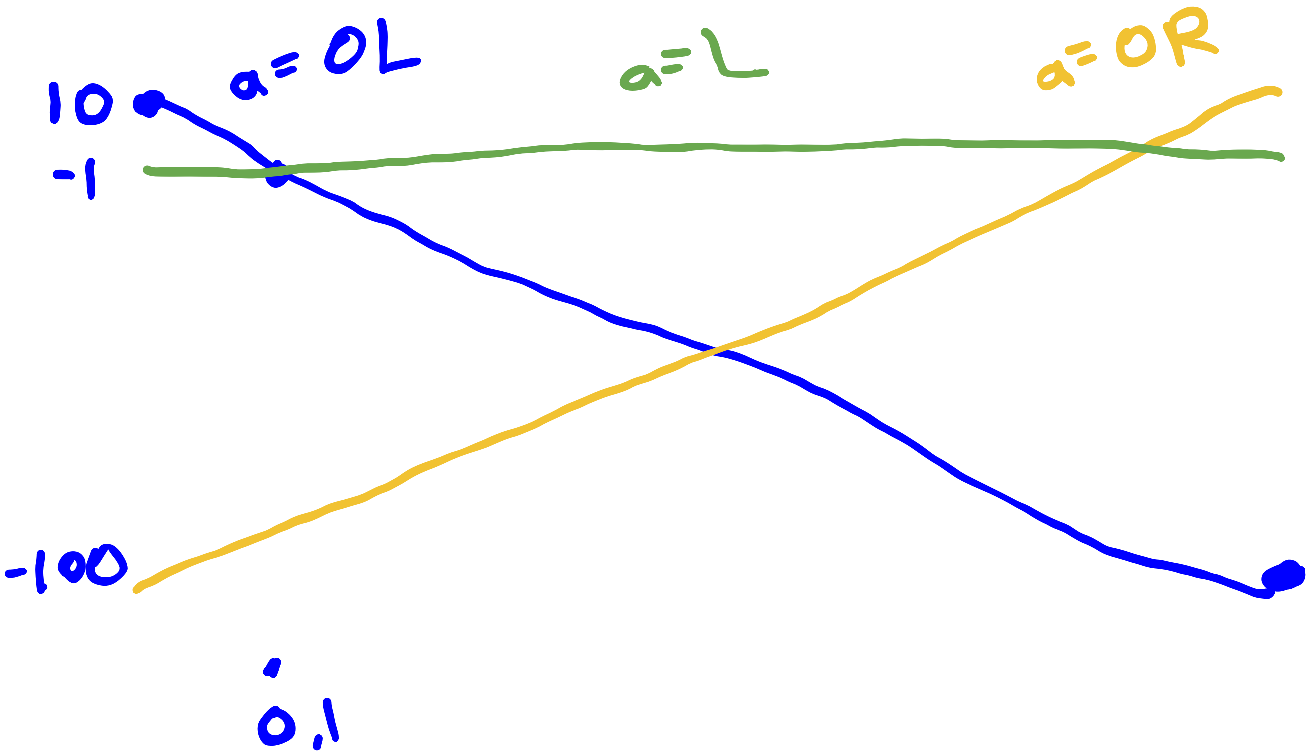
$R(b,a) = \vec{r}_a \cdot b$

α-vector

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

One-step utility

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



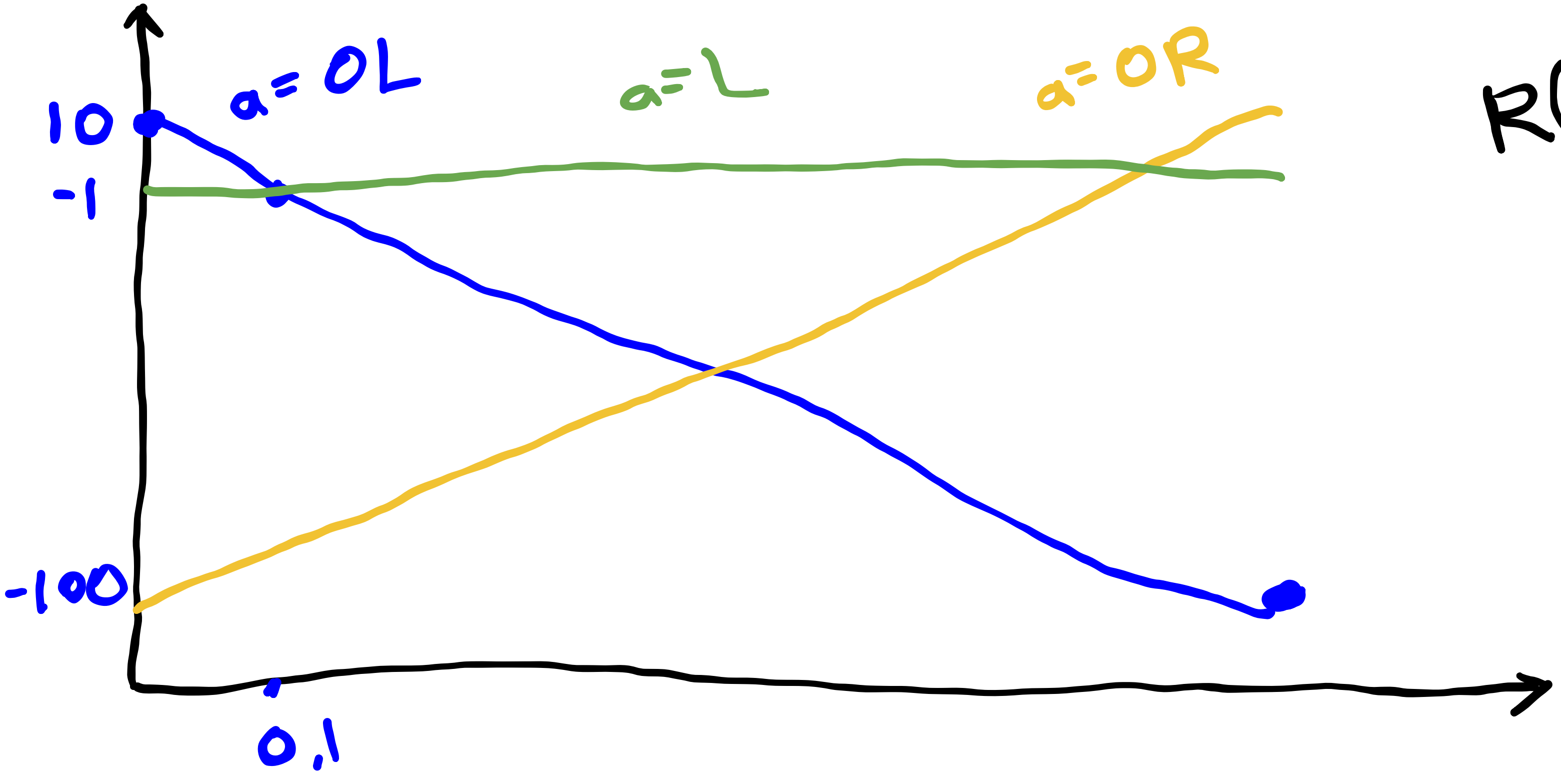
$$R(b,a) = \bar{r}_a \cdot b$$

α-vector

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

One-step utility

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



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

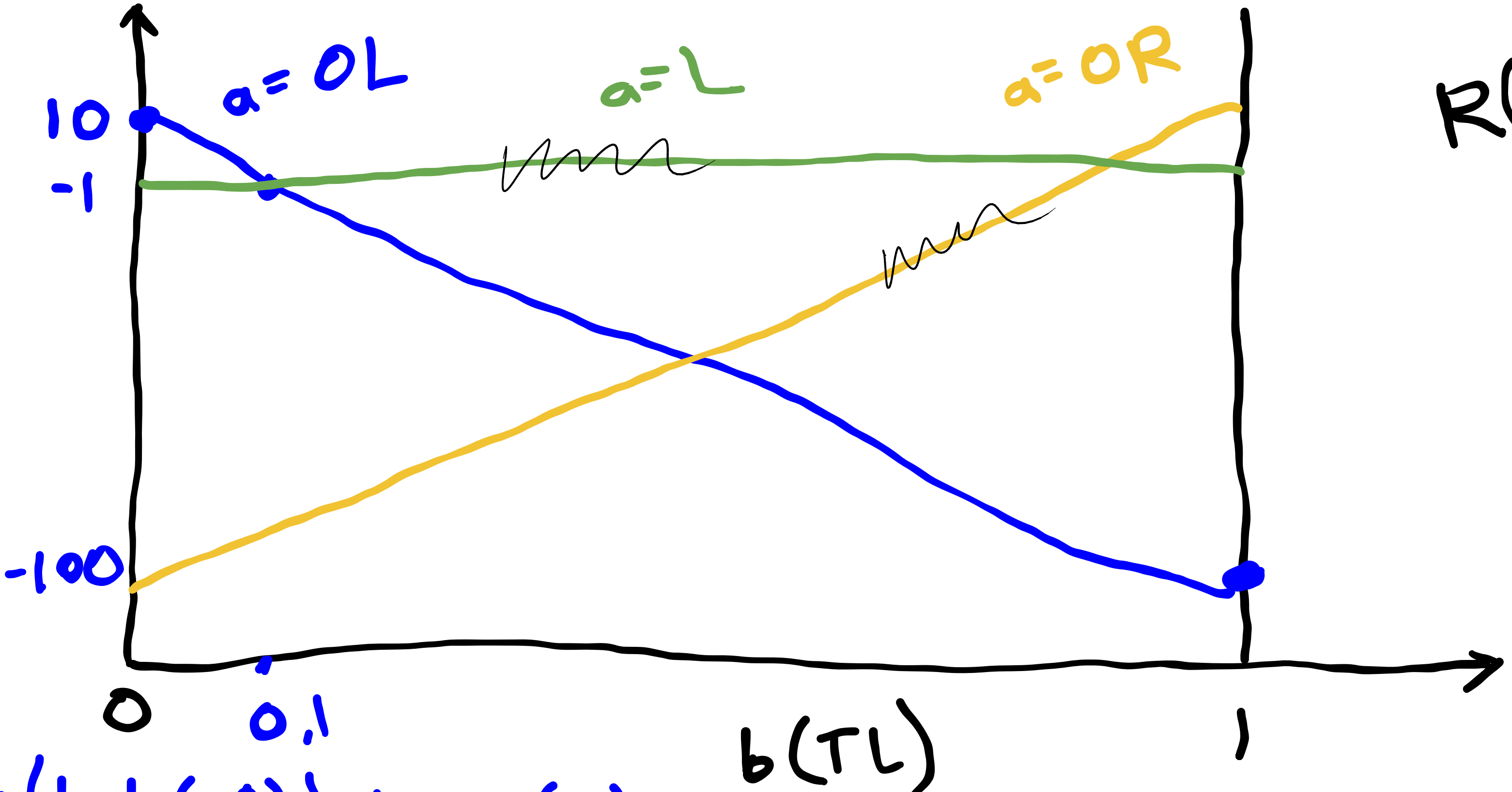
α-vector

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

One-step utility

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

1 step utility for $a = OL$ for any belief



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

α-vector

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

One-step utility

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

$a = OL$
$b(TL) = 0$

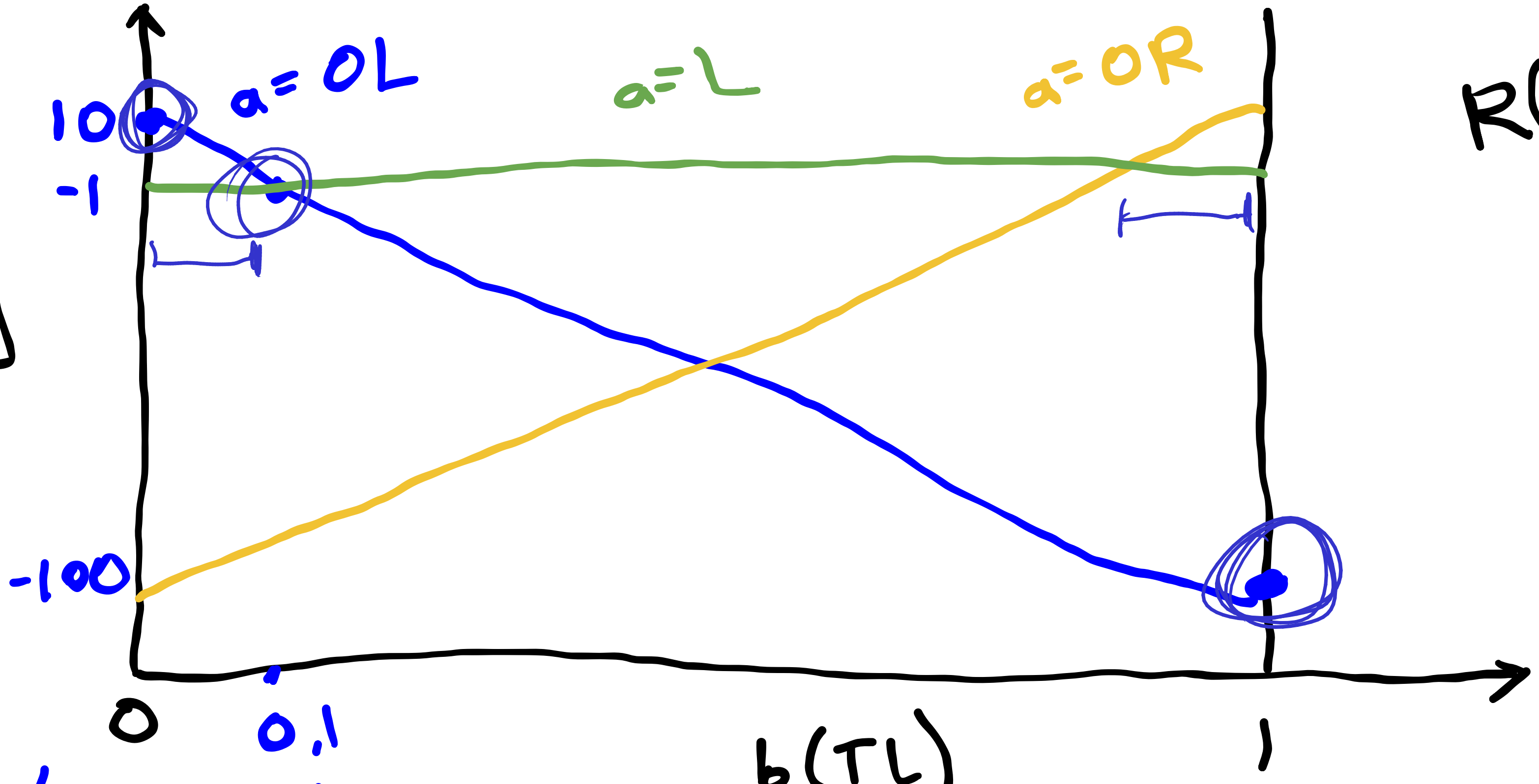
$$R(b,a) = (\underbrace{0.0}_{b(TL)} \times \underbrace{-100}_{R(TL,OL)}) + (\underbrace{1.0}_{1-b(TL)} \cdot \underbrace{+10}_{R(TR,OL)})$$

$$= 10$$

$a = OL$
$b(TL) = 1.0$
$R(b,a) = -100$

$b(TL) = 0.1$

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



$$R(b,a) = r_a \cdot b$$

α-vector

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

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: L OL OR

Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

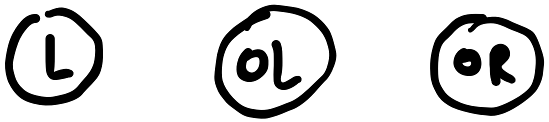
1 Step: L OL OR

2 Step:

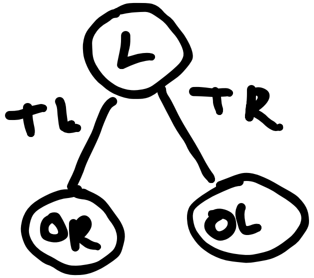
Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

1 Step:



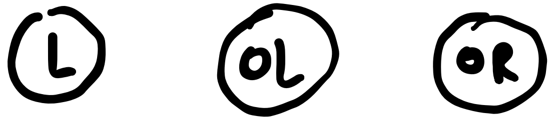
2 Step:



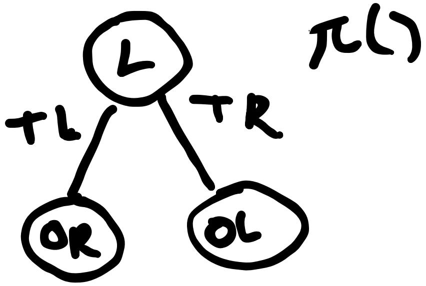
Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

