Section B | 30 minutes

Developmental Computational Psychiatry lab

HARTLEY LAB

How to Develop a Computational Model?

"All models are wrong, but some are useful"
George E. P. Box

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

Developmental Computational Psychiatry lab

HARTLEY LAB

What we will cover:

- An example for how to select the **proper model with respect to a specific task design**
- The **Rescorla Wagner** model
- The concept of **learning rate**
- The concept of **temperature**
- What is a “**softmax**” function

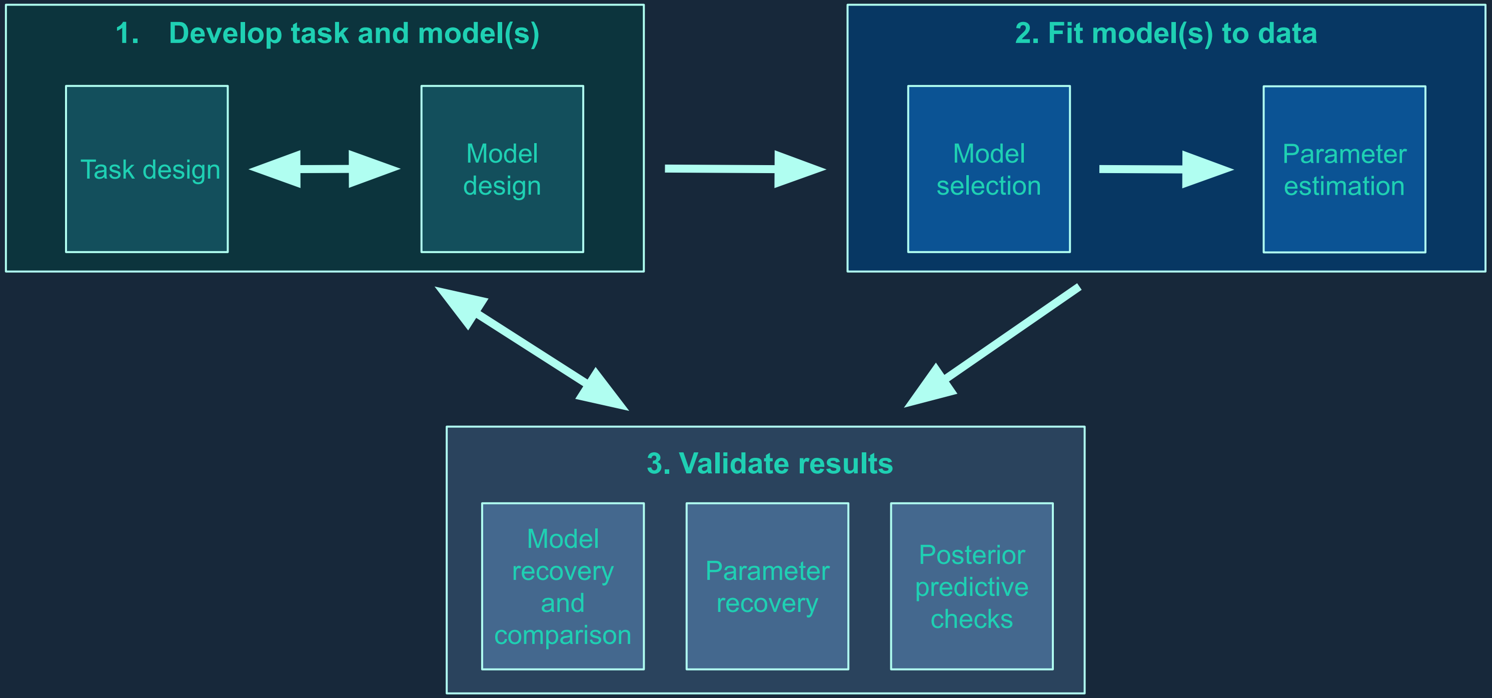
Tricia Seow | Samuel Hewitt | Noam Goldway

flux

Developmental Computational Psychiatry lab

HARTLEY LAB

Developing a computational model



Tricia Seow | Samuel Hewitt | Noam Goldway

flux

Developing a computational model

1. Develop task and model(s)

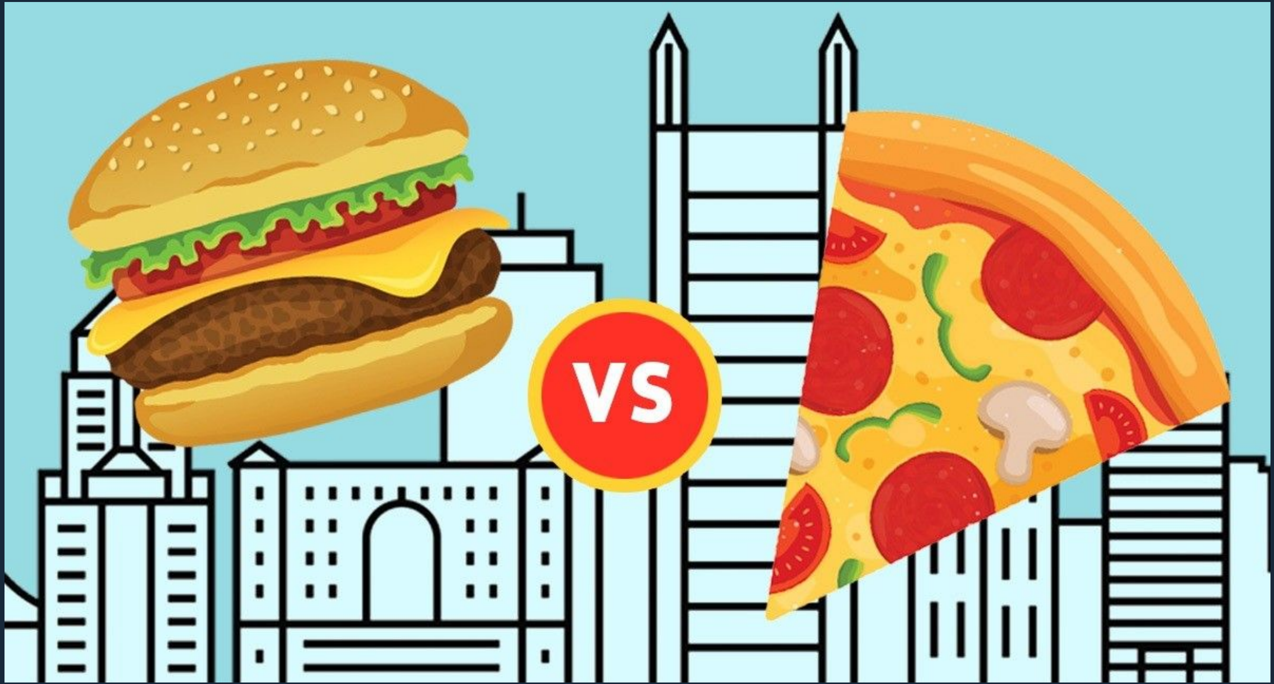
Task design ⟷ Model design

Tricia Seow | Samuel Hewitt | Noam Goldway

Section B | How to Develop a Computational Model?

Developmental Computational Psychiatry lab

HARTLEY LAB

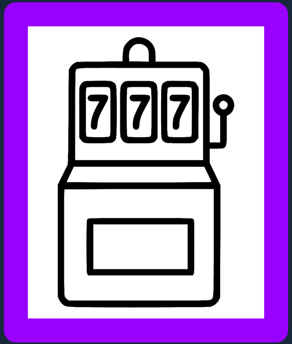


Tricia Seow | Samuel Hewitt | Noam Goldway

flux

2-arm bandit

Experimental task -



Trial	Choice	Outcome
1	Right	0
2	Left new choice	+1
3	Left	+1
4	Right	0
5	Left past experience	+1

How do you maximise reward if you do not know which slot machine is better?

- ➔ Learn expected value of each slot machine
- ➔ Make the next choice based on values learnt

Tricia Seow | Samuel Hewitt | Noam Goldway

Modelling behaviour with RL

Value function: Rescorla Wagner model

$$V_t = V_{t-1} + \alpha(R_t - V_{t-1})$$

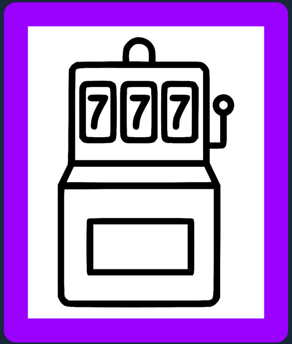
Tricia Seow | Samuel Hewitt | Noam Goldway

Prediction Error

 $V_{prepule} > V_{orange}$ 

Tricia Seow | Samuel Hewitt | Noam Goldway

Prediction Error



$$V_t = V_{t-1} + \alpha(R_t - V_{t-1})$$

Tricia Seow | Samuel Hewitt | Noam Goldway

Developmental Computational Psychiatry lab
HARTLEY LAB

Prediction Error

?

$$V_t = V_{t-1} + \alpha(R_t - 0.5)$$

Developmental Computational Psychiatry lab

HARTLEY LAB

Prediction Error

$$V_t = V_{t-1} + \alpha(1-0.5)$$

flux

Prediction Error

0.5

$$V_t = V_{t-1} + \alpha(1 - 0.5)$$

Developmental Computational Psychiatry lab

HARTLEY LAB

Modelling behaviour with RL

Value function: Rescorla Wagner model

$$V_t = V_{t-1} + \alpha (R_t - V_{t-1})$$

Value
(of the slot machine)

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

Developmental Computational Psychiatry lab

HARTLEY LAB

Modelling behaviour with RL

Value function: Rescorla Wagner model

$$V_t = V_{t-1} + \alpha \, (R_t - V_{t-1})$$

Value (of the slot machine) = Value on previous trial

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

Developmental Computational Psychiatry lab

HARTLEY LAB

Modelling behaviour with RL

Value function: Rescorla Wagner model

$$V_t = V_{t-1} + \alpha (R_t - V_{t-1})$$

Value (of the slot machine) = Value on previous trial

(Reward - Value on previous trial)

Prediction error

what you received - what you expected

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

Developmental Computational Psychiatry lab

HARTLEY LAB

Modelling behaviour with RL

Value function: Rescorla Wagner model

$$V_t = V_{t-1} + \alpha (R_t - V_{t-1})$$

Value (of the slot machine) = Value on previous trial + Learning rate (Reward - Value on previous trial)