1 Step:



2 Step:

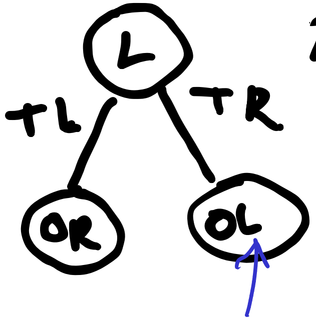


Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

1 Step: L OL OR

2 Step: $\pi(\ )$

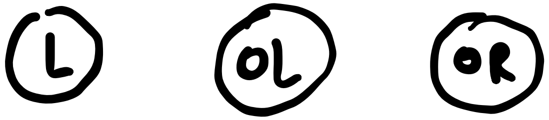


$\pi(TR) = OL$

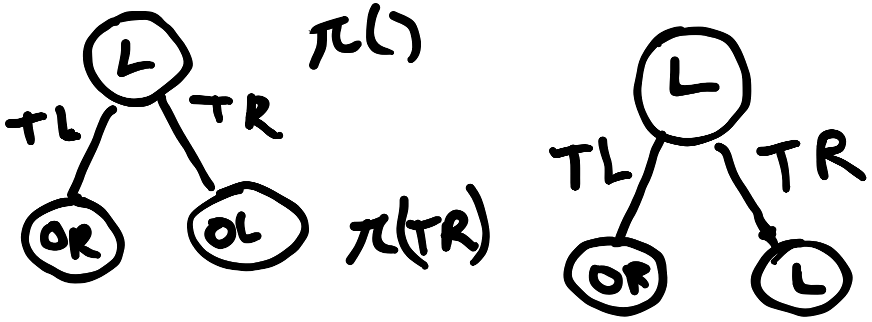
Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

1 Step:



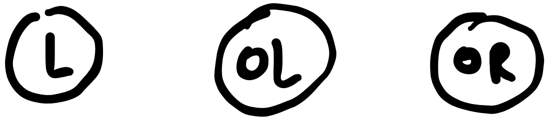
2 Step:



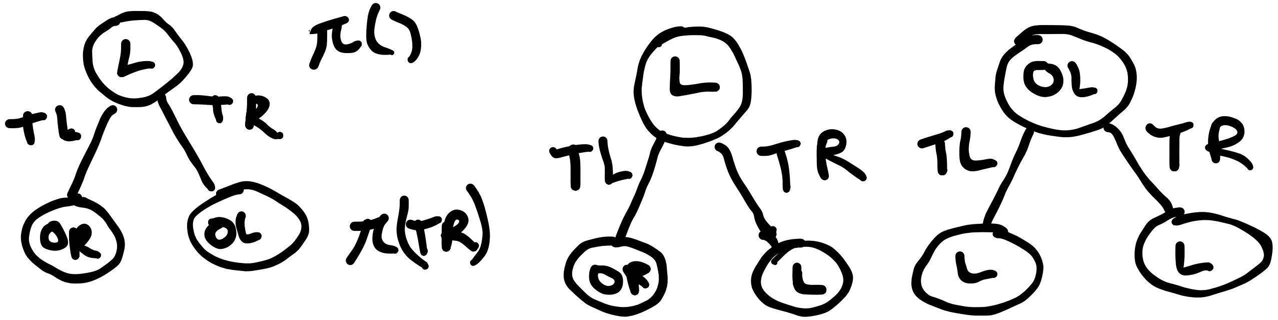
Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

1 Step:



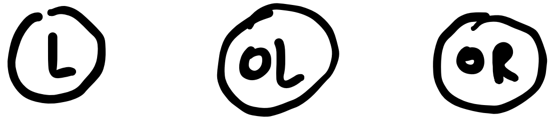
2 Step:



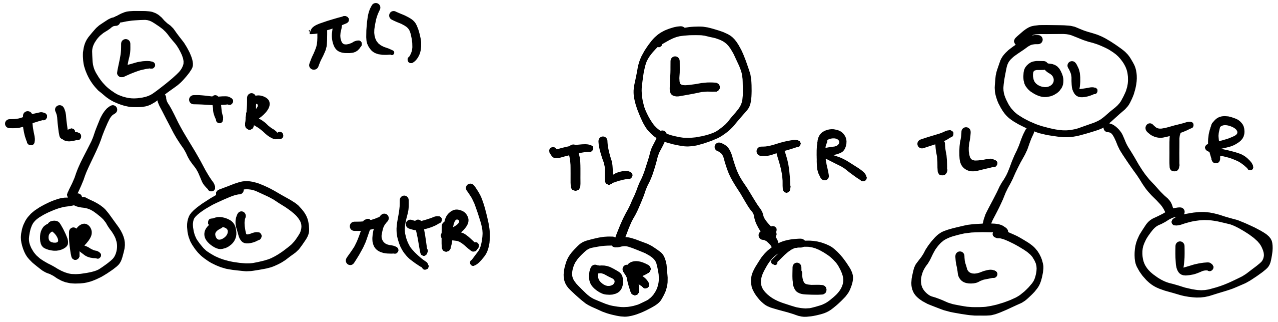
Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

1 Step:



2 Step:



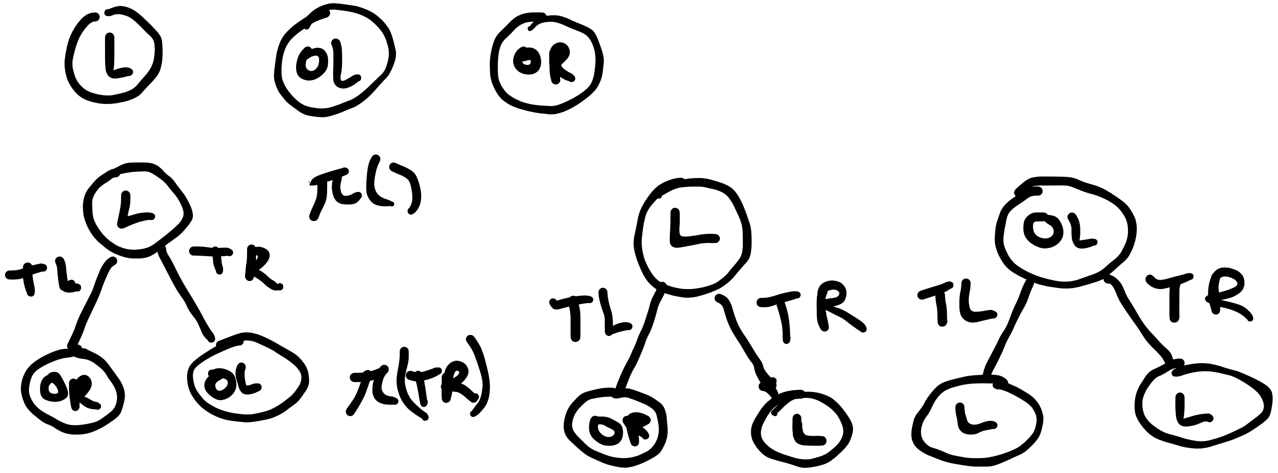
$$|A|^{\frac{(|O|^h-1)}{(|O|-1)}}$$

Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

1 Step:

2 Step:



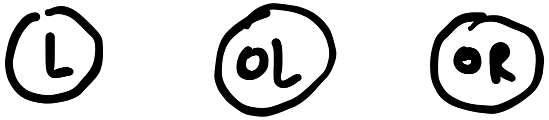
$|A|^{\frac{(|O|^h-1)}{(|O|-1)}}$

27 two step plans!

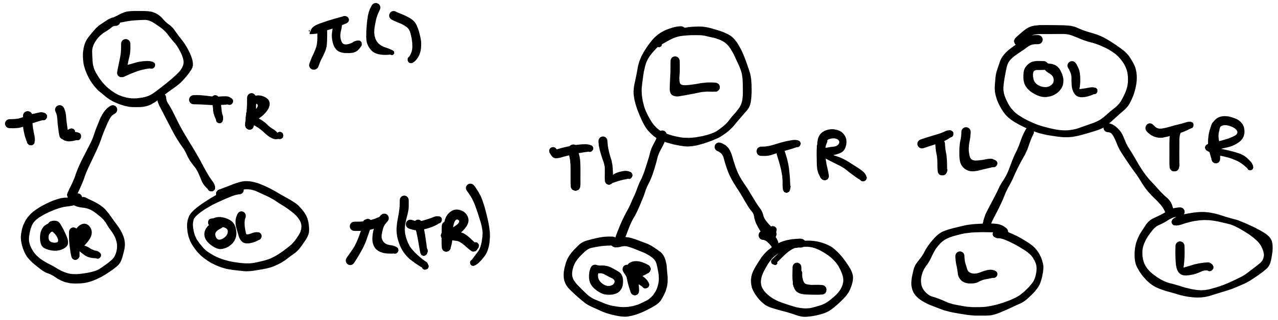
Alpha Vectors for Conditional Plans

Conditional Plans: fixed-depth history-based policies

1 Step:

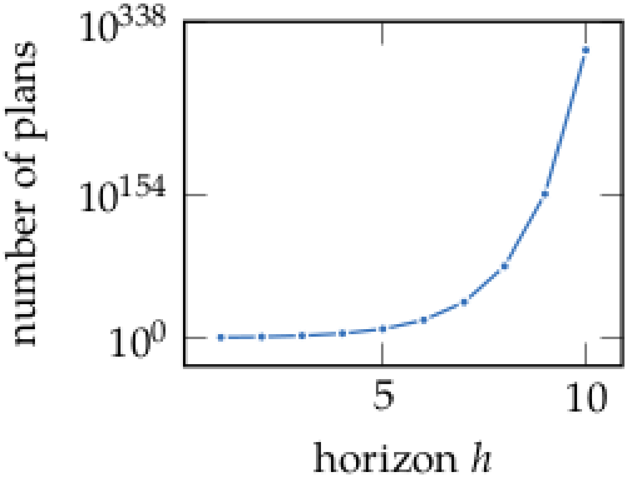


2 Step:



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

27 two step plans!

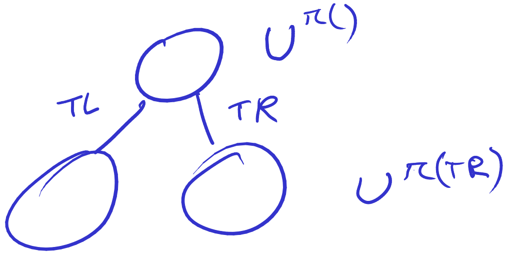


Alpha Vectors for Conditional Plans

Alpha Vectors for Conditional Plans

$U^{\pi}(s) = R(s, \pi(s)) + \gamma E(U^{\pi}(s'))$

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



Alpha Vectors for Conditional Plans

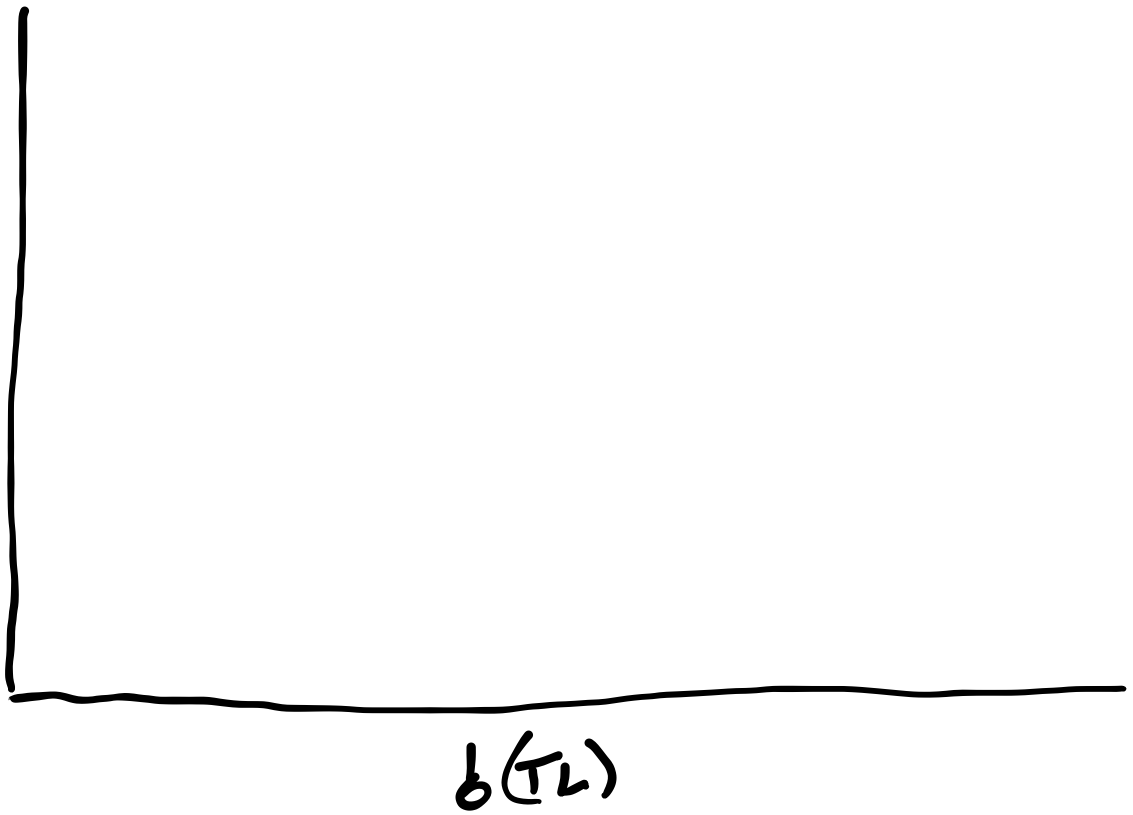
$$U^{\pi}(s) = R(s, \pi()) + \gamma \left[\sum_{s'} T\big(s' \mid s, \pi()\big) \sum_{o} O\big(o \mid \pi(), s'\big) U^{\pi(o)}(s') \right]$$

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

Alpha Vectors for Conditional Plans

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

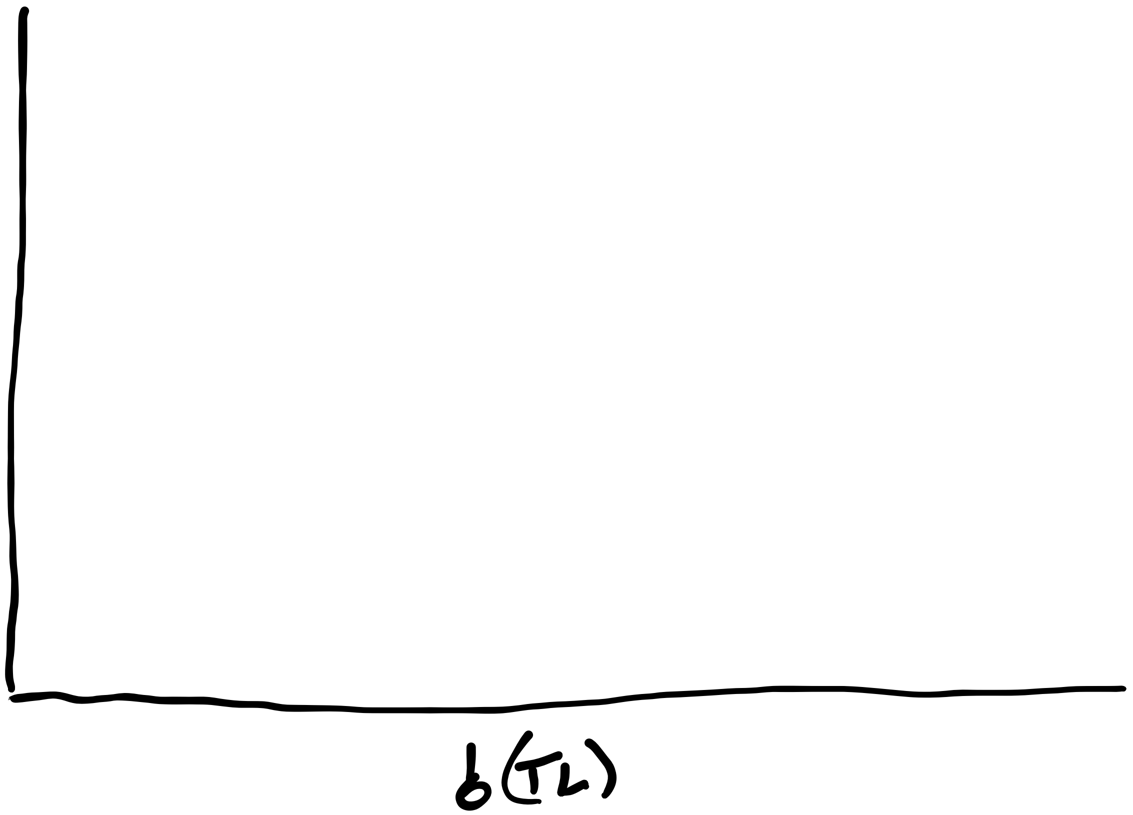
For 1-step: $U^{\pi}(s) = R(s, \pi())$

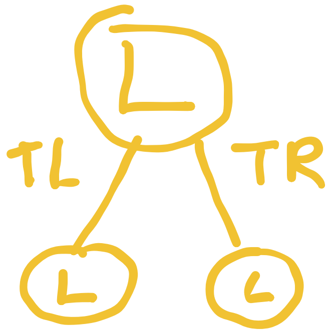


Alpha Vectors for Conditional Plans

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

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

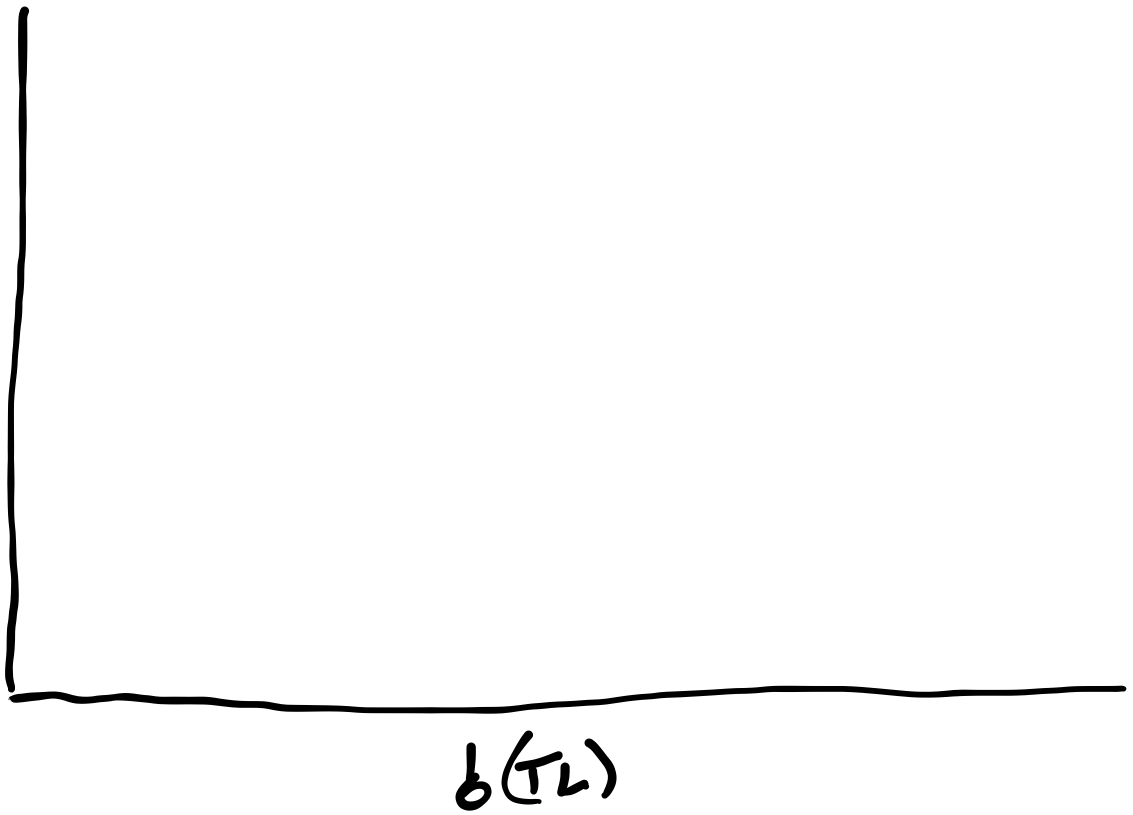


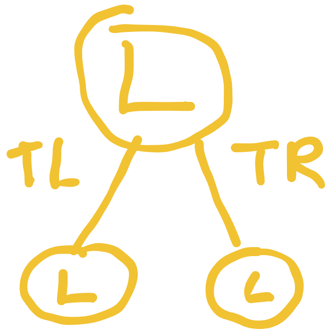


Alpha Vectors for Conditional Plans

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

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



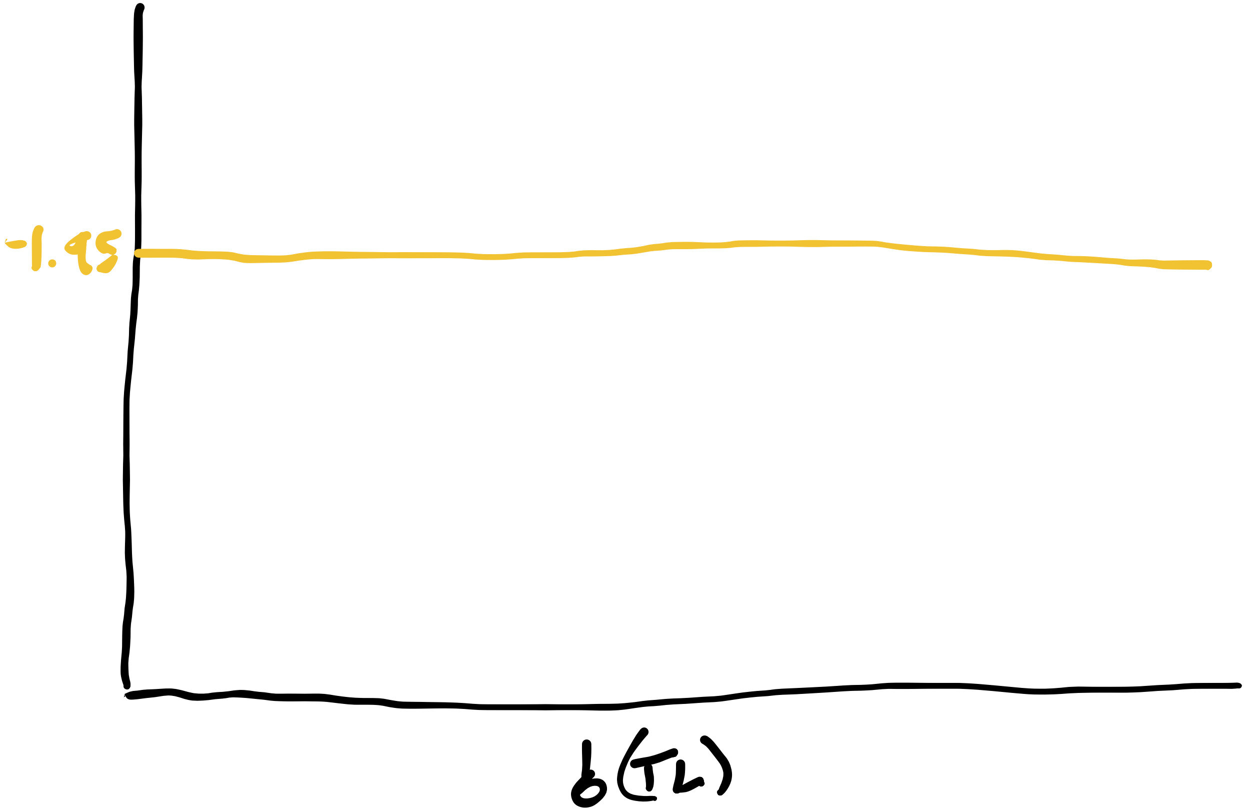


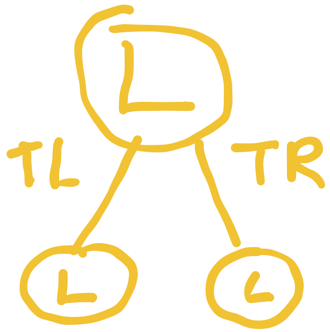
$U^{\pi}(s) = -1 + \gamma(-1)$

Alpha Vectors for Conditional Plans

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

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



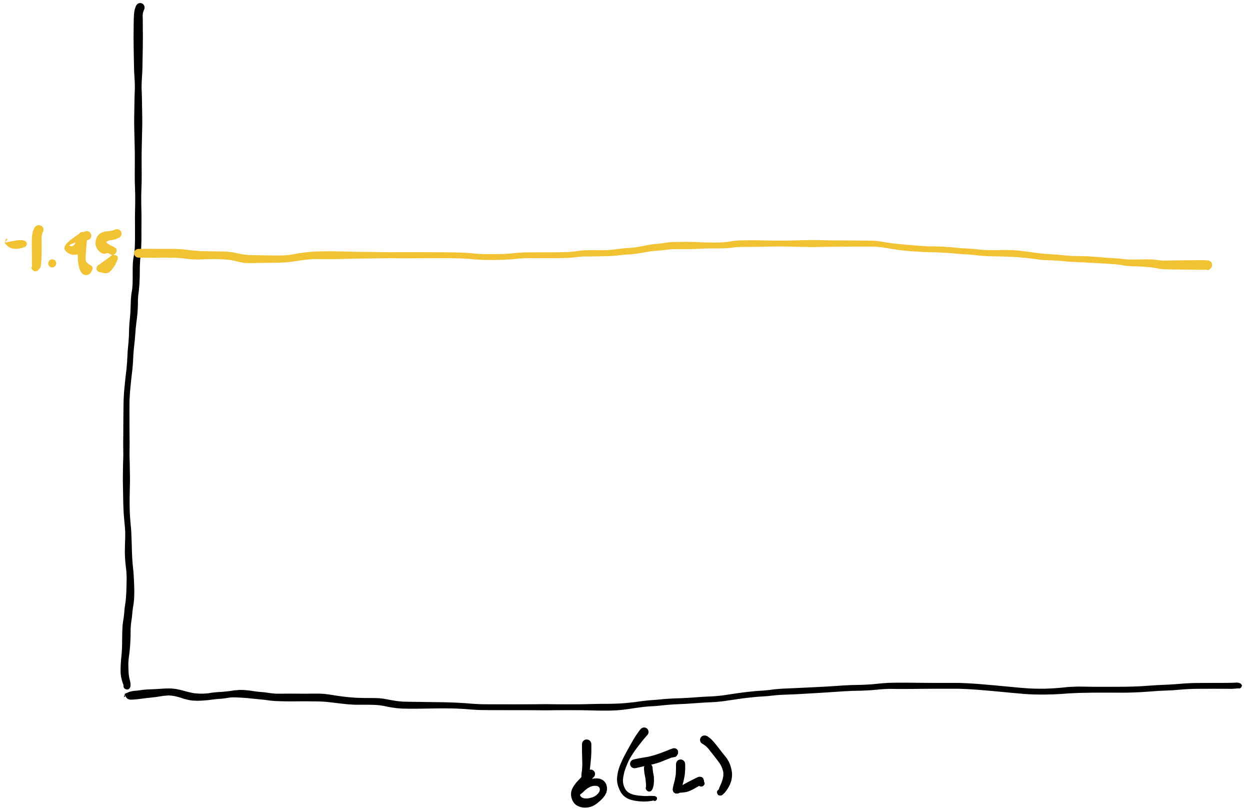


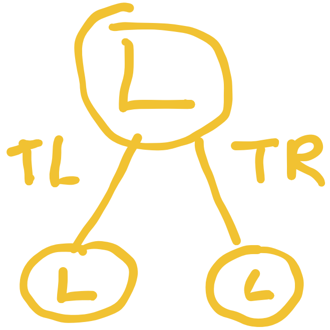
$U^{\pi}(s) = -1 + \gamma(-1)$

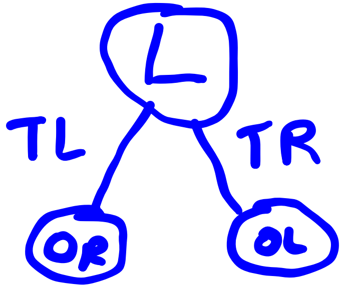
Alpha Vectors for Conditional Plans

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

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





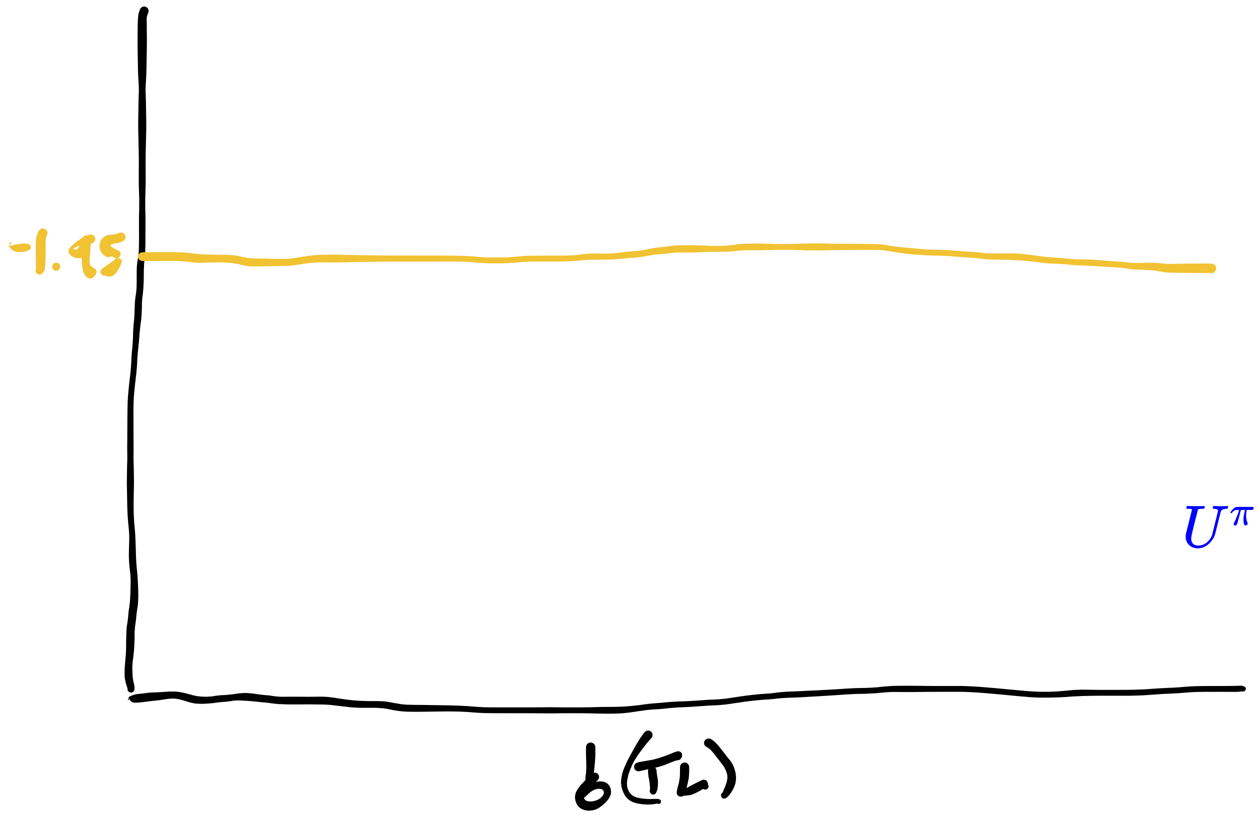


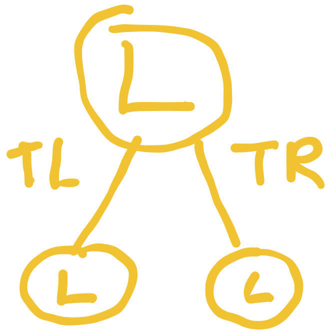
$U^{\pi}(s) = -1 + \gamma(-1)$

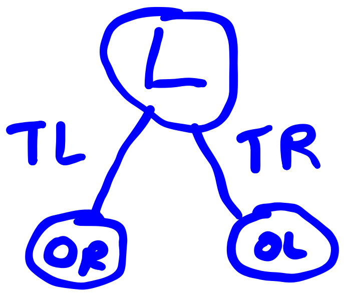
Alpha Vectors for Conditional Plans

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

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







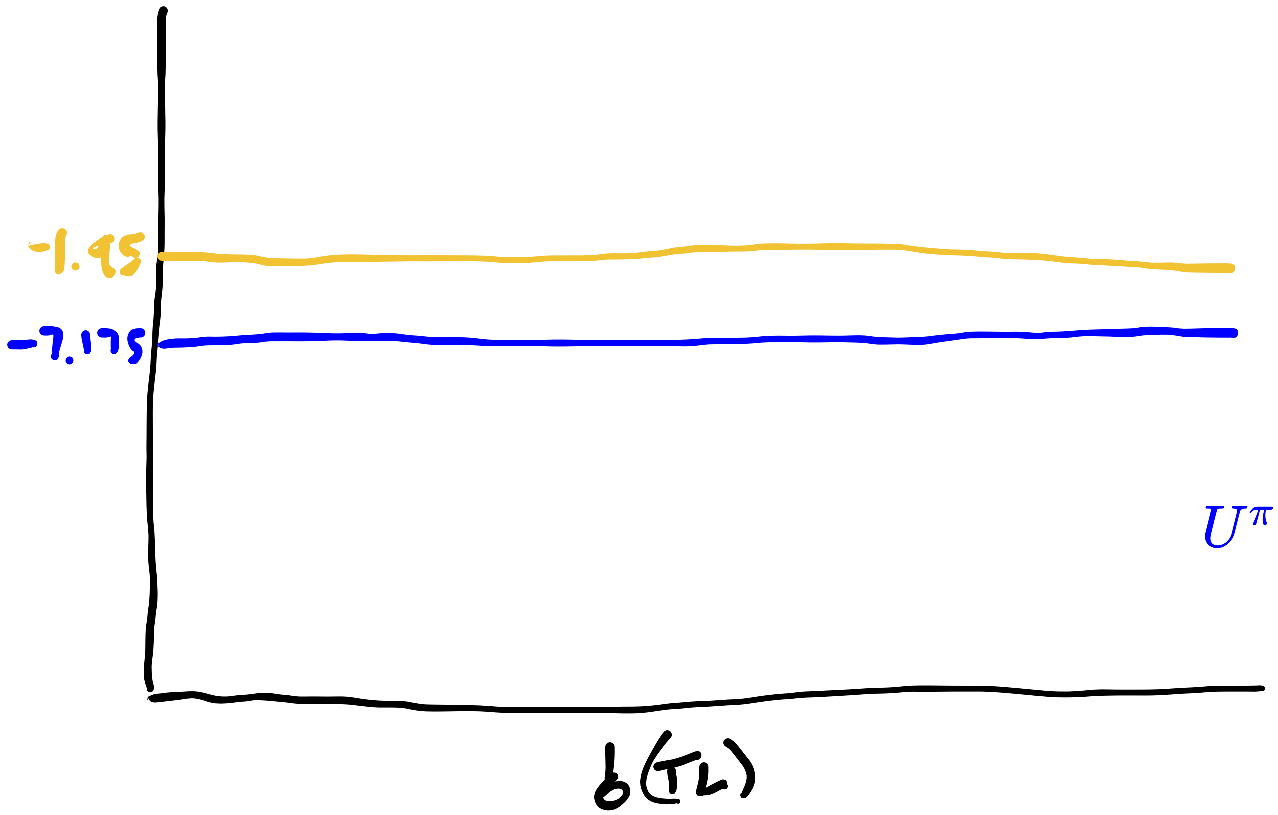
$$U^{\pi}(s) = -1 + \gamma(-1)$$

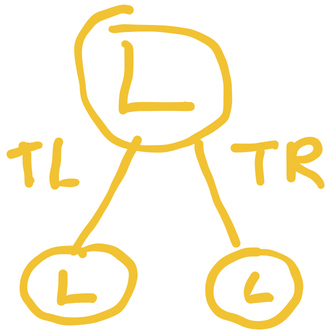
$$U^{\pi}(TL) = -1 + 0.95(1(0.85 \times 10 + 0.15 \times -100))$$