$\underbrace{\text{Reward - Value on previous trial}}_{\text{Prediction error}}$

Prediction error

what you received - what you expected

Tricia Seow | Samuel Hewitt | Noam Goldway

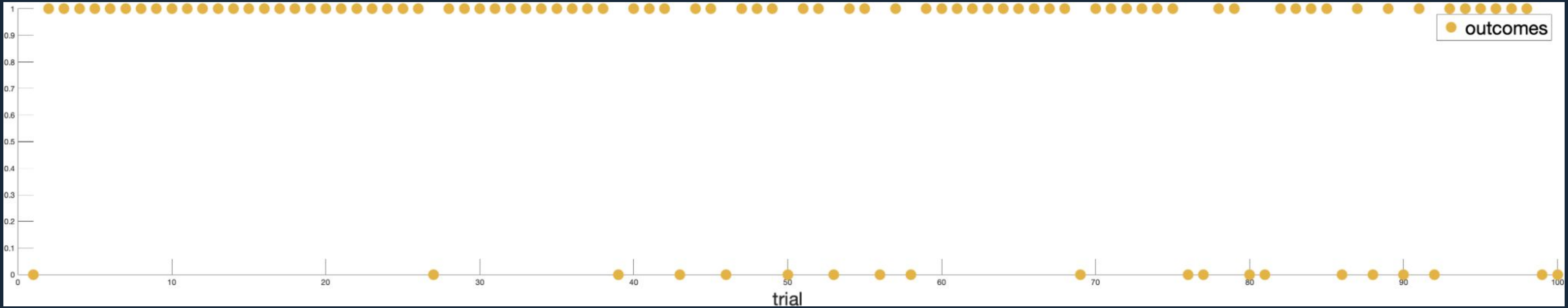
flux

Developmental Computational Psychiatry lab

HARTLEY LAB

Modelling behaviour with RL

$$\text{Value (of the slot machine)} = \text{Value on previous trial} + \text{Learning rate} \left(\overbrace{\text{Reward} - \text{Value on previous trial}}^{\substack{\text{Prediction error} \\ \text{what you received - what you expected}}} \right)$$



Tricia Seow | Samuel Hewitt | Noam Goldway

https://github.com/DevComPsy/2021FluxCompModellingWorkshop

flux

Section B | How to Develop a Computational Model?

Developmental Computational Psychiatry lab

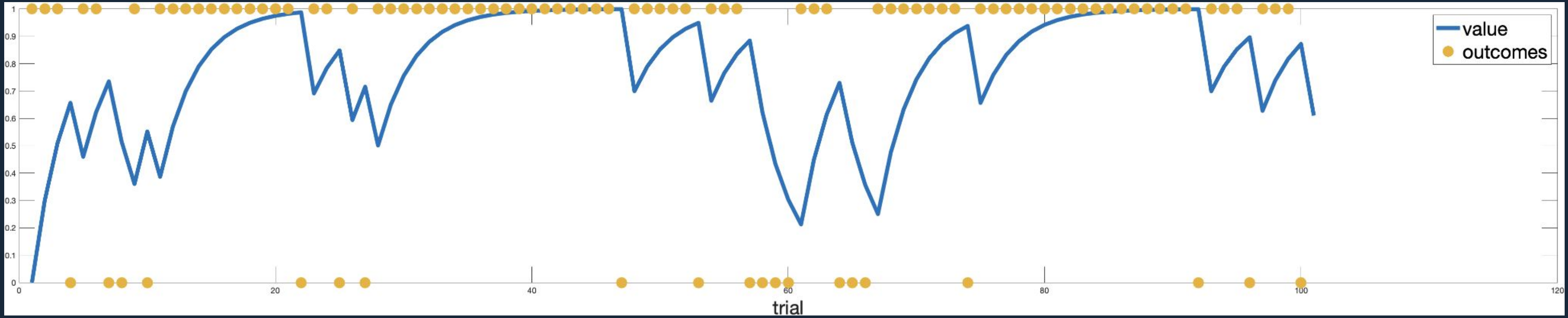
HARTLEY LAB

Modelling behaviour with RL

Prediction error

what you received - what you expected

Value (of the slot machine) = Value on previous trial + Learning rate ($\underbrace{\textbf{Reward} - \text{Value on previous trial}}_{}$)



Tricia Seow | Samuel Hewitt | Noam Goldway

https://github.com/DevComPsy/2021FluxCompModellingWorkshop

flux

Developmental Computational Psychiatry lab

HARTLEY LAB

Modelling behaviour with RL

$$\text{Value (of the slot machine)} = \text{Value on previous trial} + \text{Learning rate}\,\overbrace{(\text{Reward} - \text{Value on previous trial})}^{\substack{\text{Prediction error} \\ \text{what you received - what you expected}}}$$

Trial 1	?				

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

Modelling behaviour with RL

Prediction error

what you received - what you expected

Value (of the slot machine) = Value on previous trial + Learning rate (Reward - Value on previous trial)

Trial 1	?	*initiation:* 0.5			*initiation:* 0.5

Developmental Computational Psychiatry lab

HARTLEY LAB

Modelling behaviour with RL

Prediction error

what you received - what you expected

Value (of the slot machine) = Value on previous trial + Learning rate (Reward - Value on previous trial)

Trial 1	?	0.5		*1*	0.5

Developmental Computational Psychiatry lab

HARTLEY LAB

Modelling behaviour with RL

Prediction error

what you received - what you expected

Value (of the slot machine) = Value on previous trial + Learning rate (Reward - Value on previous trial)

0.5

Trial 1	?	0.5	+	*1*	*0.5*

Tricia Seow | Samuel Hewitt | Noam Goldway

Developmental Computational Psychiatry lab

HARTLEY LAB

Modelling behaviour with RL

$$\text{Value (of the slot machine)} = \text{Value on previous trial} + \text{Learning rate}\,(\underbrace{\text{Reward} - \text{Value on previous trial}}_{\substack{\text{Prediction error} \\ \text{what you received - what you expected}}})$$

Trial 1	*1*	0.5	+	*0.5*

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

Developmental Computational Psychiatry lab

HARTLEY LAB

Modelling behaviour with RL

$$\underset{\text{(of the slot machine)}}{\text{Value}} = \underset{\text{previous trial}}{\text{Value on}} + \underset{\text{rate}}{\text{Learning}} \Big(\overbrace{\text{Reward} - \underset{\text{previous trial}}{\text{Value on}}}^{\substack{\text{Prediction error} \\ \text{what you received - what you expected}}} \Big)$$

Trial 1	*1*	0.5	+	*0.5*
Trial 2		1	+	1

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

Developmental Computational Psychiatry lab

HARTLEY LAB

Modelling behaviour with RL

Prediction error

what you received - what you expected

Value (of the slot machine) = Value on previous trial + Learning rate (Reward - Value on previous trial)

Trial 1	*1*	0.5	+	*0.5*	
Trial 2		1	+	1	1

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

Developmental Computational Psychiatry lab

HARTLEY LAB

Modelling behaviour with RL

Prediction error

what you received - what you expected

$$\text{Value (of the slot machine)} = \text{Value on previous trial} + \text{Learning rate}\,(\underbrace{\text{Reward} - \text{Value on previous trial}}_{})$$

Trial 1	*1*	0.5	+	*0.5*
Trial 2		1	+	(1 - 1)

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

Developmental Computational Psychiatry lab

HARTLEY LAB

Modelling behaviour with RL

$$\underbrace{\text{Value (of the slot machine)}} = \text{Value on previous trial} + \text{Learning rate}\left(\overbrace{\text{Reward} - \text{Value on previous trial}}^{\substack{\text{Prediction error} \\ \text{what you received - what you expected}}}\right)$$

Trial 1	*1*	0.5	+	*0.5*
Trial 2		1	+	*0*

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

Developmental Computational Psychiatry lab

HARTLEY LAB

Modelling behaviour with RL

Value (of the slot machine) = Value on previous trial + Learning rate (Reward - Value on previous trial)

$$\underbrace{\text{Reward} - \text{Value on previous trial}}_{\substack{\text{Prediction error} \\ \text{what you received - what you expected}}}$$

Trial 1	*1*	0.5	+	*0.5*
Trial 2	***1***	1	+	*0*

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

Developmental
Computational
Psychiatry lab

HARTLEY LAB

Modelling behaviour with RL

$$\text{Value (of the slot machine)} = \text{Value on previous trial} + \text{Learning rate}\,(\underbrace{\text{Reward} - \text{Value on previous trial}}_{\text{Prediction error: what you received - what you expected}})$$

Trial 1	*1*	0.5	+	*0.5*
Trial 2	= *1*	1	+	*0*

Tricia Seow | Samuel Hewitt | Noam Goldway

Modelling behaviour with RL

$$\text{Value (of the slot machine)} = \text{Value on previous trial} + \text{Learning rate}\,(\underbrace{\text{Reward} - \text{Value on previous trial}}_{\substack{\text{Prediction error} \\ \text{what you received - what you expected}}})$$

Trial 1	*1*	0.5	+	*0.5*
Trial 2	*1*	1	+	*0*
Trial 3		1	+	1

Tricia Seow | Samuel Hewitt | Noam Goldway

Modelling behaviour with RL

Prediction error

what you received - what you expected

Value (of the slot machine) = Value on previous trial + Learning rate (Reward - Value on previous trial)

Trial 1	*1*	0.5	+	*0.5*
Trial 2	*1*	1	+	*0*
Trial 3		1	+	*0* 1

Tricia Seow | Samuel Hewitt | Noam Goldway

Developmental Computational Psychiatry lab

HARTLEY LAB

Modelling behaviour with RL

Prediction error

what you received - what you expected

Value (of the slot machine) = Value on previous trial + Learning rate (Reward - Value on previous trial)

Trial 1	*1*	0.5	+	*0.5*
Trial 2	*1*	1	+	*0*
Trial 3		1	+	-1

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

Developmental Computational Psychiatry lab

HARTLEY LAB

Modelling behaviour with RL

$$\text{Value (of the slot machine)} = \text{Value on previous trial} + \text{Learning rate}\,\Big(\overbrace{\text{Reward} - \text{Value on previous trial}}^{\substack{\text{Prediction error} \\ \text{what you received - what you expected}}}\Big)$$

Trial 1	*1*	0.5	+	*0.5*
Trial 2	*1*	1	+	*0*
Trial 3	***0***	1	+	-1

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

Developmental Computational Psychiatry lab
HARTLEY LAB

Modelling behaviour with RL

Prediction error

what you received - what you expected

Value (of the slot machine) = Value on previous trial + Learning rate (Reward - Value on previous trial)

$$V_t = V_{t-1} + \alpha \, (R_t - V_{t-1})$$

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

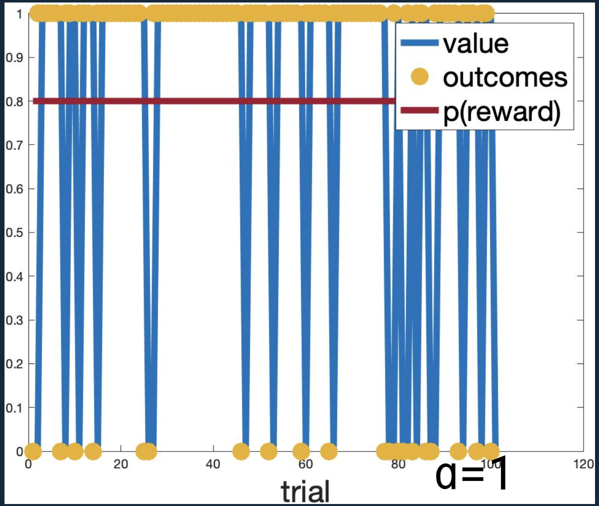
Developmental Computational Psychiatry lab

HARTLEY LAB

How much should we learn?

$$V_t = V_{t-1} + \alpha(R_t - V_{t-1})$$

What happens if we manipulate learning rate?