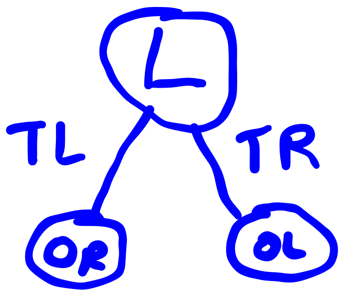
Alpha Vectors for Conditional Plans

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

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







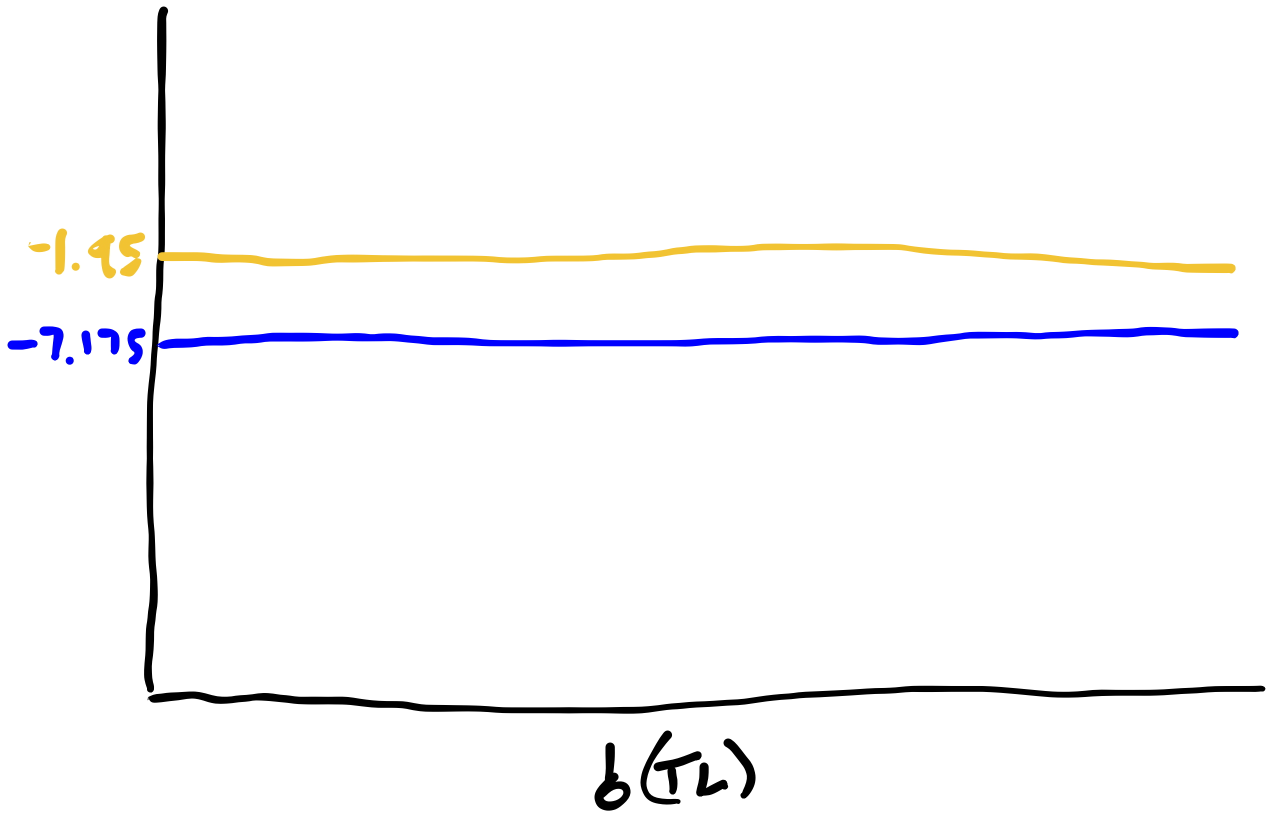
$$U^{\pi}(s) = -1 + \gamma(-1)$$

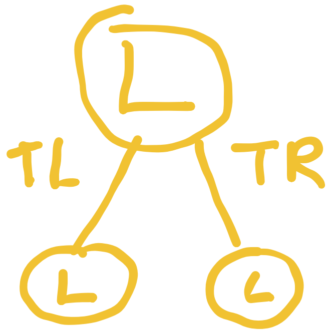
$$\begin{aligned} U^{\pi}(TL) &= -1 + 0.95(1(0.85 \times 10 + 0.15 \times -100)) \\ &= -7.175 \end{aligned}$$

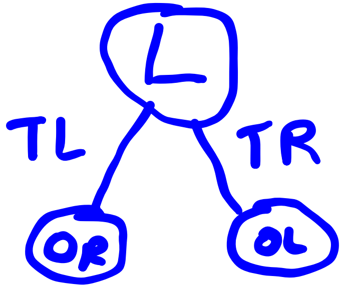
Alpha Vectors for Conditional Plans

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

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



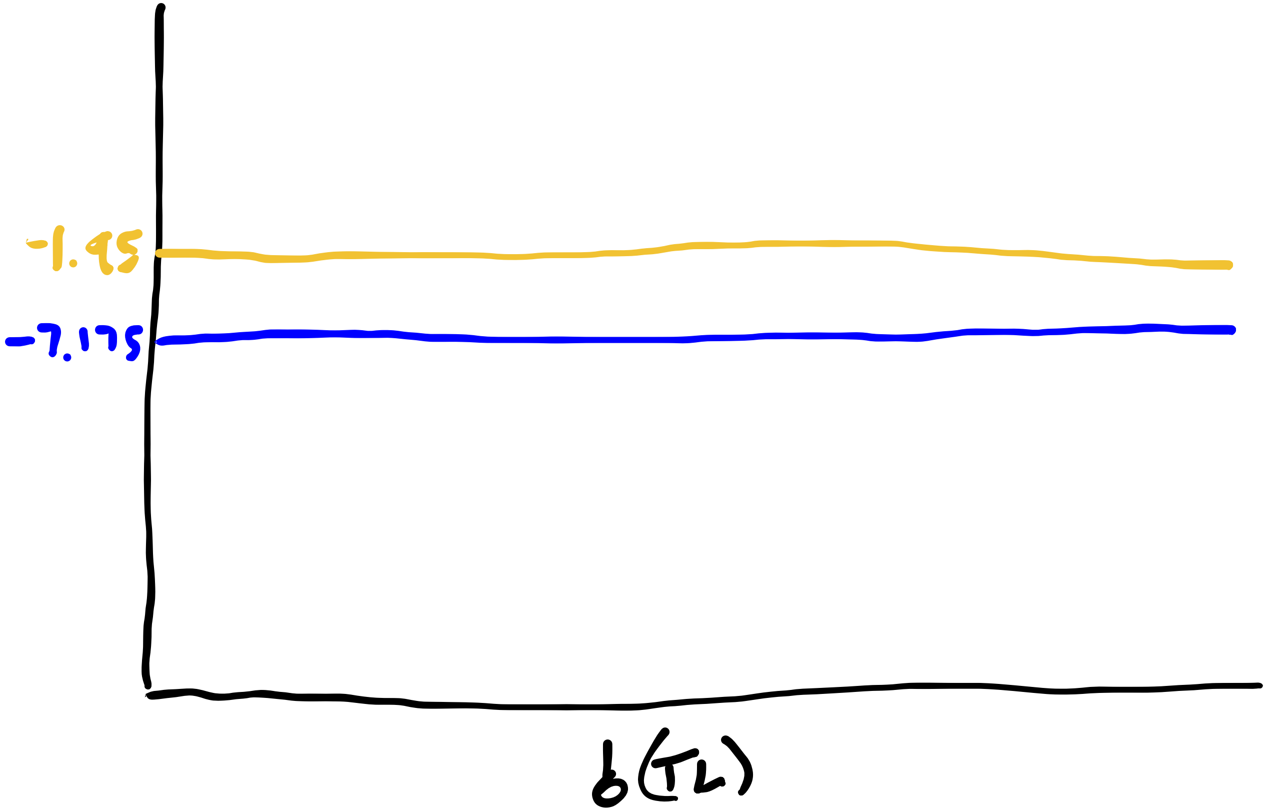


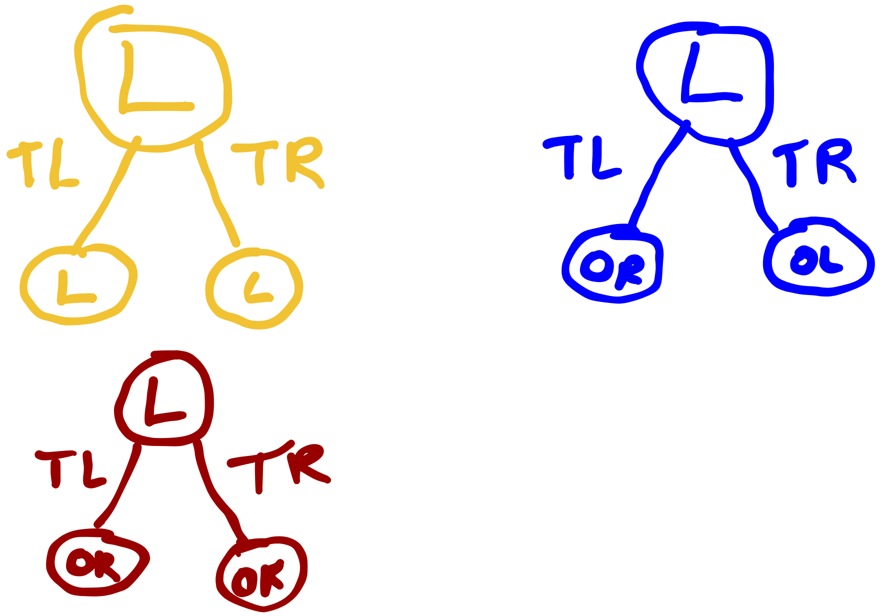


Alpha Vectors for Conditional Plans

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

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

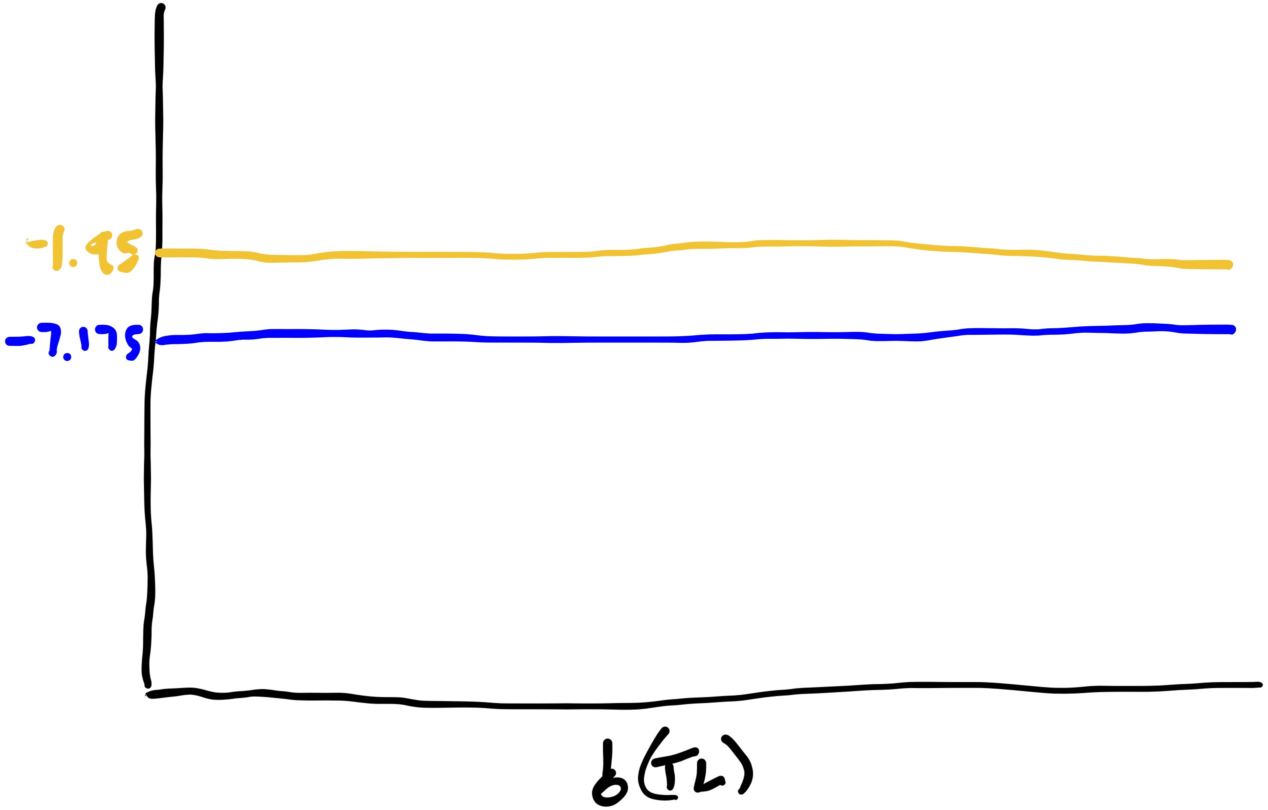


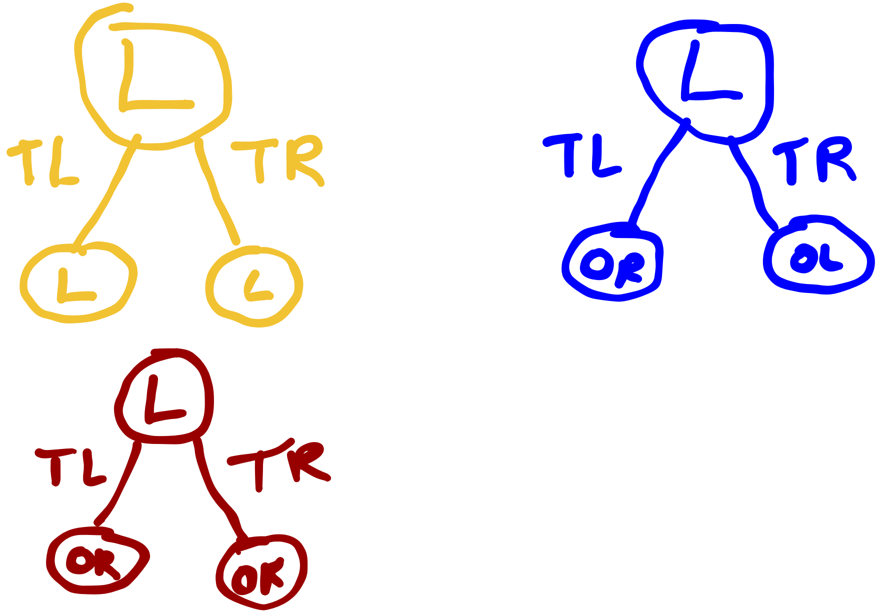


Alpha Vectors for Conditional Plans

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

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



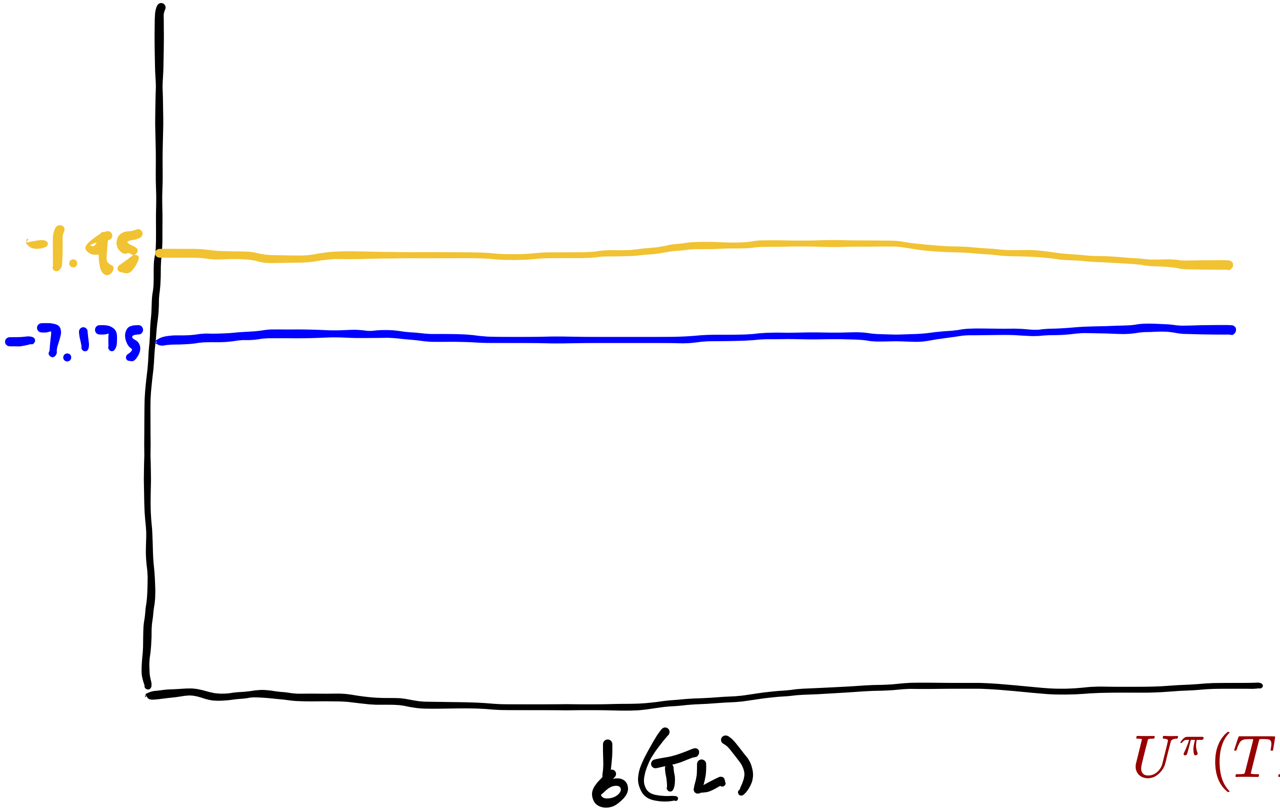


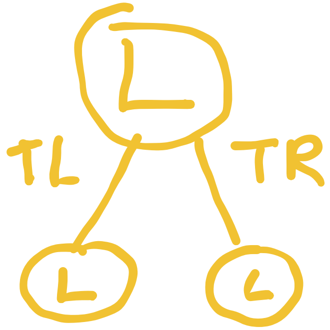
$U^{\pi}(TL) = -1 + \gamma\, 10$

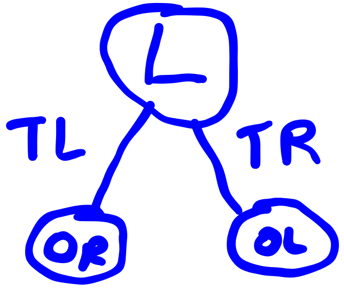
Alpha Vectors for Conditional Plans

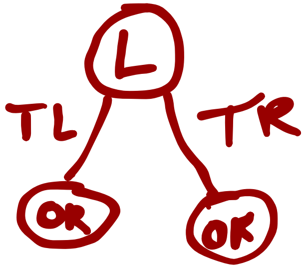
$$U^{\pi}(s) = R(s, \pi()) + \gamma \left[\sum_{s'} T(s' \mid s, \pi()) \sum_{o} O(o \mid \pi(), s') U^{\pi(o)}(s') \right]$$

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









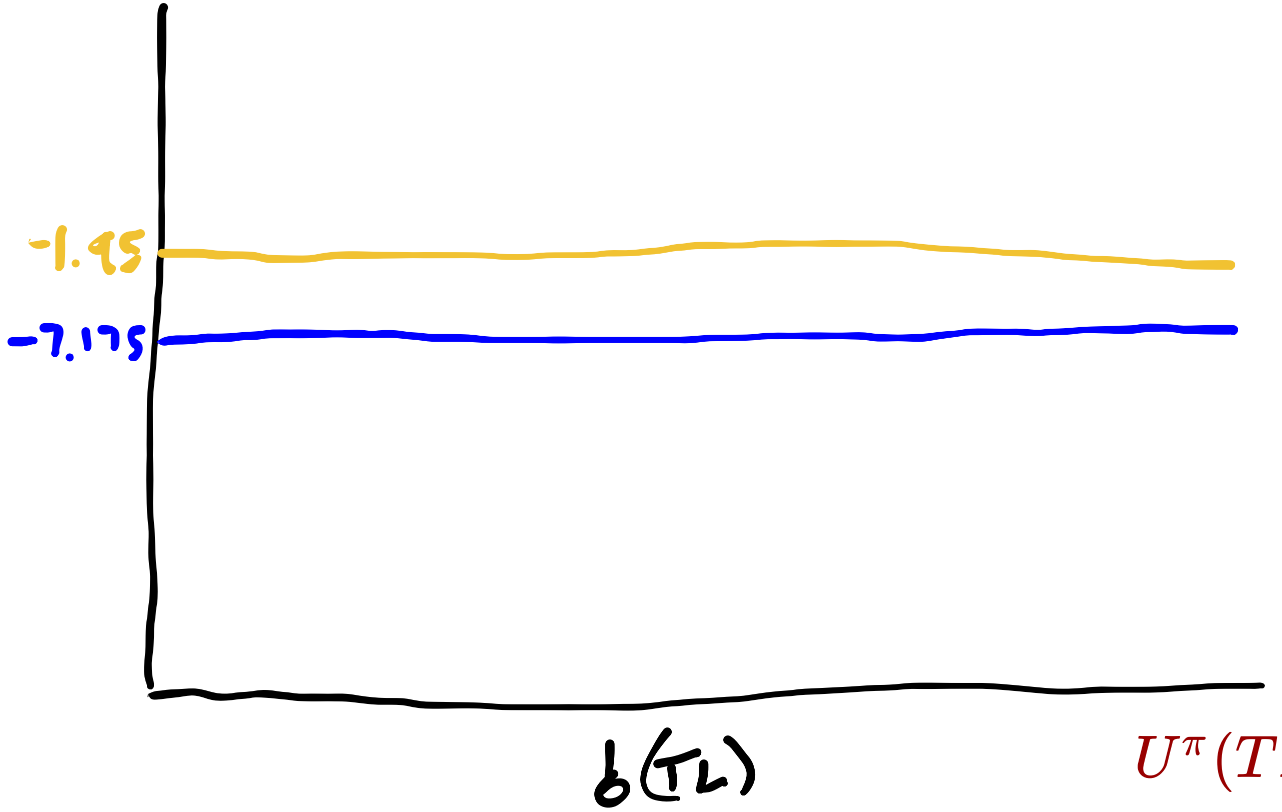
$$U^{\pi}(TL) = -1 + \gamma\, 10$$

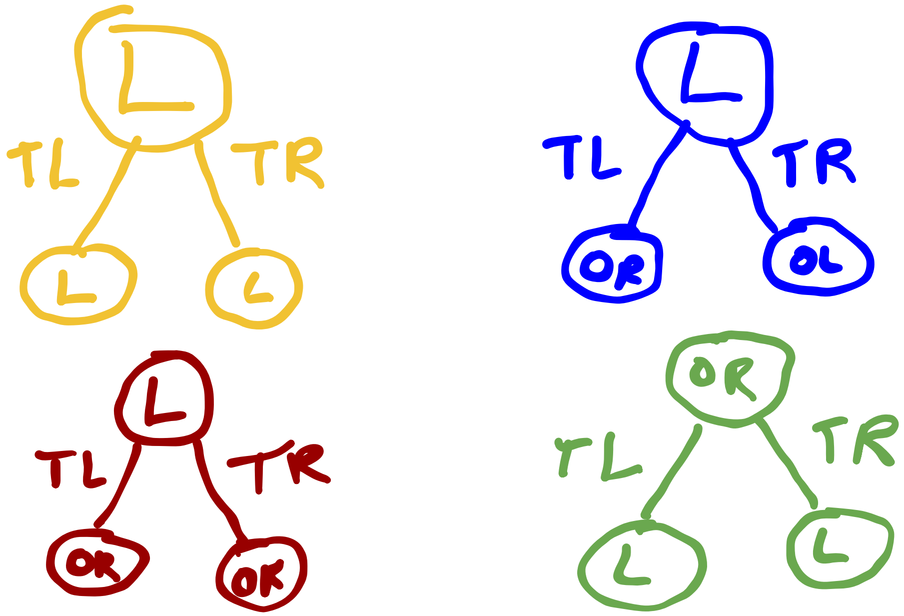
$$U^{\pi}(TR) = -1 + \gamma(-100)$$

Alpha Vectors for Conditional Plans

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

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





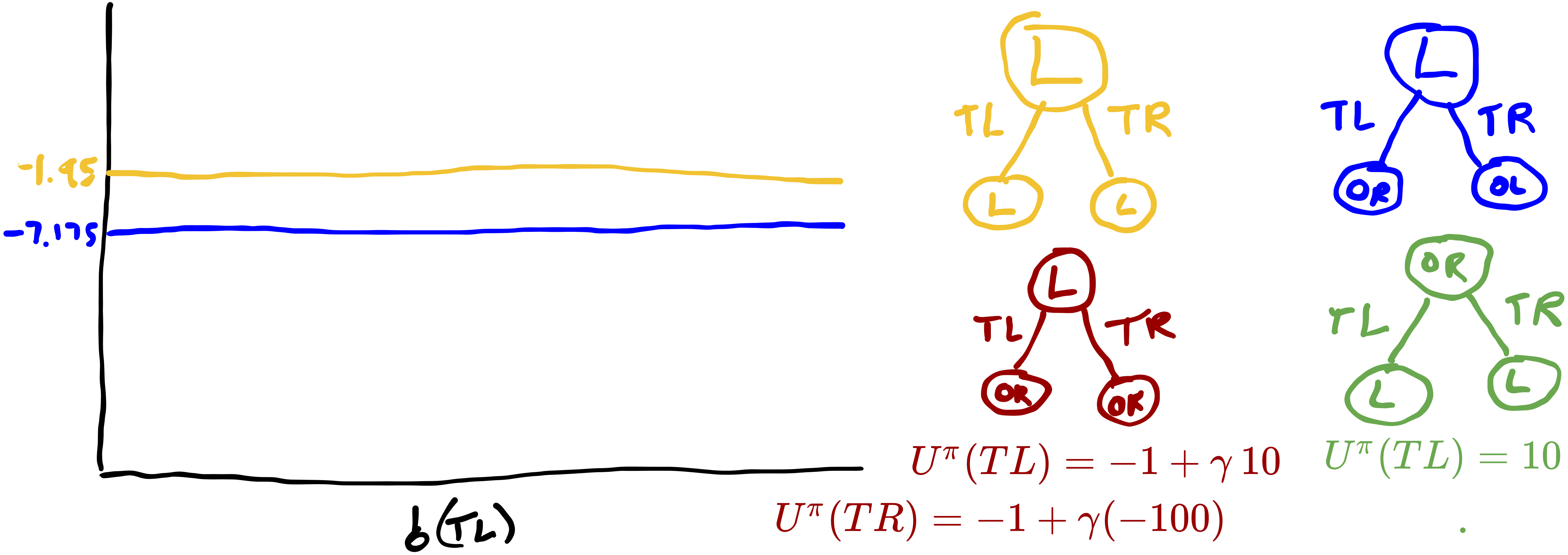
$U^{\pi}(TL) = -1 + \gamma 10$

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

Alpha Vectors for Conditional Plans

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

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



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

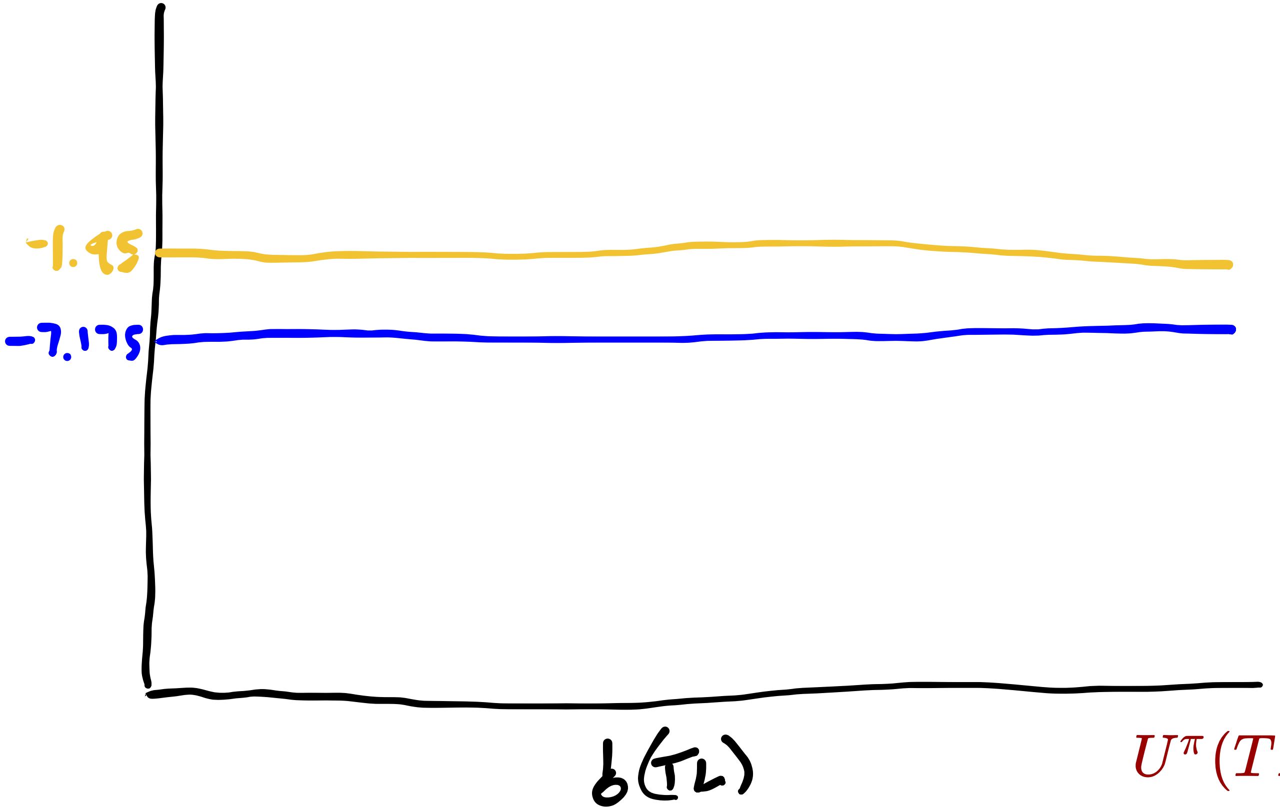
$U^{\pi}(TR) = -1 + \gamma(-100)$

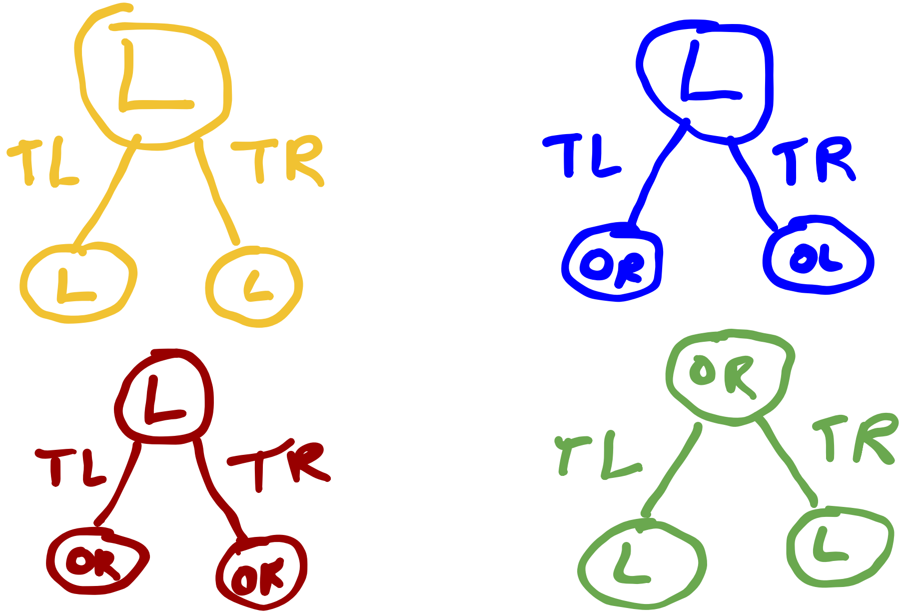
$U^{\pi}(TL) = 10 + \gamma(-1)$

Alpha Vectors for Conditional Plans

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

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





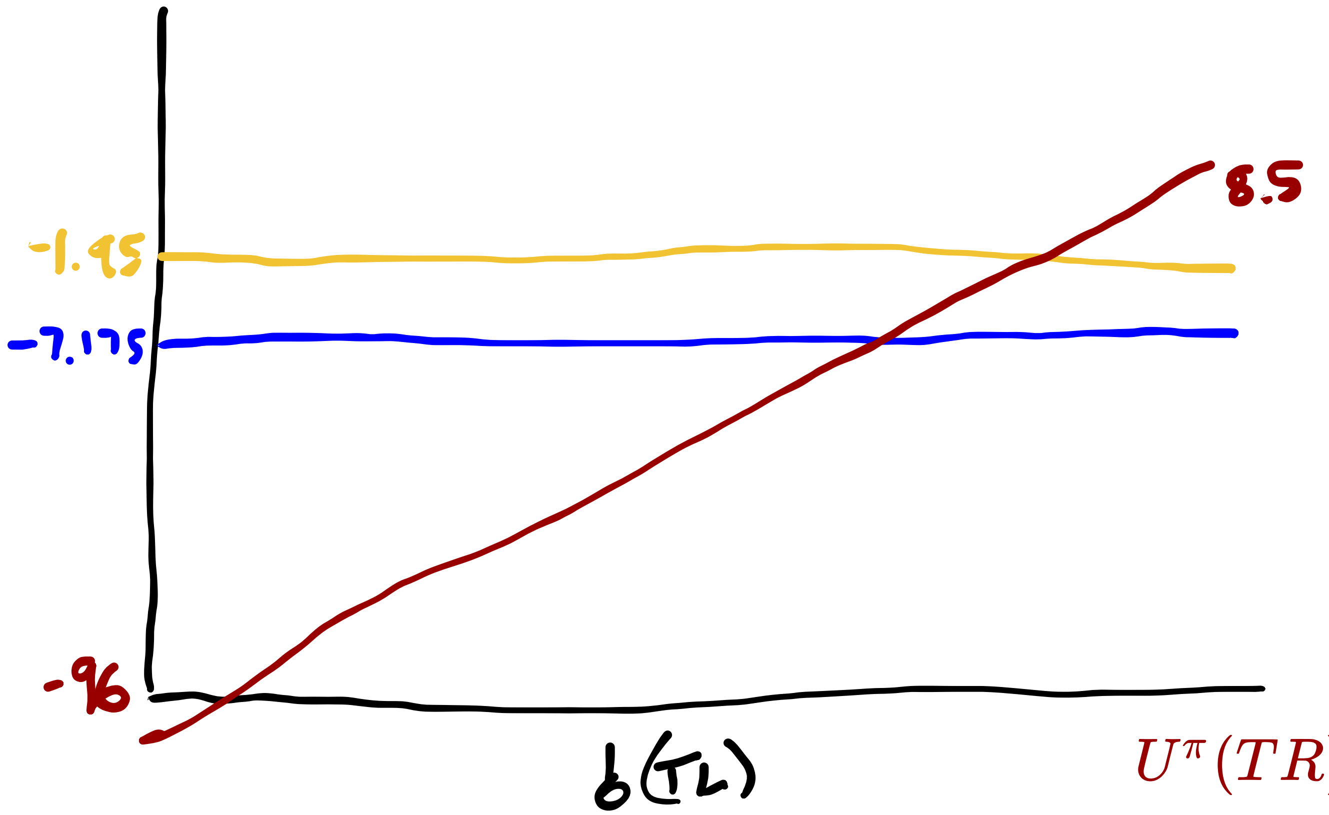
$U^{\pi}(TL) = -1 + \gamma 10$ $\quad U^{\pi}(TL) = 10 + \gamma(-1)$