Tricia Seow | Samuel Hewitt | Noam Goldway

flux

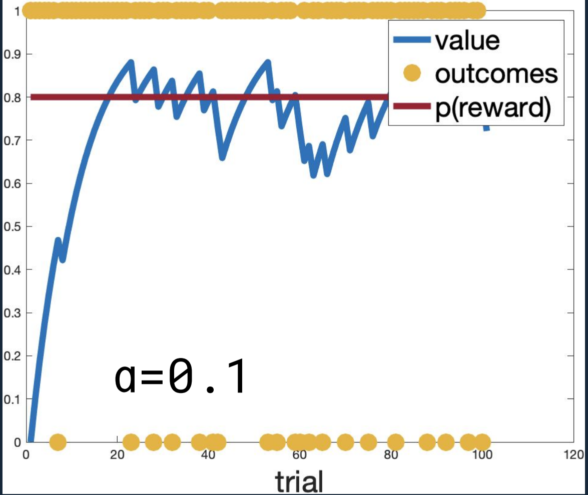
Developmental Computational Psychiatry lab

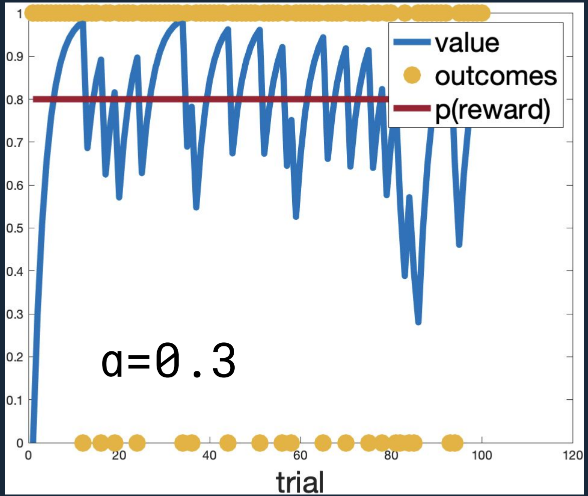
HARTLEY LAB

How much should we learn?

What happens if we manipulate learning rate?

$V_t = V_{t-1} + \alpha(R_t - V_{t-1})$





Tricia Seow | Samuel Hewitt | Noam Goldway

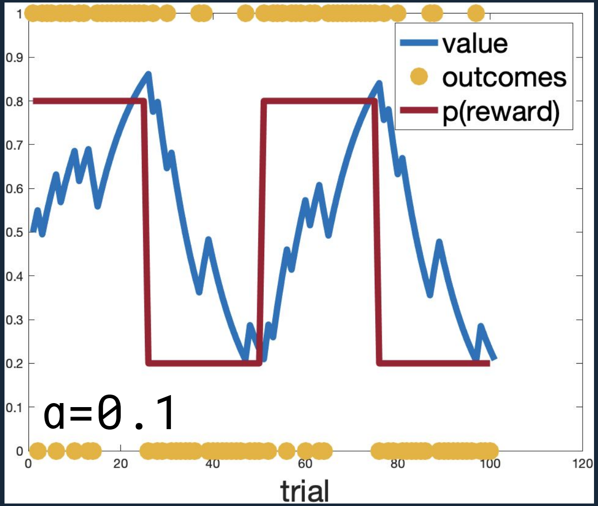
Developmental Computational Psychiatry lab

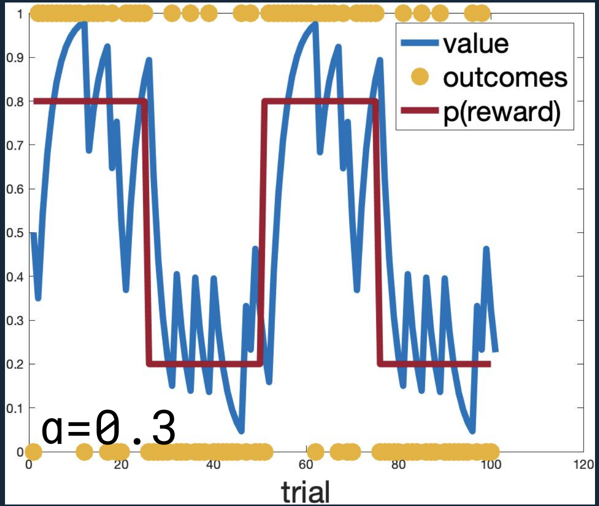
HARTLEY LAB

How much should we learn?

What happens if we manipulate learning rate?

$$V_t = V_{t-1} + \alpha(R_t - V_{t-1})$$





Tricia Seow | Samuel Hewitt | Noam Goldway

flux

Developmental Computational Psychiatry lab

HARTLEY LAB

Is low learning rate always better?

$$V_t = V_{t-1} + \alpha(R_t - V_{t-1})$$

→ Depend on the statistics of the environment
- Low volatility-> low α is better
- High volatility-> high α is better

Tricia Seow | Samuel Hewitt | Noam Goldway

Developmental Computational Psychiatry lab

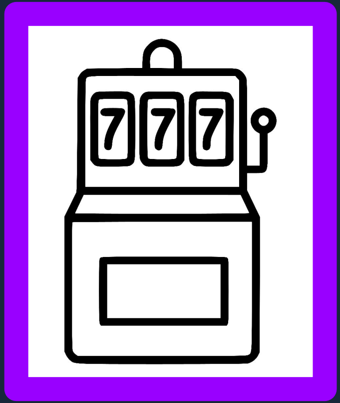
HARTLEY LAB

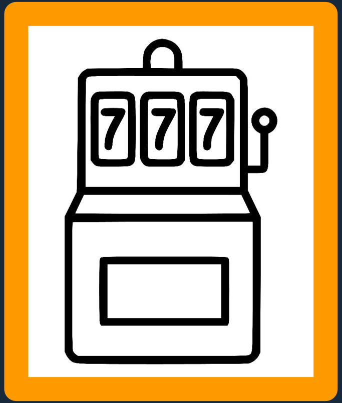
What did we learn so far

- ➔ What are multi arm bandit tasks
- ➔ How RL and, specifically Rescorla Wagner model can help us to 'solve' such problems
- ➔ Expected value
- ➔ Prediction error
- ➔ High vs low learning rate

How should we choose?

How should we choose?





80% reward

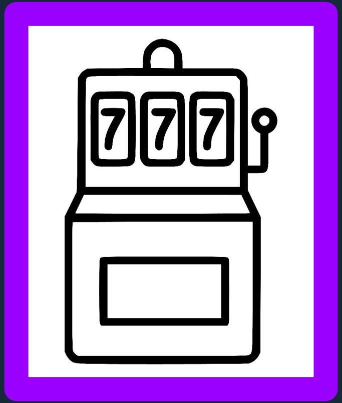
20% reward

← learnt via trial and error (value function) →

Developmental Computational Psychiatry lab

HARTLEY LAB

How should we choose?



80% reward

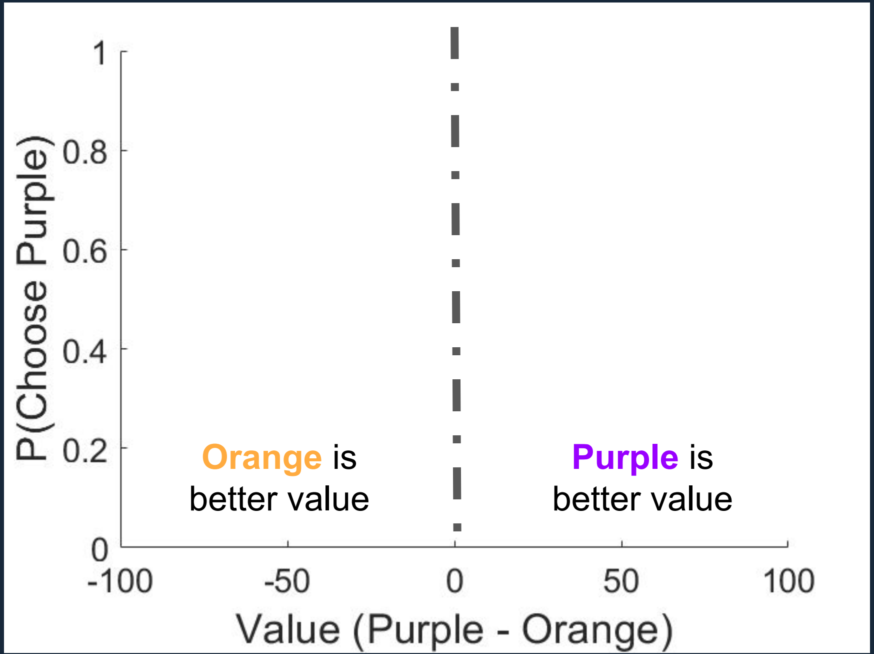
Maximise rewards

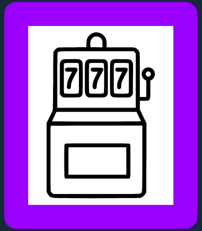
- ➔ Pick slot machine with largest likelihood of reward
- ➔ Exploit

Developmental Computational Psychiatry lab

HARTLEY LAB

How should we choose?