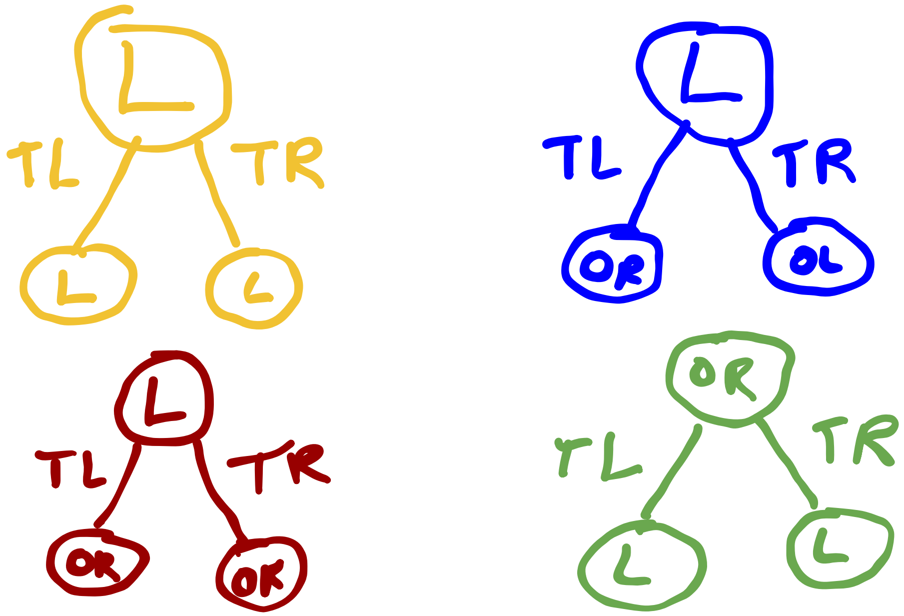
$U^{\pi}(TR) = -1 + \gamma(-100)$ $\quad U^{\pi}(TR) =. -100 + \gamma(-1)$

Alpha Vectors for Conditional Plans

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

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





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

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

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

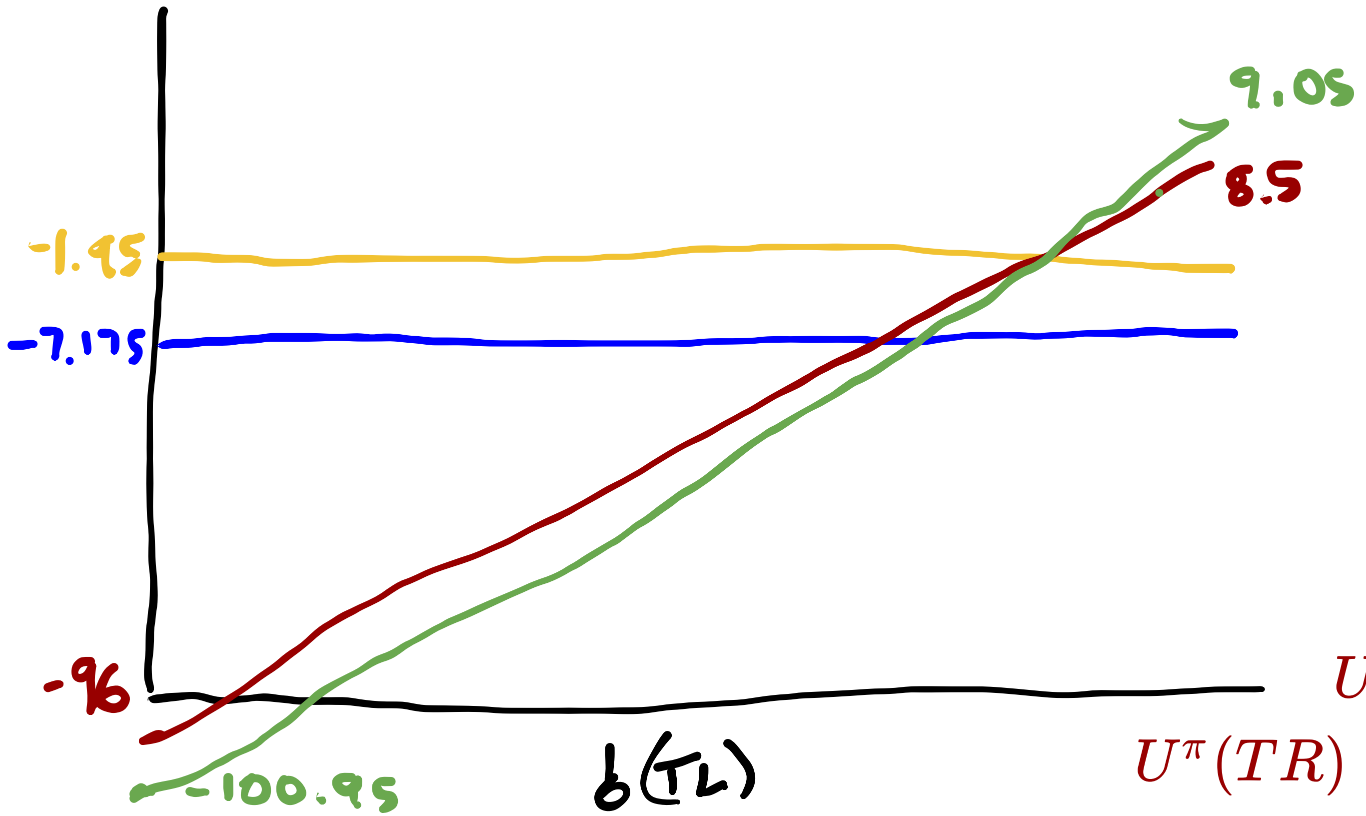
$U^{\pi}(TR) = .-100 + \gamma(-1)$

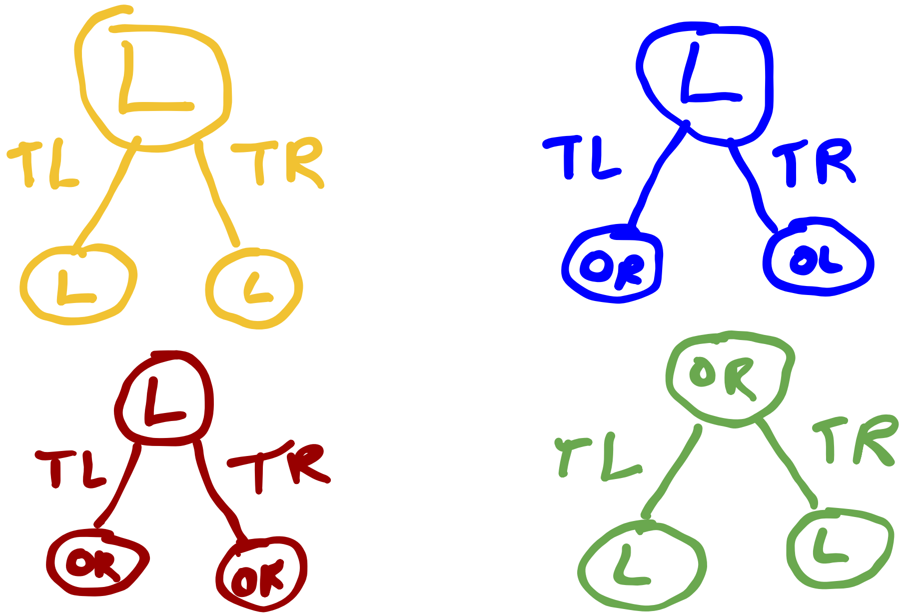
Alpha Vectors for Conditional Plans

$U^{\pi}(b) = \underset{s \sim b}{E}\left[U^{\pi}(s)\right]$

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

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





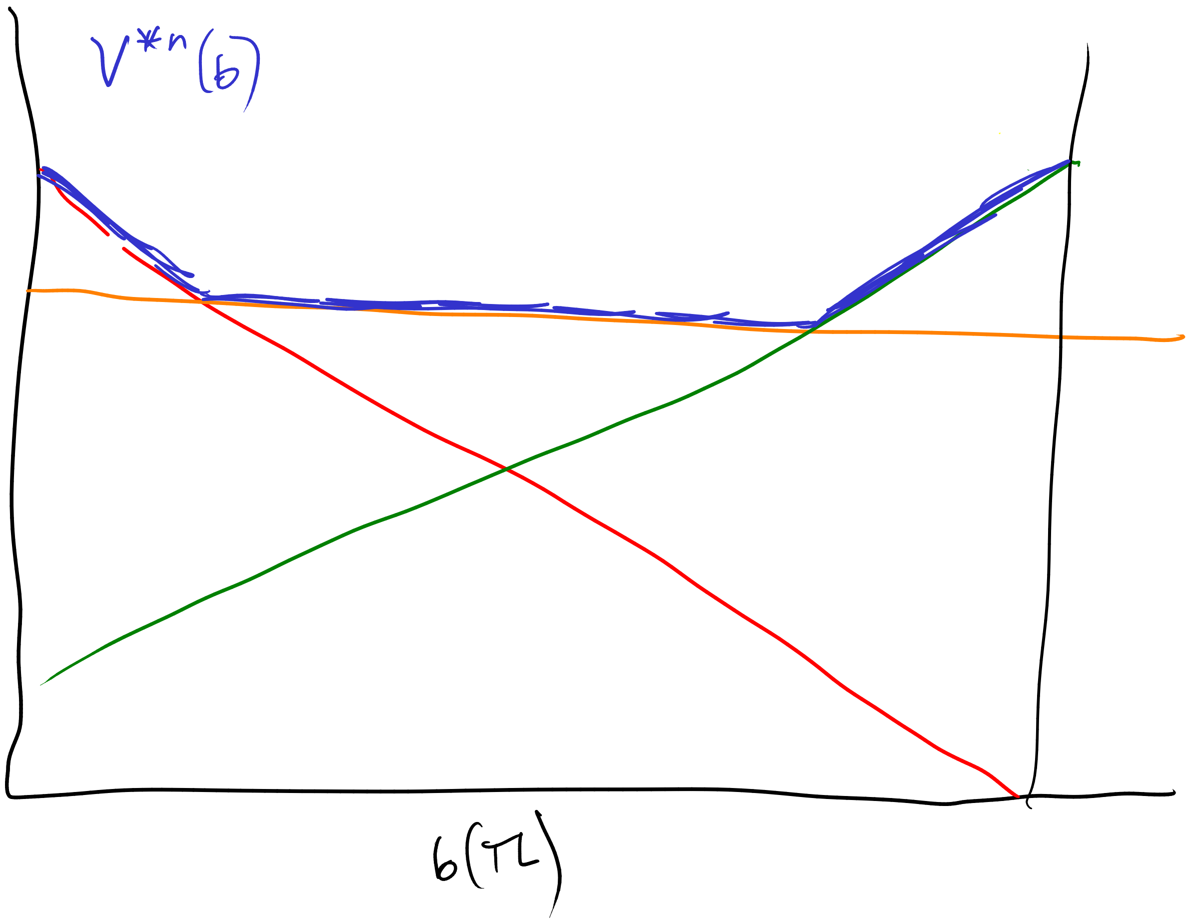
$U^{\pi}(TL) = -1 + \gamma 10$ $\quad U^{\pi}(TL) = 10 + \gamma(-1)$

$U^{\pi}(TR) = -1 + \gamma(-100)$ $\quad U^{\pi}(TR) =. -100 + \gamma(-1)$

POMDP Value Functions

Each conditional plan has a corresponding vector

$\alpha^\pi(s) = U^\pi(s)$



n = number of steps

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

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

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

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

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

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

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

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

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

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

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

Exercise: 2-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

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

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

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

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

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

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

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

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

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

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

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

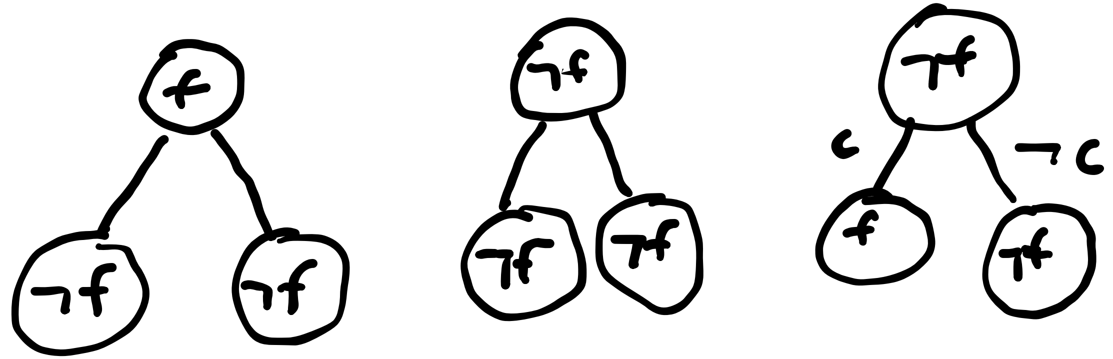
$$\gamma = 0.9$$

Exercise: 2-Step Crying Baby α Vectors

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

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

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



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

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

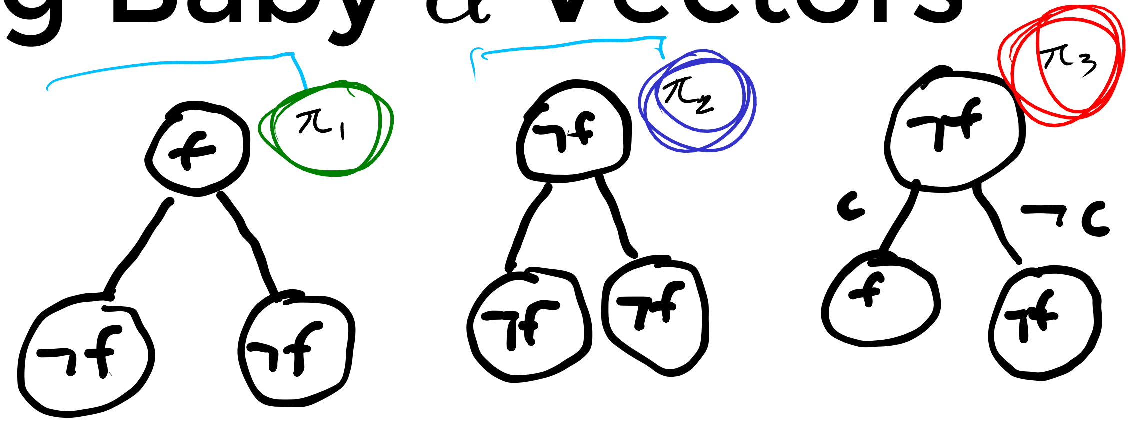
$$R(a) = \begin{cases} -5 \text{ if } a = f \\ 0 \text{ otherwise} \end{cases}$$

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

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

$$\gamma = 0.9$$

Exercise: 2-Step Crying Baby α Vectors



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

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

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

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

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

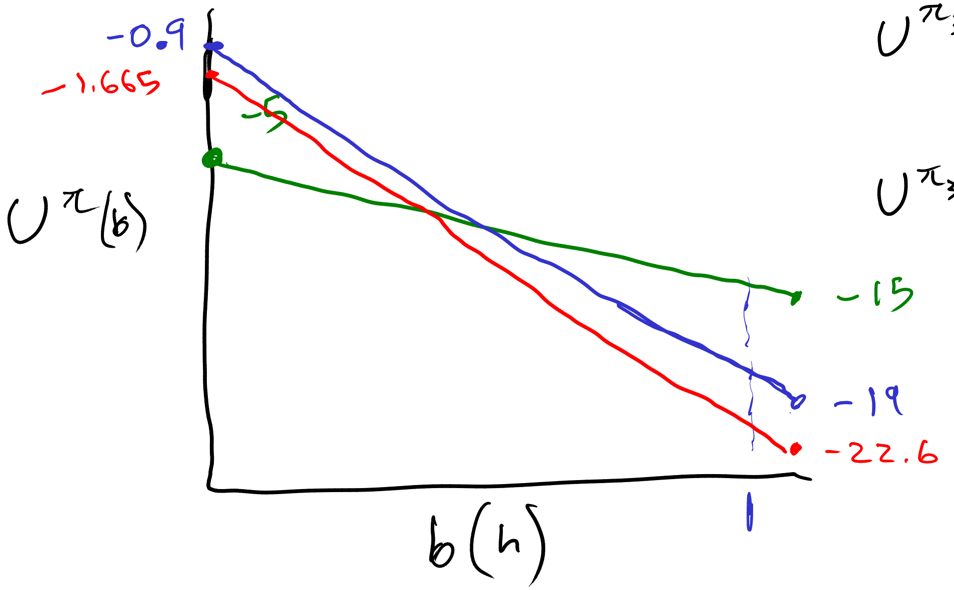
$$R(s) = \begin{cases} -10 \text{ if } s = h \\ 0 \text{ otherwise} \end{cases}$$

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

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

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

$$\gamma = 0.9$$



$U^{\pi_3}(h) = -10 + \gamma(0.8(-15) + 0.2(-10)) =$

-22.6

$U^{\pi_3}(\neg h) = 0$

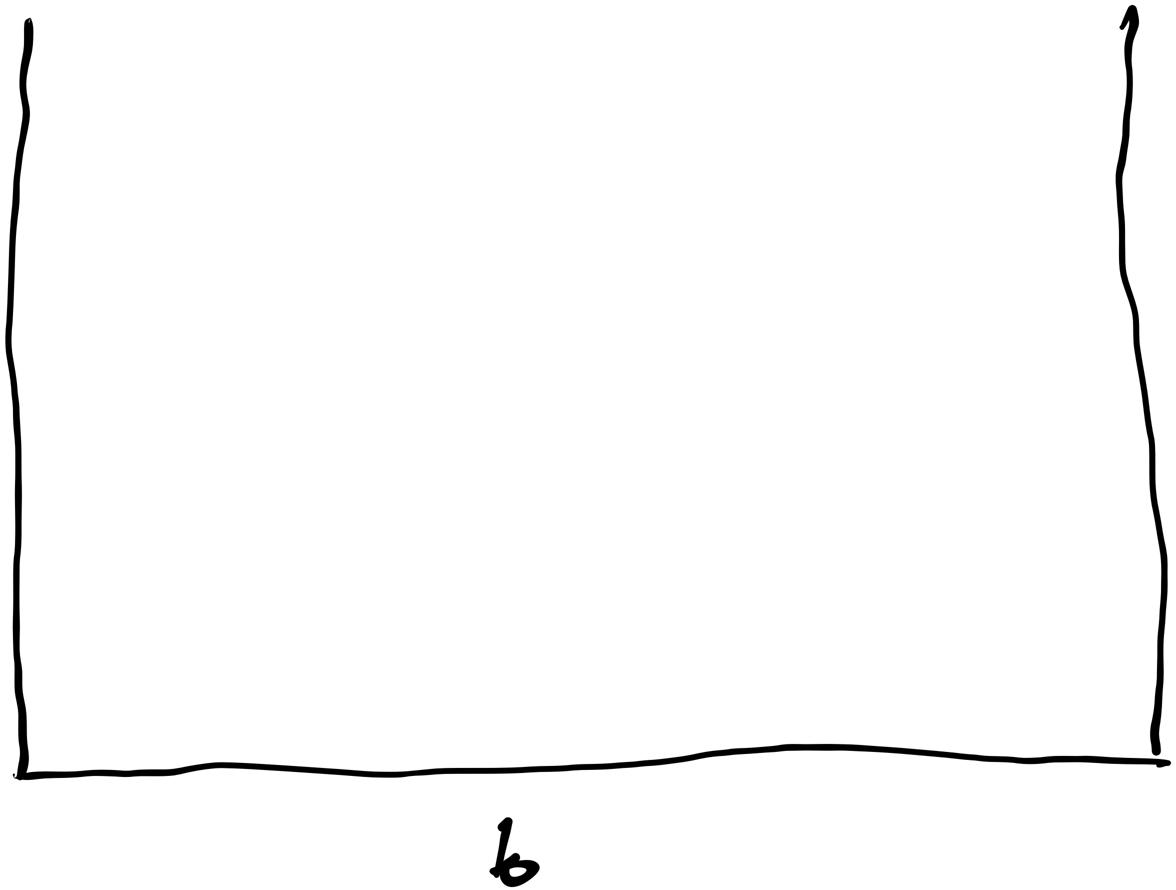
$-\gamma\big(0.1(\underbrace{0.2(-10)}_{\neg c} + \underbrace{0.8(-15)}_{c})\big) \;\}h$