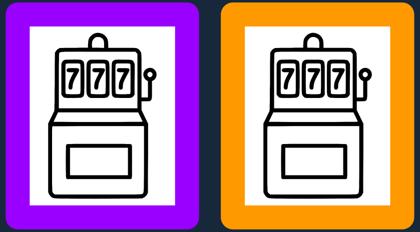


flux

Developmental Computational Psychiatry lab

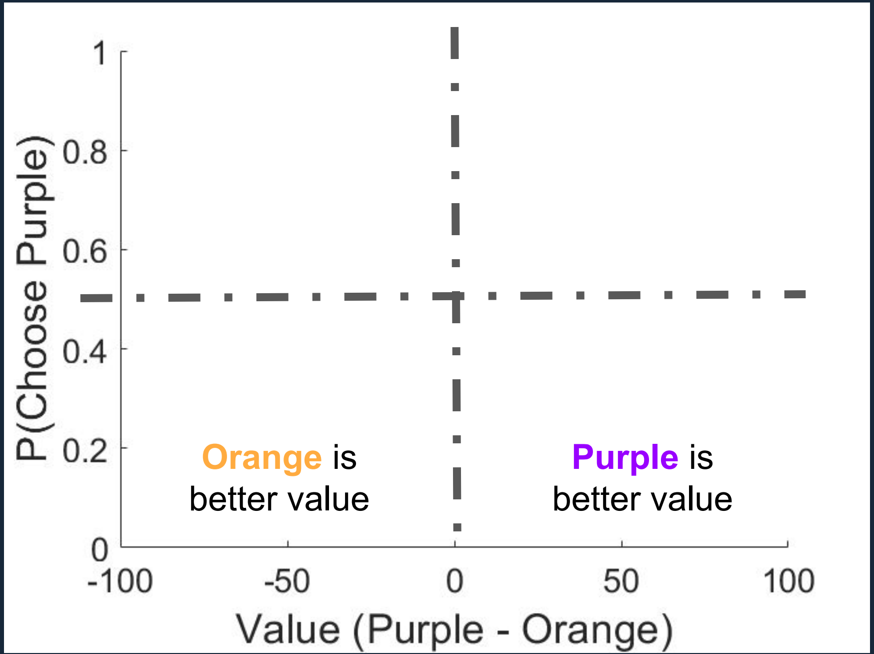
HARTLEY LAB

How should we choose?



Choose **purple** is better

Choose **orange** is better



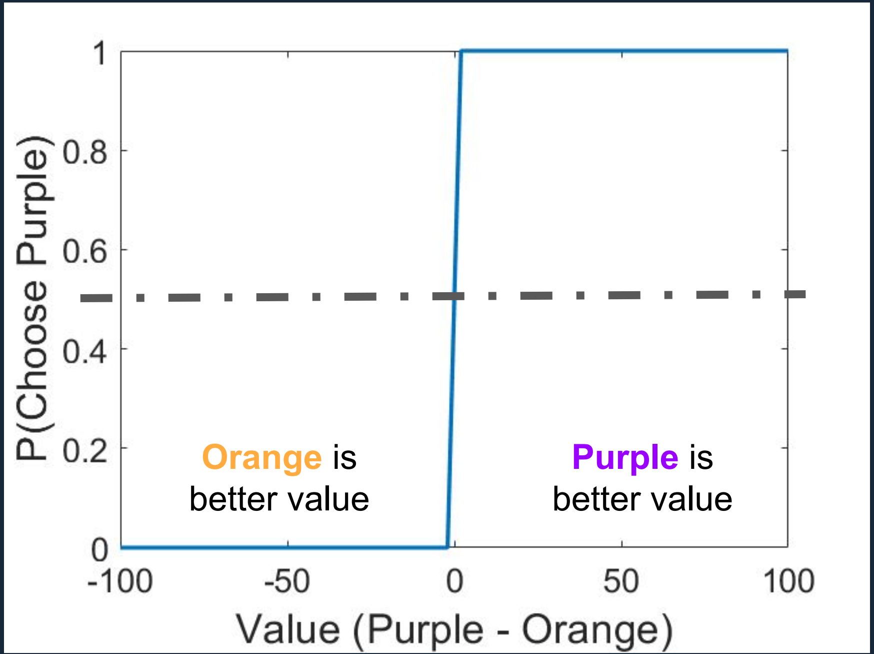
Tricia Seow | Samuel Hewitt | Noam Goldway

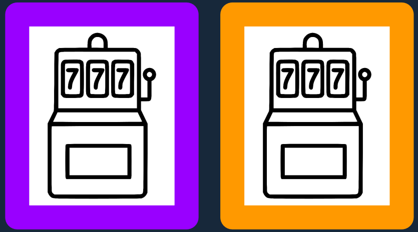
flux

How should we choose?

Choose **purple** is better

Choose **orange** is better





Exploit

➔ Choose slot machine when reward is better than the other

Tricia Seow | Samuel Hewitt | Noam Goldway

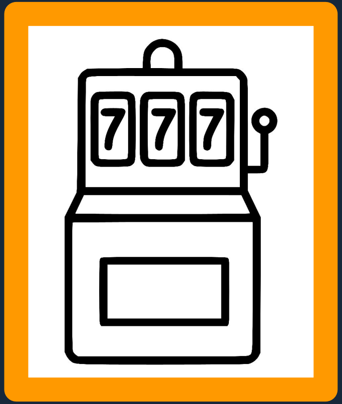
Developmental Computational Psychiatry lab

HARTLEY LAB

How should we choose?

Try other options

- Sample the outcomes of the other slot machine
- Explore



20% reward

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

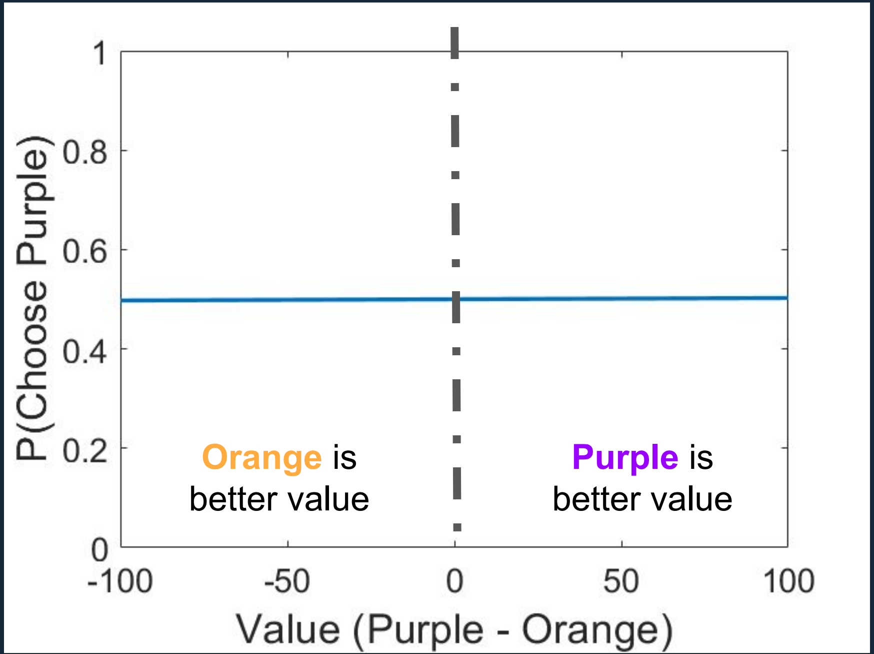
Developmental Computational Psychiatry lab

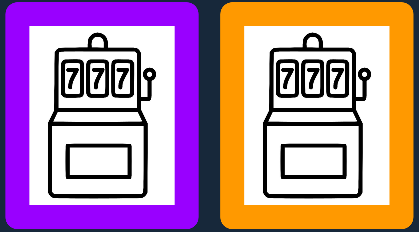
HARTLEY LAB

How should we choose?

Choose **purple** is better

Choose **orange** is better





Explore

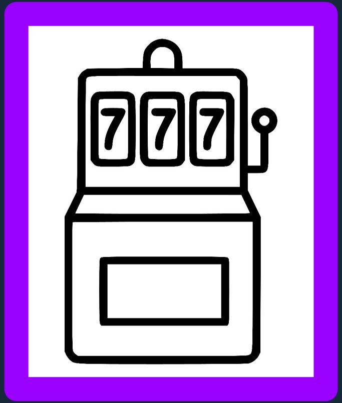
➔ Choose slot machine equally

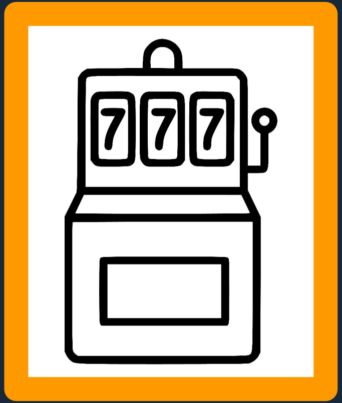
Tricia Seow | Samuel Hewitt | Noam Goldway

Developmental Computational Psychiatry lab

HARTLEY LAB

How should we choose?





Exploit → an individual difference ← *Explore*

we can model as a **free parameter**

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

How should we choose?

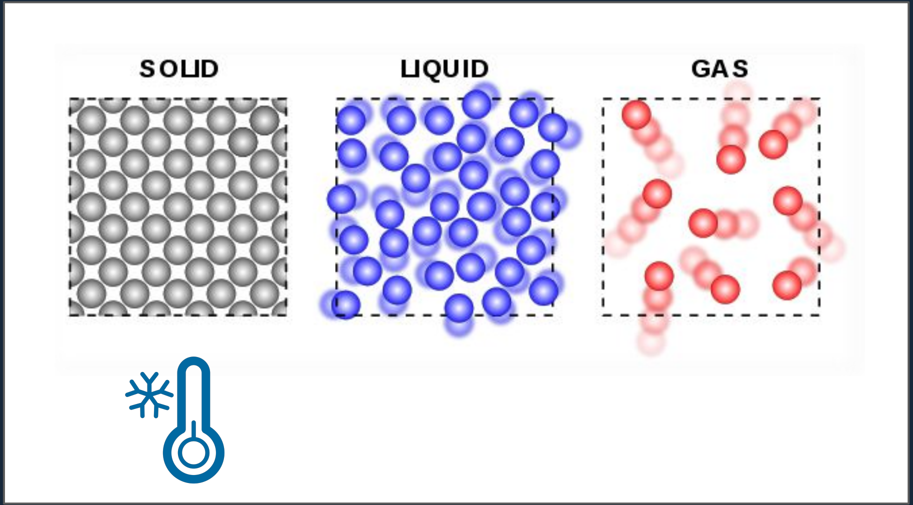
Temperature (τ): parameter that determines the extent to which value estimates influence choice behaviour

How should we choose?

Temperature (τ): parameter that determines the extent to which value estimates influence choice behaviour

Low temperature

➔ Choices are less noisy



Developmental
Computational
Psychiatry lab

HARTLEY LAB

How should we choose?

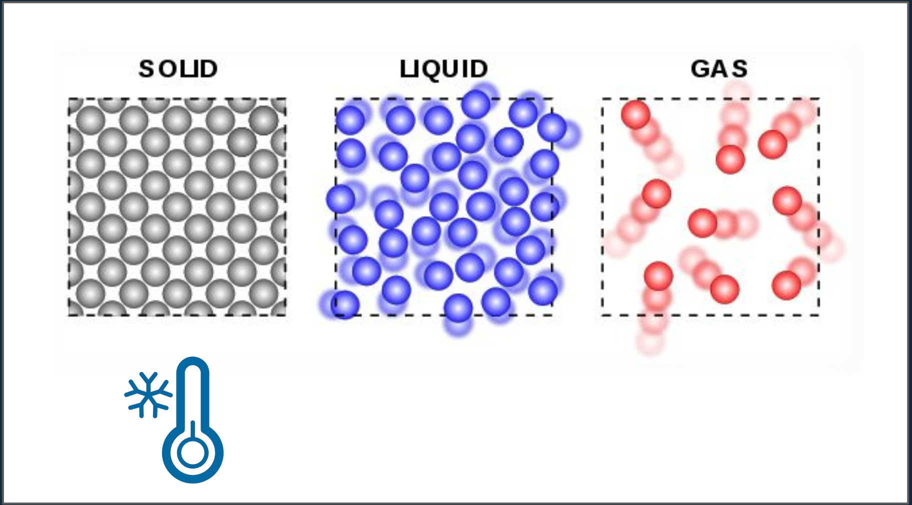
Temperature (τ): parameter that determines the extent to which value estimates influence choice behaviour

Low temperature

→ Choices are less noisy

→ More affected by value

→ More deterministic



Developmental Computational Psychiatry lab

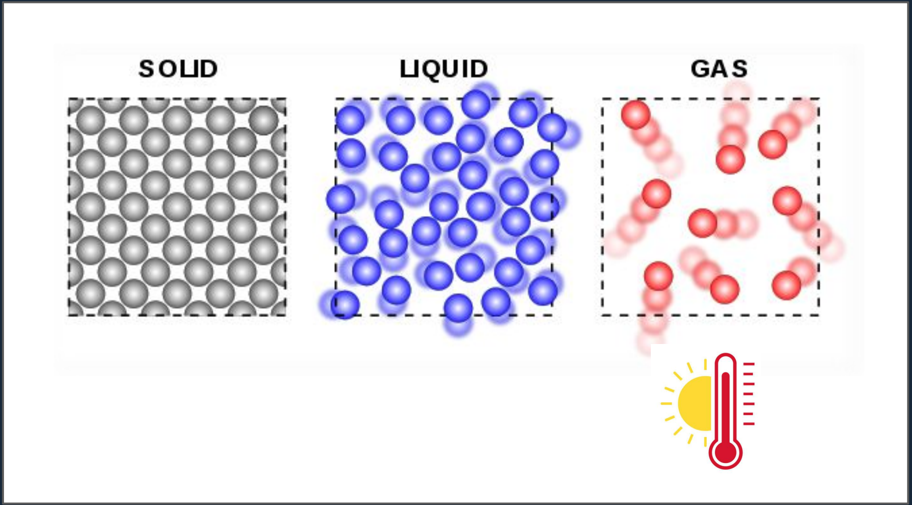
HARTLEY LAB

How should we choose?

Temperature (τ): parameter that determines the extent to which value estimates influence choice behaviour

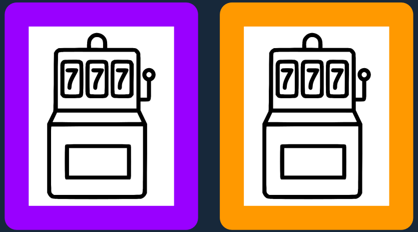
High temperature

- ➔ Choices are more noisy
- ➔ Less affected by value
- ➔ Less deterministic



How should we choose?

Temperature (τ): parameter that determines the extent to which value estimates influence choice behaviour



➔ Let's assume that if we don't pick purple we will pick orange; and vice versa

Developmental Computational Psychiatry lab

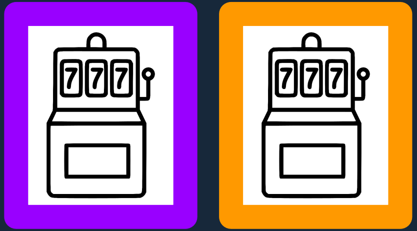
HARTLEY LAB

How should we choose?

Temperature (τ): parameter that determines the extent to which value estimates influence choice behaviour

Softmax equation:

$$P(\text{purple}) = \frac{\exp([\,V_{\text{purple}}\,]/\,\tau)}{\text{SUM}[\exp([\,V_{\text{purple}}, V_{\text{orange}}\,]/\,\tau)]}$$



→ Let's assume that if we don't pick purple we will pick orange; and vice versa

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

How should we choose?

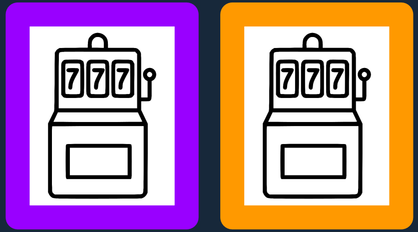
Temperature (τ): parameter that determines the extent to which value estimates influence choice behaviour

Softmax equation:

$$P(\text{purple}) = \frac{\exp([V_{\text{purple}}]/\tau)}{\text{SUM}[\exp([V_{\text{purple}}, V_{\text{orange}}]/\tau)]}$$

Probability of choosing purple

→ P(orange) = 1 - P(purple)



→ Let's assume that if we don't pick purple we will pick orange; and vice versa

Tricia Seow | Samuel Hewitt | Noam Goldway

Developmental Computational Psychiatry lab

HARTLEY LAB

How should we choose?

Temperature (τ): parameter that determines the extent to which value estimates influence choice behaviour

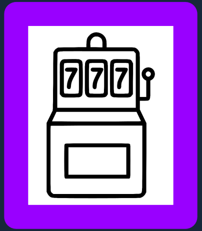
Softmax equation:

Value of machines

$$P(\text{purple}) = \frac{\exp([\ V_{purple}\ ]/\ \tau)}{\text{SUM}[\exp([\ V_{purple}, V_{orange}\ ]/\ \tau)]}$$

Probability of choosing **purple**

➔ P(orange) = 1 - P(purple)



➔ Let's assume that if we don't pick **purple** we will pick orange; and vice versa

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

Developmental Computational Psychiatry lab

HARTLEY LAB

How should we choose?

Temperature (τ): parameter that determines the extent to which value estimates influence choice behaviour

Softmax equation:

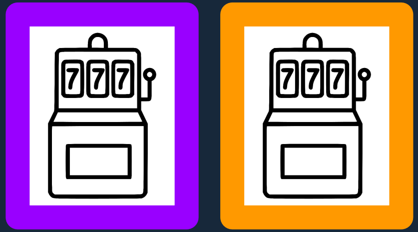
Value of machines

$$P(\text{purple}) = \frac{\exp([\,V_{purple}\,]/\,\tau)}{\text{SUM}[\exp([\,V_{purple}\ V_{orange}\,]/\,\tau)]}$$

Probability of choosing purple

➔ $P(\text{orange}) = 1 - P(\text{purple})$

Free parameter temperature



➔ Let's assume that if we don't pick purple we will pick orange; and vice versa

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

How should we choose?

$$P(\text{purple}) = \frac{\exp([V_{purple}]/\tau)}{\text{SUM}[\exp([V_{purple}, V_{orange}]/\tau)]}$$

Exploit

➔ Choose slot machine when reward is better than the other

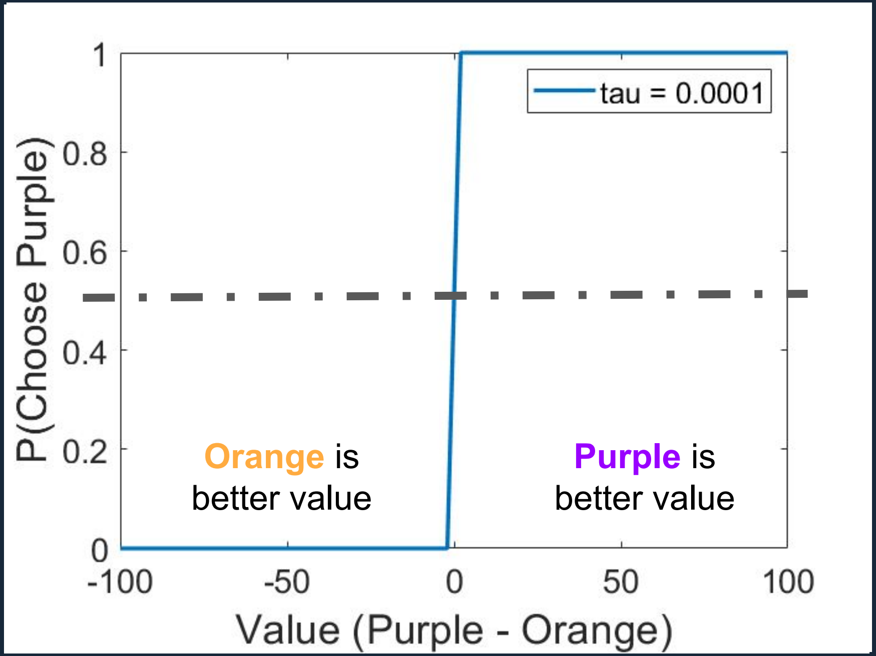
Developmental Computational Psychiatry lab

HARTLEY LAB

How should we choose?