$+0.9(\underbrace{0.9(0)}_{\neg c} + \underbrace{0.1(-5)}_{c}))\big) \;\}\neg h$

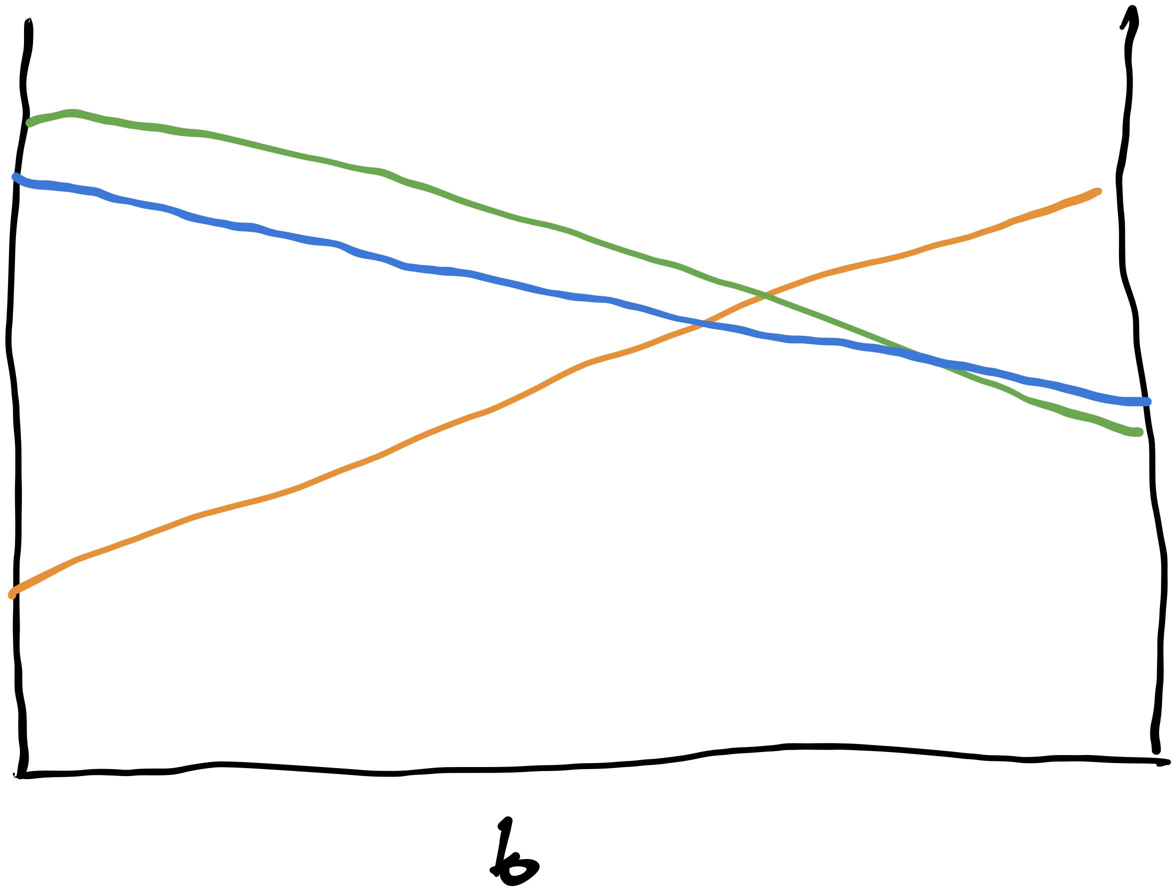
$= -1.665$

α-Vector Pruning

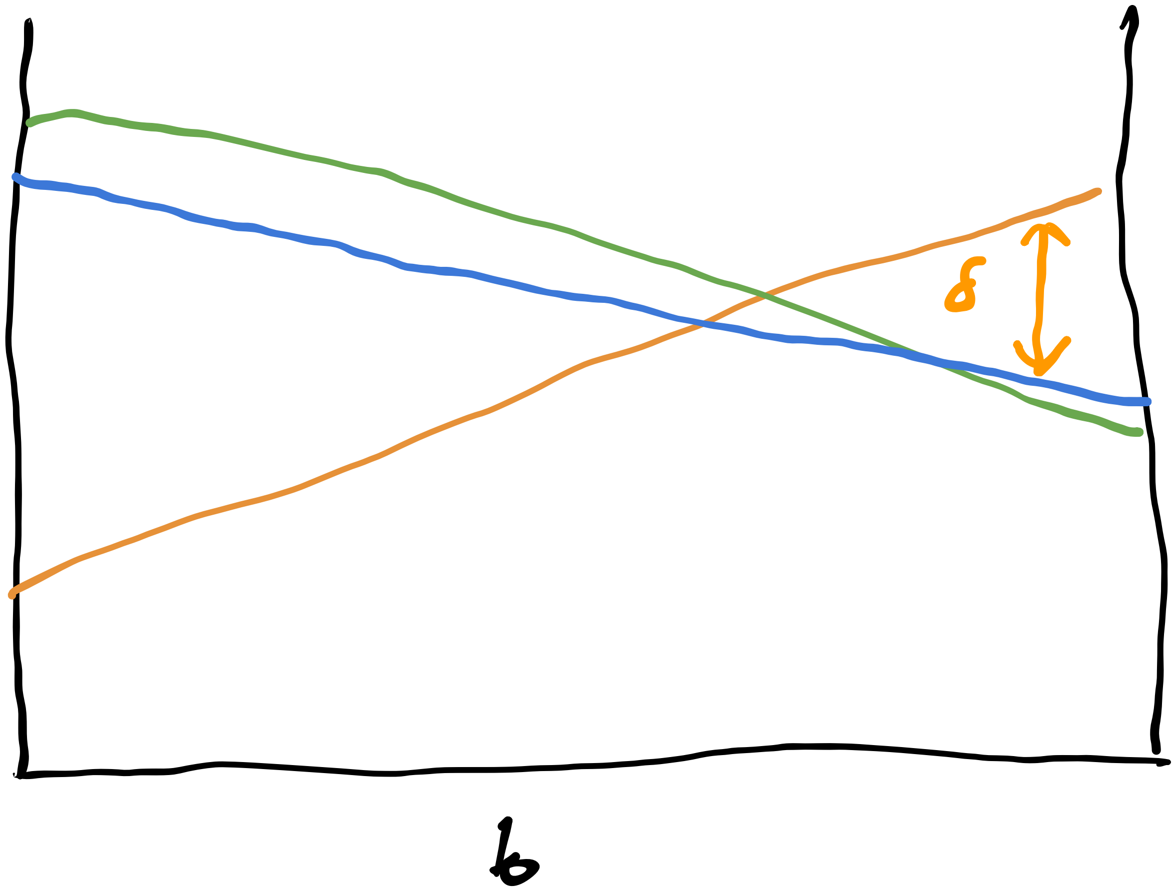
α-Vector Pruning



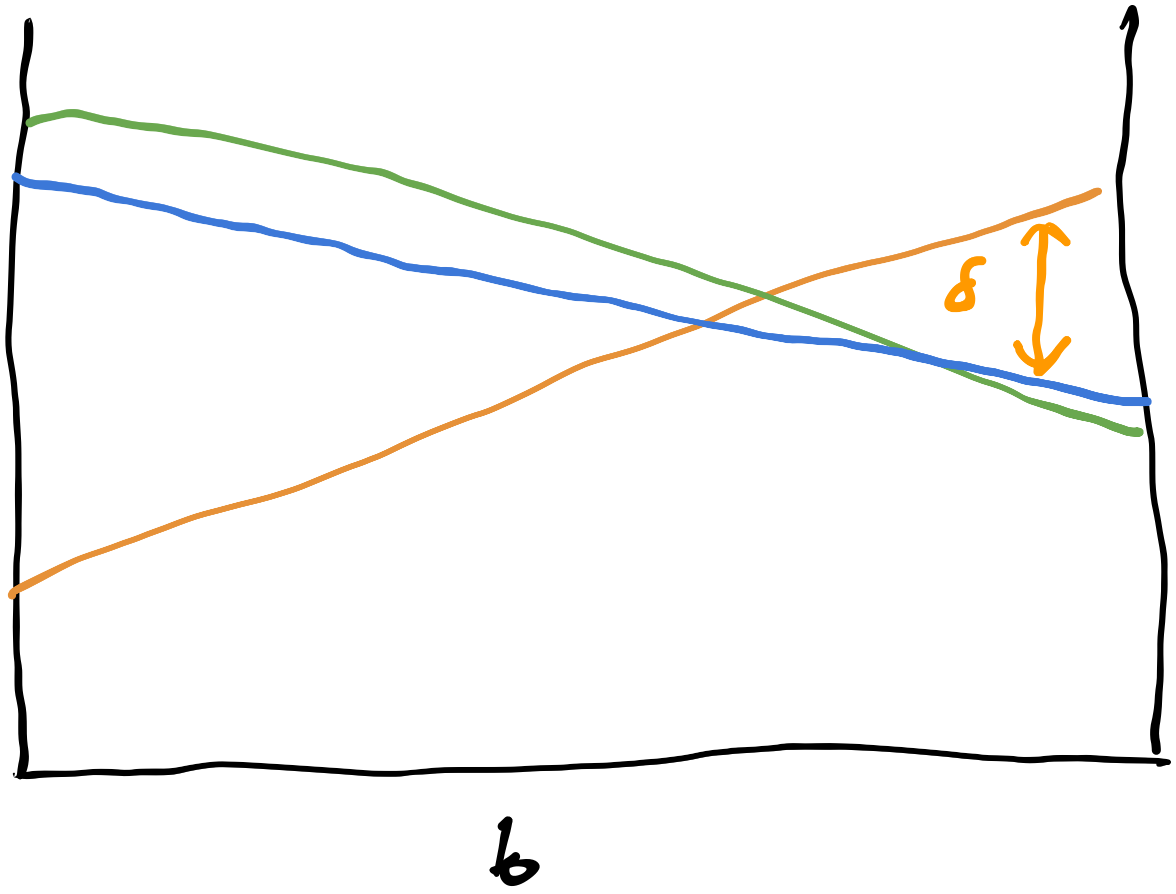
α-Vector Pruning



α-Vector Pruning

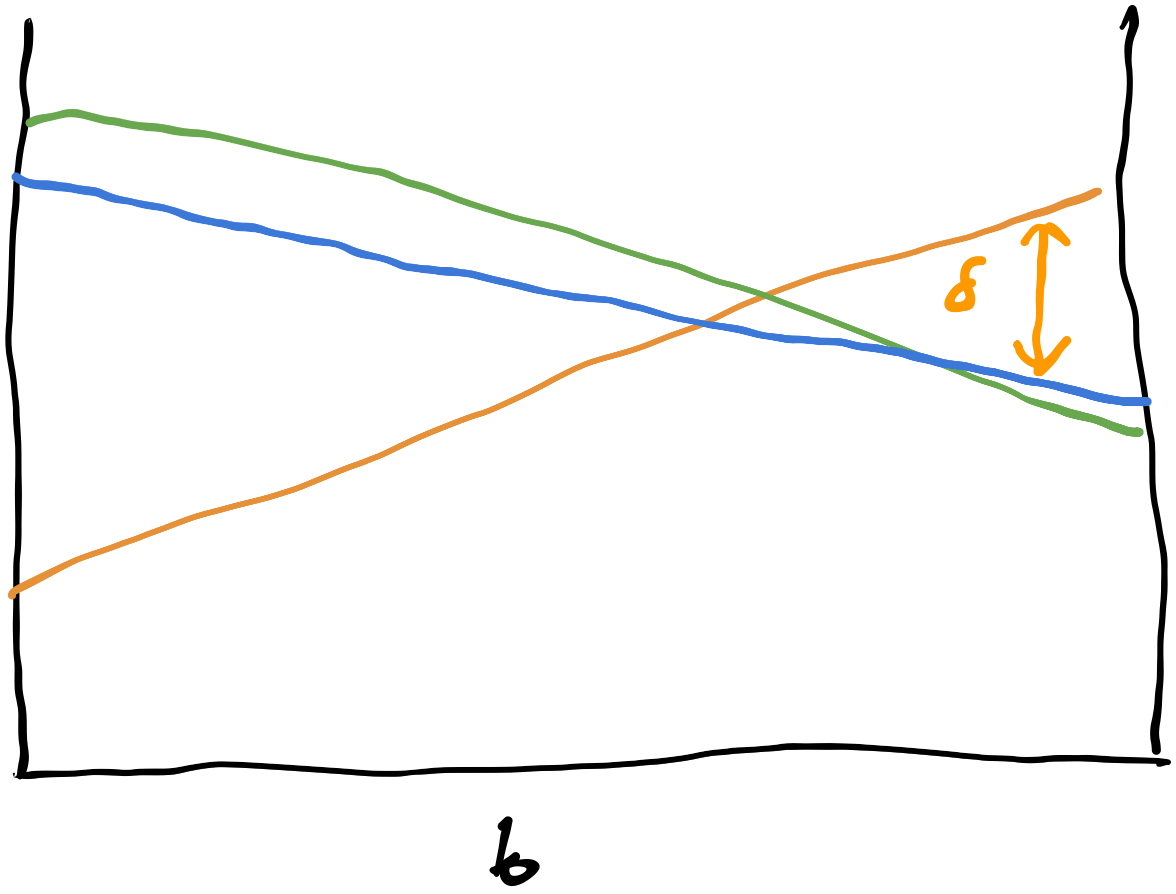


α-Vector Pruning



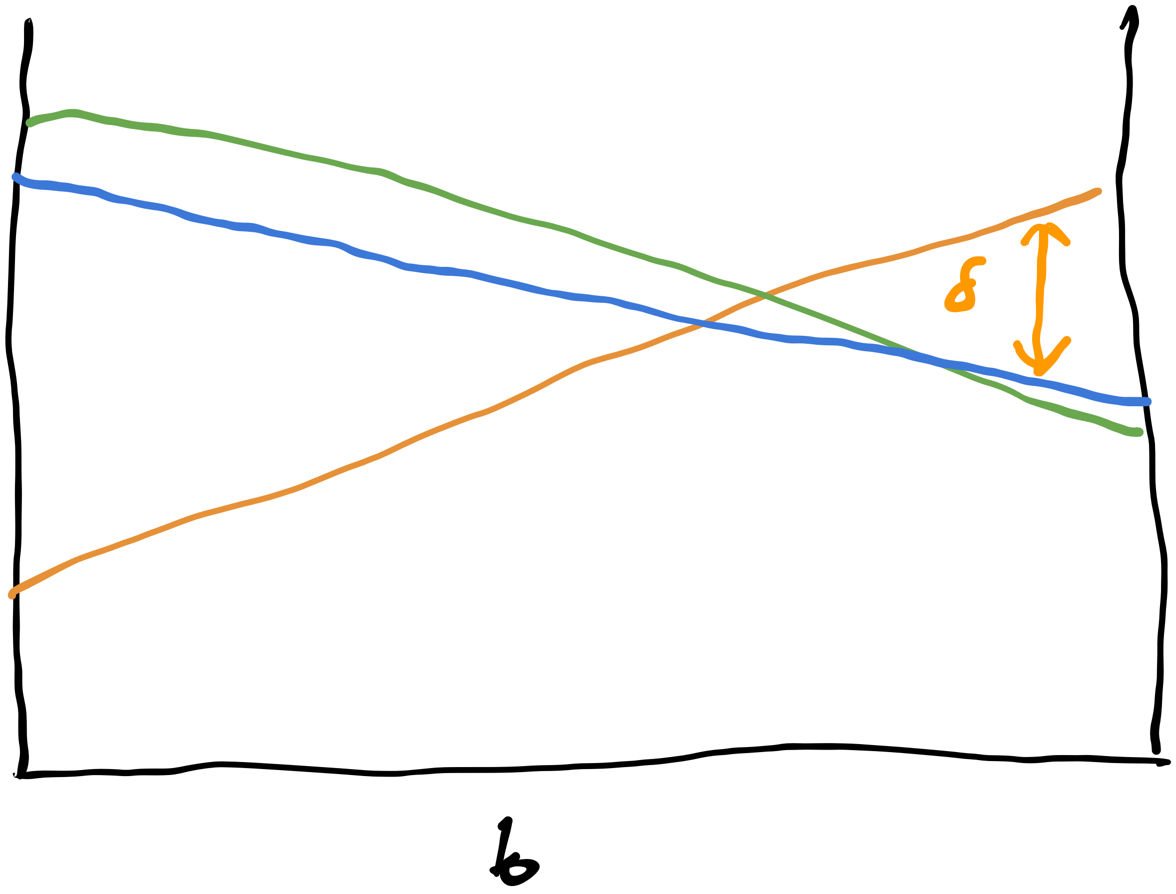
$$\underset{\delta, b}{\text{maximize}} \quad \delta$$

α-Vector Pruning



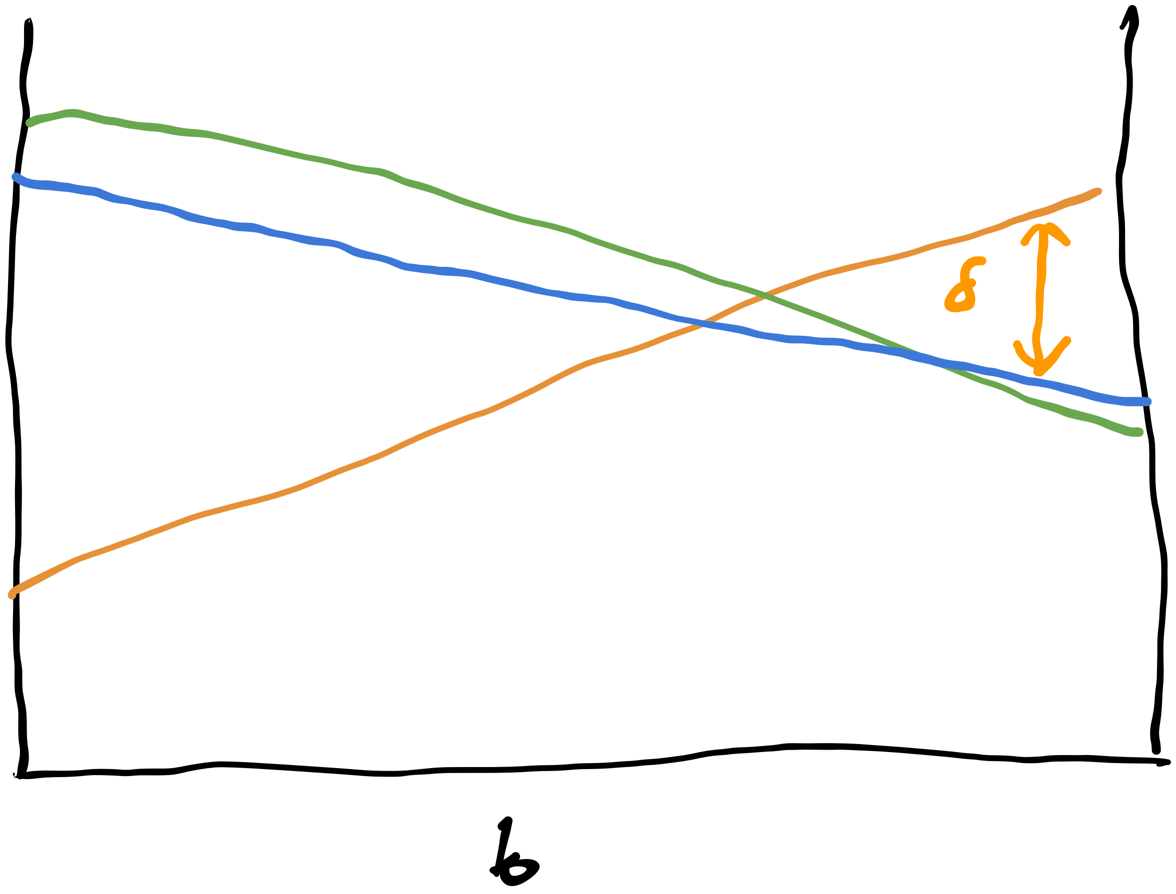
$$\begin{aligned} \underset{\delta, b}{\text{maximize}} \quad & \delta \\ \text{subject to} \quad & b \geq 0 \end{aligned}$$

α-Vector Pruning



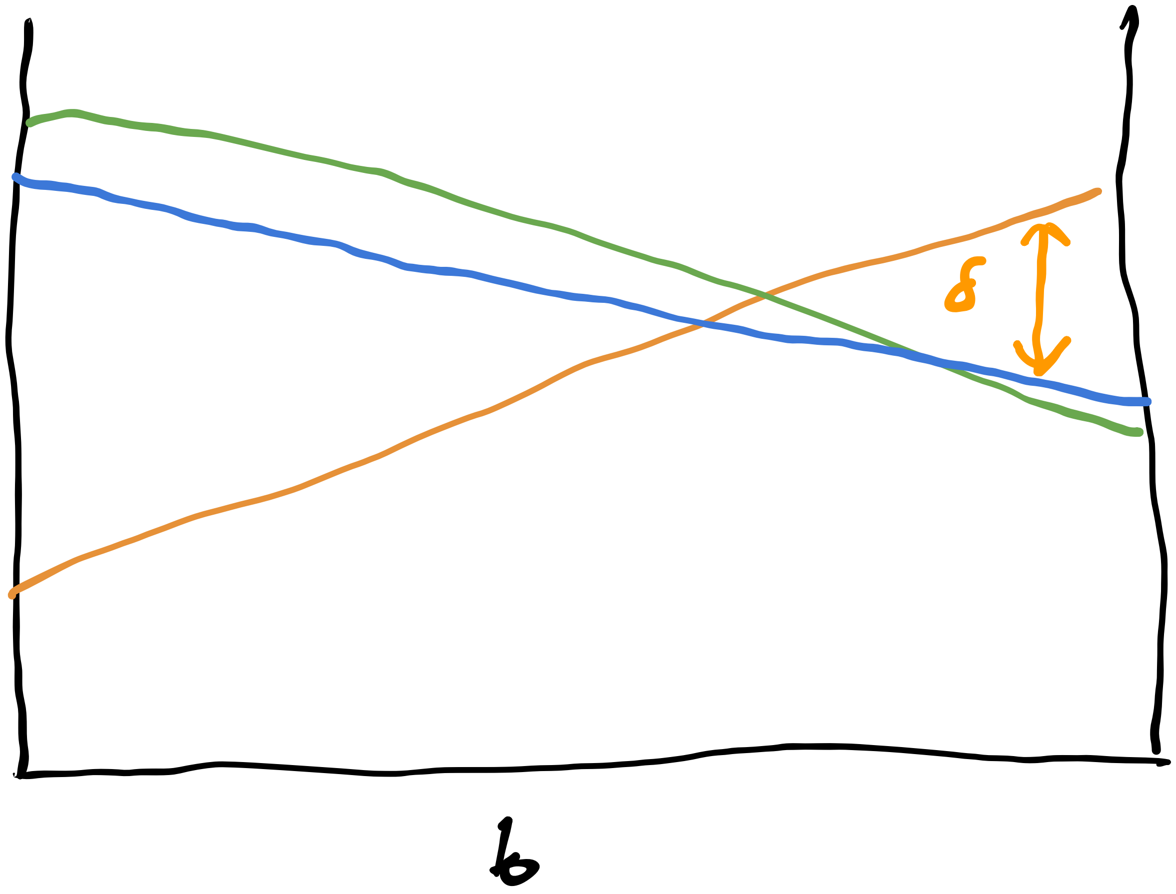
$$
\begin{aligned}
\underset{\delta,b}{\text{maximize}} \quad & \delta \\
\text{subject to} \quad & b \geq 0 \\
& \mathbf{1}^\top b = 1
\end{aligned}
$$

α-Vector Pruning



$$
\begin{aligned}
\underset{\delta, b}{\text{maximize}} \quad & \delta \\
\text{subject to} \quad & b \geq 0 \\
& \mathbf{1}^\top b = 1 \\
& \alpha^\top b \geq \alpha'^\top b + \delta \quad \forall \alpha' \in \Gamma
\end{aligned}
$$

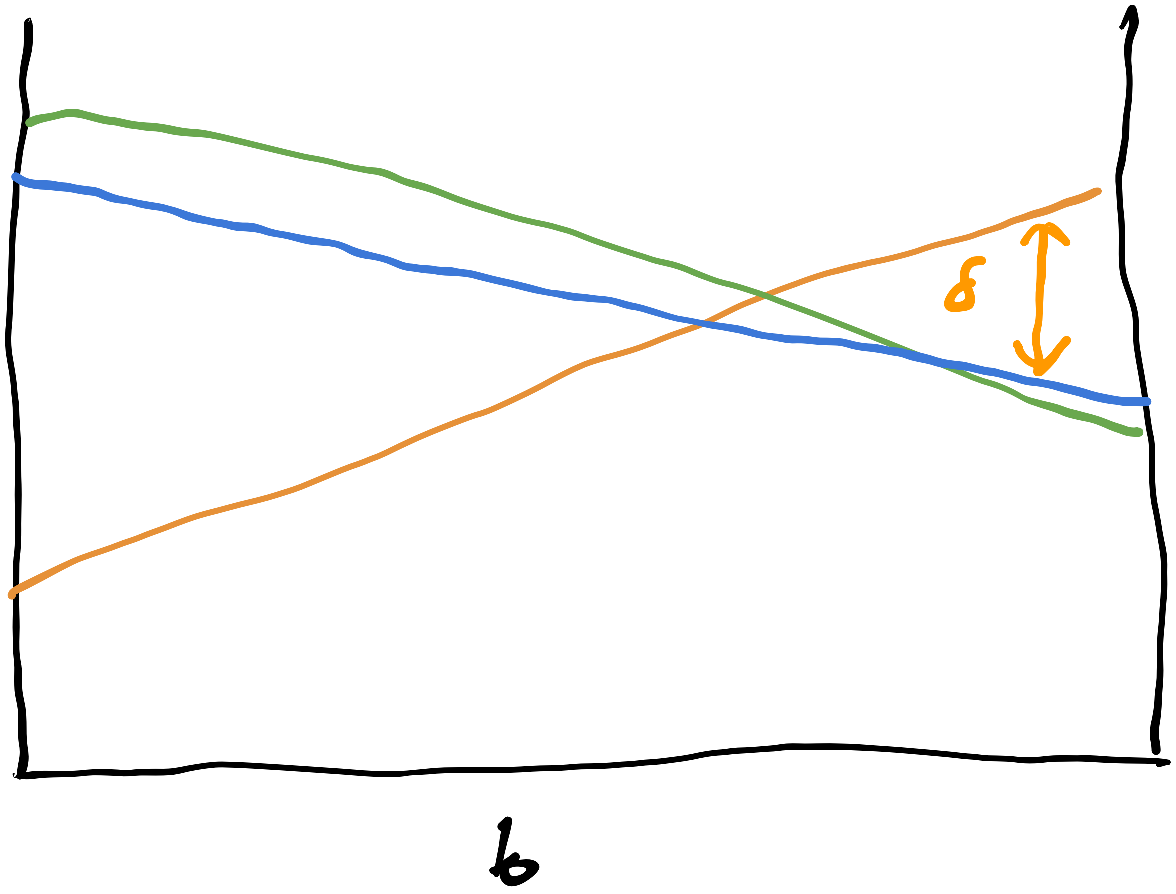
α-Vector Pruning



$$\begin{aligned} \underset{\delta, b}{\text{maximize}} \quad & \delta \\ \text{subject to} \quad & b \geq 0 \\ & \mathbf{1}^\top b = 1 \\ & \alpha^\top b \geq \alpha'^\top b + \delta \quad \forall \alpha' \in \Gamma \end{aligned}$$

"Linear Program"

α-Vector Pruning



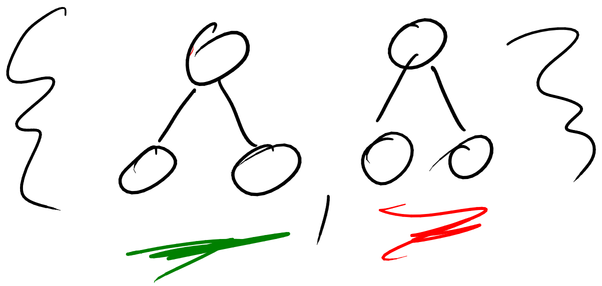
$$
\begin{aligned}
\underset{\delta, b}{\text{maximize}} \quad & \delta \\
\text{subject to} \quad & b \geq 0 \\
& \mathbf{1}^\top b = 1 \\
& \alpha^\top b \geq \alpha'^\top b + \delta \quad \forall \alpha' \in \Gamma
\end{aligned}
$$

"Linear Program"

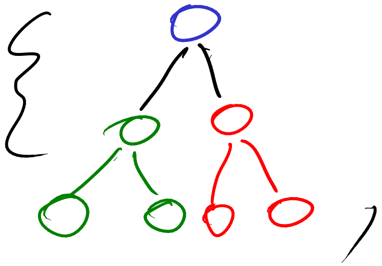
If there is a solution, α is not dominated; b solution sometimes called "witness".

Alpha Vector Expansion

Γ^2



Γ^3



POMDP Value Iteration

$\Gamma^0 = \emptyset$

for $n \in 1 \ldots d$

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

Prune Γ^n

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