Choose **purple** is better

Choose **orange** is better



$$P(\text{purple}) = \frac{\exp([V_{purple}]/\tau)}{\text{SUM}[\exp([V_{purple}, V_{orange}]/\tau)]}$$

Exploit

➔ Choose slot machine when reward is better than the other

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

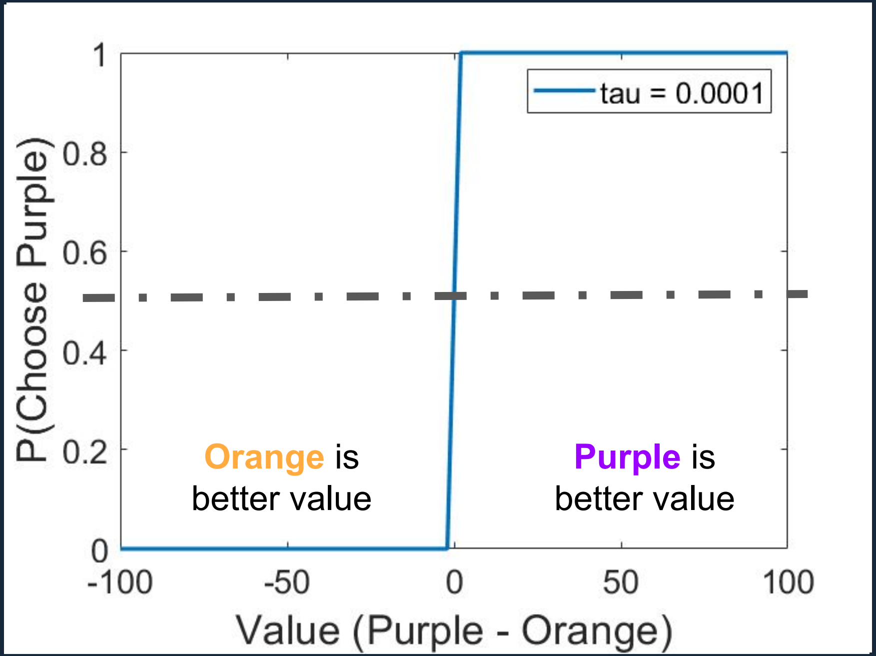
Developmental Computational Psychiatry lab

HARTLEY LAB

How should we choose?

Choose **purple** is better

Choose **orange** is better



$$P(\text{purple}) = \frac{\exp([V_{purple}]/\tau)}{\text{SUM}[\exp([V_{purple}, V_{orange}]/\tau)]}$$

Exploit

- ➔ Choose slot machine when reward is better than the other

Temperature is low

- ➔ Choices are less noisy
- ➔ More affected by value
- ➔ More deterministic

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

Developmental Computational Psychiatry lab

HARTLEY LAB

How should we choose?

$$P(\text{purple}) = \frac{\exp([\,V_{purple}\,]/\tau)}{\mathrm{SUM}[\exp([\,V_{purple}, V_{orange}\,]/\tau)]}$$

Explore

➔ Random choice

flux

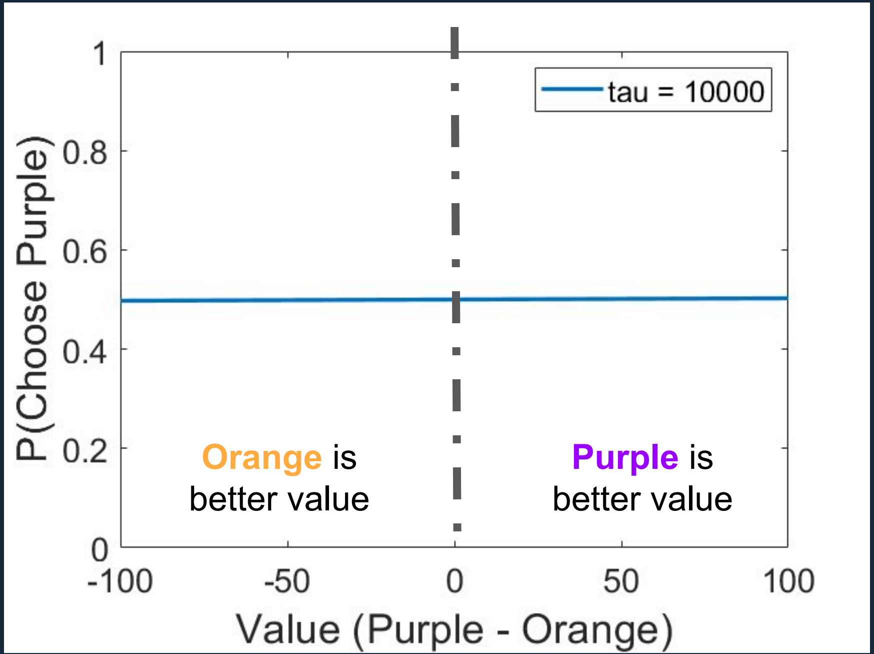
Developmental Computational Psychiatry lab

HARTLEY LAB

How should we choose?

Choose purple is better

Choose orange is better



$$P(\text{purple}) = \frac{\exp([V_{\text{purple}}]/\tau)}{\mathrm{SUM}[\exp([V_{\text{purple}}, V_{\text{orange}}]/\tau)]}$$

Explore

➔ Random choice

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

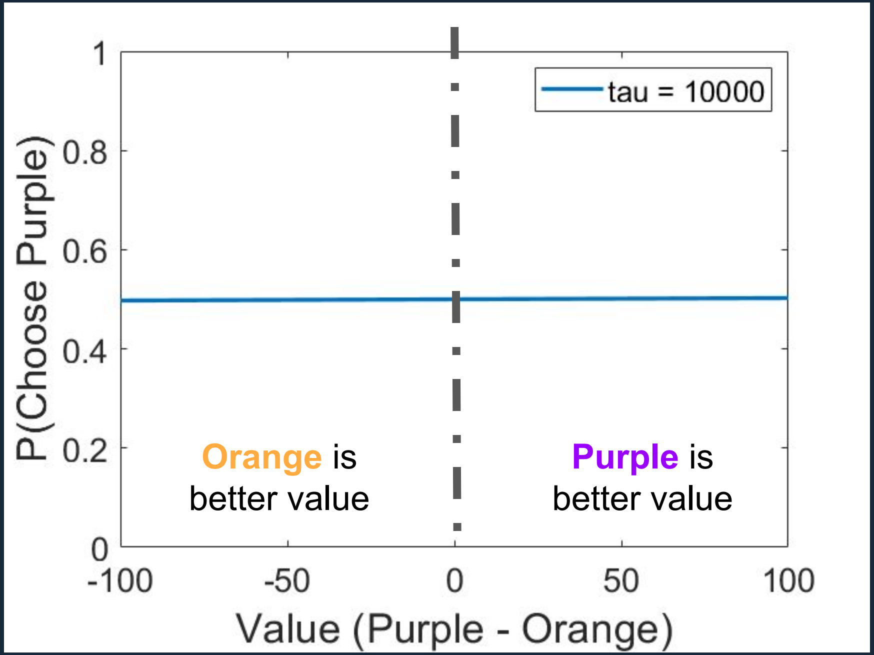
Developmental Computational Psychiatry lab

HARTLEY LAB

How should we choose?

Choose **purple** is better

Choose **orange** is better



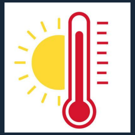
$$P(\text{purple}) = \frac{\exp([V_{purple}]/\tau)}{\text{SUM}[\exp([V_{purple}, V_{orange}]/\tau)]}$$

Explore

- ➔ Random choice

Temperature is high

- ➔ Choices are more noisy
- ➔ Less affected by value
- ➔ More random



Tricia Seow | Samuel Hewitt | Noam Goldway

flux

Developmental Computational Psychiatry lab

HARTLEY LAB

How should we choose?

$$P(\text{purple}) = \frac{\exp([V_{purple}]/\tau)}{\text{SUM}[\exp([V_{purple}, V_{orange}]/\tau)]}$$

Temperature is low

- ➔ Choices are less noisy
- ➔ More affected by value
- ➔ More deterministic

Temperature is high

- ➔ Choices are more noisy
- ➔ Less affected by value
- ➔ Less deterministic

Tricia Seow | Samuel Hewitt | Noam Goldway

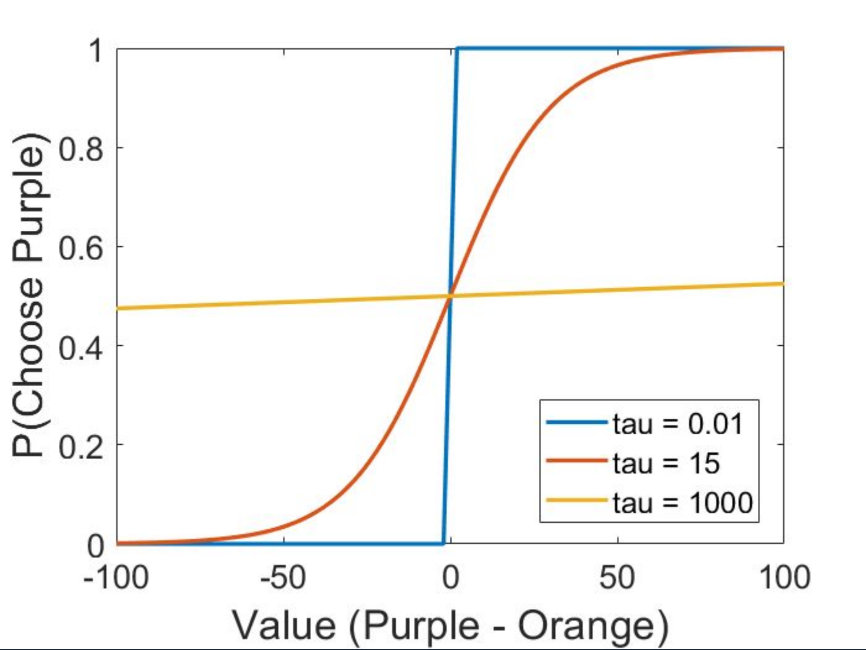
Developmental Computational Psychiatry lab

HARTLEY LAB

How should we choose?

Choose **purple** is better

Choose **orange** is better



$$P(\text{purple}) = \frac{\exp([\,V_{purple}\,]/\tau)}{\text{SUM}[\exp([\,V_{purple}, V_{orange}\,]/\tau)]}$$

Temperature is low

- ➔ Choices are less noisy
- ➔ More affected by value
- ➔ More deterministic

Temperature is high

- ➔ Choices are more noisy
- ➔ Less affected by value
- ➔ Less deterministic

Tricia Seow | Samuel Hewitt | Noam Goldway

Developmental Computational Psychiatry lab

HARTLEY LAB

How should we choose?

Softmax

$$P(\text{purple}) = \frac{\exp([\,V_{purple}\,]/\tau)}{\mathrm{SUM}[\exp([\,V_{purple}, V_{orange}\,]/\tau)]}$$

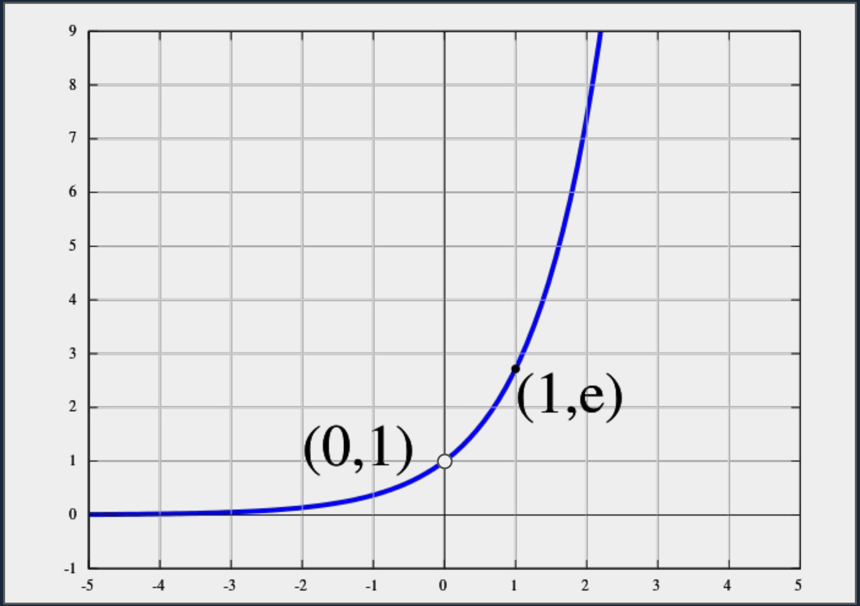
What does the **exponential (exp)** do?

Tricia Seow | Samuel Hewitt | Noam Goldway

Developmental Computational Psychiatry lab

HARTLEY LAB

How should we choose?

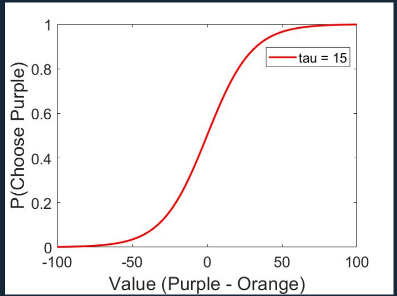


Softmax

$$P(\text{purple}) = \frac{\exp([\,V_{\text{purple}}\,]/\,\tau)}{\text{SUM}[\exp([\,V_{\text{purple}}, V_{\text{orange}}\,]/\,\tau)]}$$

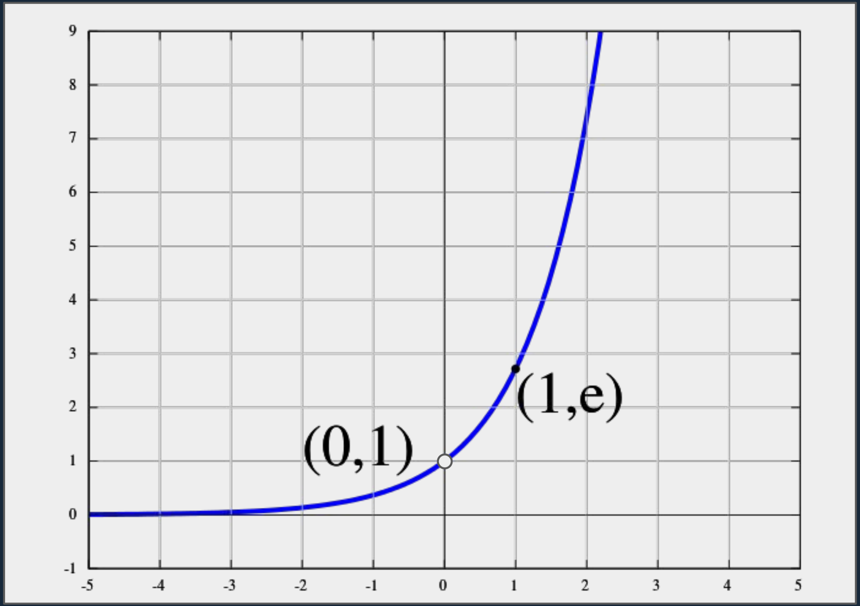
What does the **exponential** (exp) do?

- ➔ Non-linear transformation of value
- ➔ Deals with negative values



Tricia Seow | Samuel Hewitt | Noam Goldway

How should we choose?



Softmax

$$P(\text{purple}) = \frac{\exp([\ V_{purple}\ ]/\ \tau)}{\text{SUM}[\exp([\ V_{purple}, V_{orange}\ ]/\ \tau)]}$$

What does the **exponential (exp)** do?

➔ Non-linear transformation of value

➔ Deals with negative values

What does the **division by SUM** do?

➔ Normalizes values to between **0 to 1**

Developmental
Computational
Psychiatry lab

HARTLEY LAB

How should we choose?

$$P(\text{purple}) = \frac{\exp([\ V_{purple}\ ])}{\text{SUM}[\exp([\begin{smallmatrix} V_{purple} \\ V_{orange} \end{smallmatrix}])]}$$

Softmax

➔ Transforms value input into values between 0 to 1

Assume temperature = 1

Tricia Seow | Samuel Hewitt | Noam Goldway

Developmental Computational Psychiatry lab

HARTLEY LAB

How should we choose?

$$P(\text{purple}) = \frac{\exp([\,V_{purple}\,])}{\text{SUM}[\exp([\,V_{purple}\;V_{orange}\,])]}$$

Softmax

➔ Transforms value input into values between 0 to 1

Assume temperature = 1

For my next slot machine play...

V_{purple} = [60] V_{orange} = [40]

Tricia Seow | Samuel Hewitt | Noam Goldway

Developmental Computational Psychiatry lab

HARTLEY LAB

How should we choose?

$$P(\text{purple}) = \frac{\exp([\ V_{purple}\ ])}{\text{SUM}[\exp([\begin{smallmatrix} V_{purple} \\ V_{orange} \end{smallmatrix}])]}$$

$$= \frac{\exp([\ 60\ ])}{\text{SUM}[\exp([\begin{smallmatrix} 60 \\ 40 \end{smallmatrix}])]}$$

Softmax

➔ Transforms value input into values between 0 to 1

Assume temperature = 1

For my next slot machine play...

V_{purple} = [60] V_{orange} = [40]

Developmental Computational Psychiatry lab

HARTLEY LAB

How should we choose?

$$P(\text{purple}) = \frac{\exp([\ V_{\text{purple}}\ ])}{\text{SUM}[\exp(\begin{bmatrix} V_{\text{purple}} \\ V_{\text{orange}} \end{bmatrix})]}$$

$$= \frac{\exp([\ 60\ ])}{\text{SUM}[\exp(\begin{bmatrix} 60 \\ 40 \end{bmatrix})]}$$

$$= \frac{e^{60}}{e^{60}+e^{40}}$$

$$= 1$$

Softmax

→ Transforms value input into values between 0 to 1

Assume temperature = 1

For my next slot machine play...

$V_{\text{purple}} = [\ 60\ ]$ $V_{\text{orange}} = [\ 40\ ]$

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

Developmental Computational Psychiatry lab

HARTLEY LAB

How should we choose?

$$P(\text{purple}) = \frac{\exp([\,V_{purple}\,])}{SUM[\exp(\begin{bmatrix} V_{purple} \\ V_{orange} \end{bmatrix})]}$$

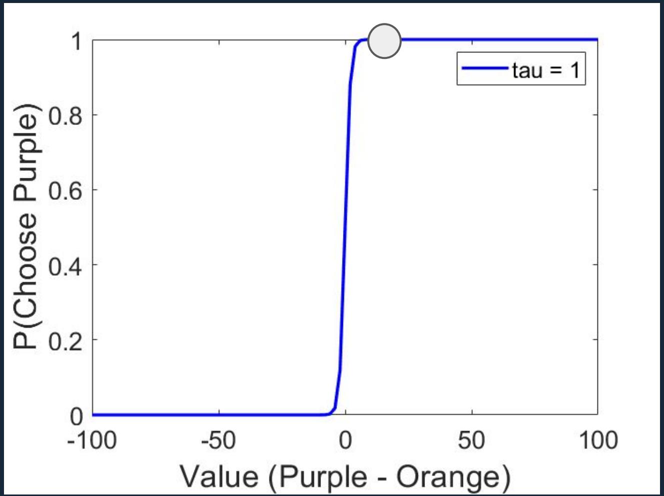
$$= \frac{\exp([\,60\,])}{SUM[\exp(\begin{bmatrix} 60 \\ 40 \end{bmatrix})]}$$

$$= \frac{e^{60}}{e^{60}+e^{40}}$$

$$= 1$$

Softmax

➔ Transforms value input into values between 0 to 1



$V_{purple} = [\,60\,]$ $V_{orange} = [\,40\,]$

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

How should we choose?

$$P(\text{purple}) = \frac{\exp([\ V_{purple}\ ])}{\text{SUM}[\exp([\ V_{purple}\ ;\ V_{orange}\ ])]}$$

$$= \frac{\exp([\ 60\ ])}{\text{SUM}[\exp([\ 60\ ;\ 40\ ])]}$$

$$= \frac{e^{60}}{e^{60}+e^{40}}$$

$$= 1$$

Softmax

- ➔ Transforms value input into values between 0 to 1

Assume temperature = 1

$$P(\text{orange}) = \frac{\exp([\ V_{orange}\ ])}{\text{SUM}[\exp([\ V_{purple}\ ;\ V_{orange}\ ])]}$$

$V_{purple} = [\ 60\ ]$ $V_{orange} = [\ 40\ ]$

Tricia Seow | Samuel Hewitt | Noam Goldway

How should we choose?

$$P(\text{purple}) = \frac{\exp([\,V_{purple}\,])}{\text{SUM}[\exp([\begin{smallmatrix}V_{purple}\\ V_{orange}\end{smallmatrix}])]} = \frac{\exp([\,60\,])}{\text{SUM}[\exp([\begin{smallmatrix}60\\ 40\end{smallmatrix}])]} = \frac{e^{60}}{e^{60}+e^{40}} = 1$$

Softmax

→ Transforms value input into values between 0 to 1

Assume temperature = 1

$$P(\text{orange}) = \frac{\exp([\,V_{orange}\,])}{\text{SUM}[\exp([\begin{smallmatrix}V_{purple}\\ V_{orange}\end{smallmatrix}])]} = \frac{e^{40}}{e^{60}+e^{40}} = 0$$

$V_{purple} = [\,60\,]$ $V_{orange} = [\,40\,]$

Developmental Computational Psychiatry lab

HARTLEY LAB

How should we choose?

$$P(\text{purple}) = \frac{\exp([\ V_{purple}\ ])}{\text{SUM}[\exp([\begin{smallmatrix} V_{purple} \\ V_{orange} \end{smallmatrix}])]}$$

$$= \frac{\exp([\ 60\ ])}{\text{SUM}[\exp([\begin{smallmatrix} 60 \\ 40 \end{smallmatrix}])]}$$

$$= \frac{e^{60}}{e^{60}+e^{40}}$$

$$= 1$$

$$P(\text{orange}) = \frac{\exp([\ V_{orange}\ ])}{\text{SUM}[\exp([\begin{smallmatrix} V_{purple} \\ V_{orange} \end{smallmatrix}])]}$$

$$= \frac{e^{40}}{e^{60}+e^{40}}$$

$$= 0$$

probability equals to 1

Softmax

➔ Transforms value input into values between 0 to 1

Assume temperature = 1

$V_{purple} = [\ 60\ ]$ $V_{orange} = [\ 40\ ]$

Developmental Computational Psychiatry lab

HARTLEY LAB

How should we choose?

$$P(\text{purple}) = \frac{\exp([V_{purple}]/\tau)}{\text{SUM}[\exp([V_{purple}, V_{orange}]/\tau)]}$$

Temperature (τ)

→ how much value affects choices

Assume temperature = 15

V_{purple} = [60] V_{orange} = [40]

Tricia Seow | Samuel Hewitt | Noam Goldway

Developmental Computational Psychiatry lab

HARTLEY LAB

How should we choose?

$$P(\text{purple}) = \frac{\exp([V_{purple}]/\tau)}{\text{SUM}[\exp([V_{purple}; V_{orange}]/\tau)]}$$

$$= \frac{\exp([60]/15)}{\text{SUM}[\exp([60; 40]/15)]}$$

Temperature (τ)

→ how much value affects choices

Assume temperature = 15

V_{purple} = [60] V_{orange} = [40]

Tricia Seow | Samuel Hewitt | Noam Goldway

Developmental Computational Psychiatry lab

HARTLEY LAB

How should we choose?

$$P(\text{purple}) = \frac{\exp([V_{purple}]/\tau)}{\text{SUM}\left[\exp\left(\begin{bmatrix} V_{purple} \\ V_{orange} \end{bmatrix}/\tau\right)\right]}$$

$$= \frac{\exp([60]/15)}{\text{SUM}\left[\exp\left(\begin{bmatrix} 60 \\ 40 \end{bmatrix}/15\right)\right]}$$

$$= \frac{e^{60/15}}{e^{60/15}+e^{40/15}}$$

$$= 0.79$$

Temperature (τ)

→ how much value affects choices

Assume temperature = 15

$V_{purple} = [\,60\,]$ $V_{orange} = [\,40\,]$

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

Developmental Computational Psychiatry lab

HARTLEY LAB

How should we choose?

$$P(\text{purple}) = \frac{\exp([V_{purple}]/\tau)}{\text{SUM}[\exp([V_{purple}; V_{orange}]/\tau)]}$$

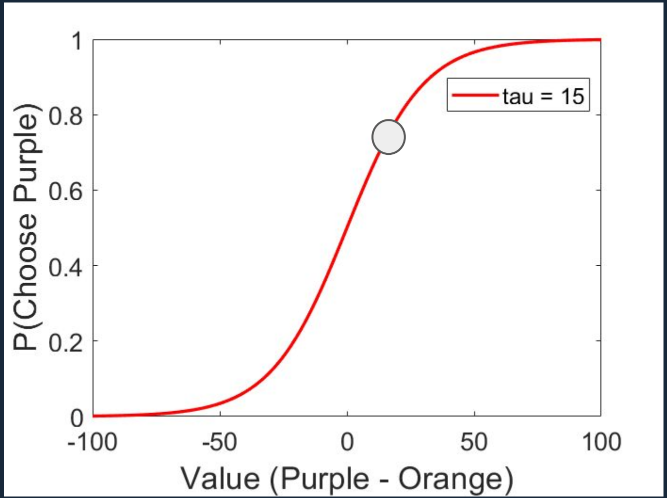
$$= \frac{\exp([60]/15)}{\text{SUM}[\exp([60; 40]/15)]}$$

$$= \frac{e^{60/15}}{e^{60/15}+e^{40/15}}$$

$$= 0.79$$

Temperature (τ)

→ how much value affects choices



$V_{purple} = [\,60\,]$ $V_{orange} = [\,40\,]$

Tricia Seow | Samuel Hewitt | Noam Goldway

Developmental Computational Psychiatry lab

HARTLEY LAB

How should we choose?

Temperature (τ)

→ how much value affects choices

Assume temperature = 15

$$P(\text{purple}) = \frac{\exp([V_{purple}]/\tau)}{\text{SUM}[\exp([V_{purple}, V_{orange}]/\tau)]} = \frac{\exp([60]/15)}{\text{SUM}[\exp([60, 40]/15)]} = \frac{e^{60/15}}{e^{60/15}+e^{40/15}} = 0.79$$

$$P(\text{orange}) = \frac{\exp([V_{orange}]/\tau)}{\text{SUM}[\exp([V_{purple}, V_{orange}]/\tau)]} = \frac{e^{40/15}}{e^{60/15}+e^{40/15}} = 0.21$$

probability equals to 1

$V_{purple} = [60]$ $V_{orange} = [40]$

Tricia Seow | Samuel Hewitt | Noam Goldway

flux

Developmental Computational Psychiatry lab

HARTLEY LAB

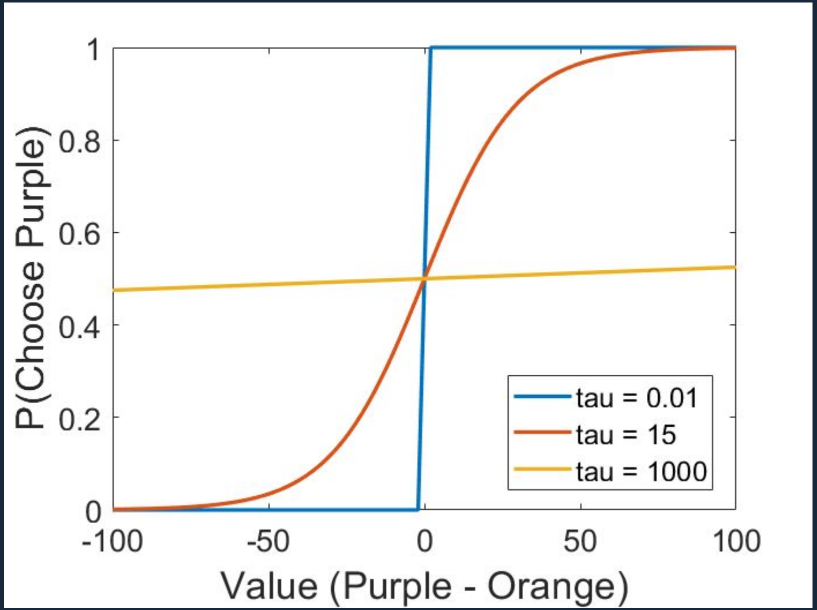
What have we learnt about choice?

Softmax equation:

$$P(\text{purple}) = \frac{\exp([V_{\text{purple}}]/\tau)}{\text{SUM}[\exp([V_{\text{purple}}, V_{\text{orange}}]/\tau)]}$$

Temperature (τ): parameter that determines the extent to which value estimates influence choice behaviour

Exploit or Explore



Tricia Seow | Samuel Hewitt | Noam Goldway

flux

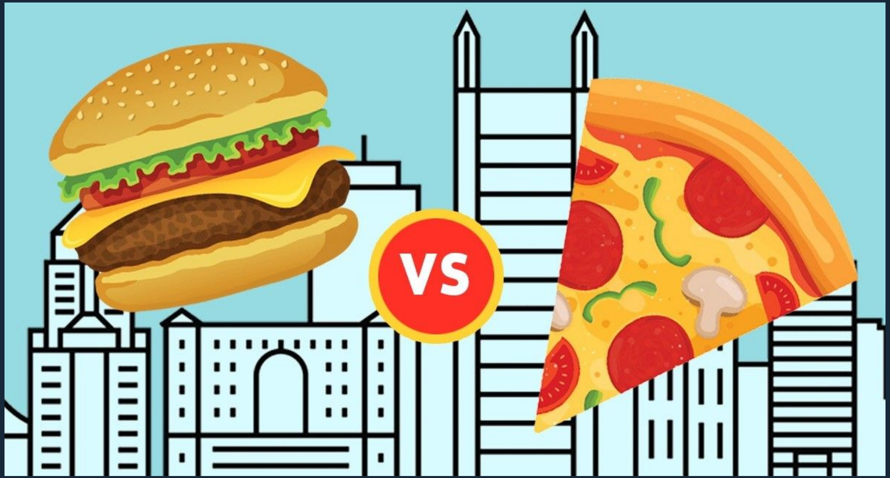
Developmental Computational Psychiatry lab

HARTLEY LAB

What have we learnt about choice?

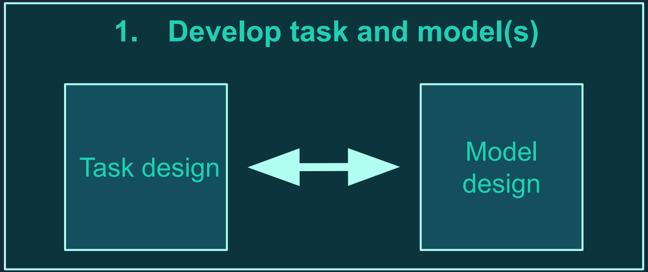
Exploit versus Explore:

- Discover "what works" by alternating between exploration and exploitation
- In uncertain environments, more exploration could be useful



flux

Summary



Task	**Trial and error learning**	**Reinforcement Learning**
Value Function	**Subjective value from objective outcomes**	**Rescorla-Wagner**
Choice Function	**Choice probabilities from value**	**Softmax**