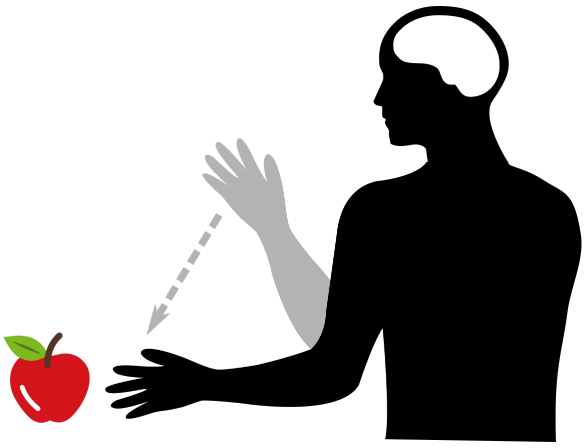
① "Mental space continuation" of reinforcement learning

In conventional RL, the **environment** is physically observable. I propose to extend it to the **internal** mental space.

From the traditional RL perspective: one reaches out for an apple, the apple is the **reward**, moving the arm is an **action**. These are all observable:

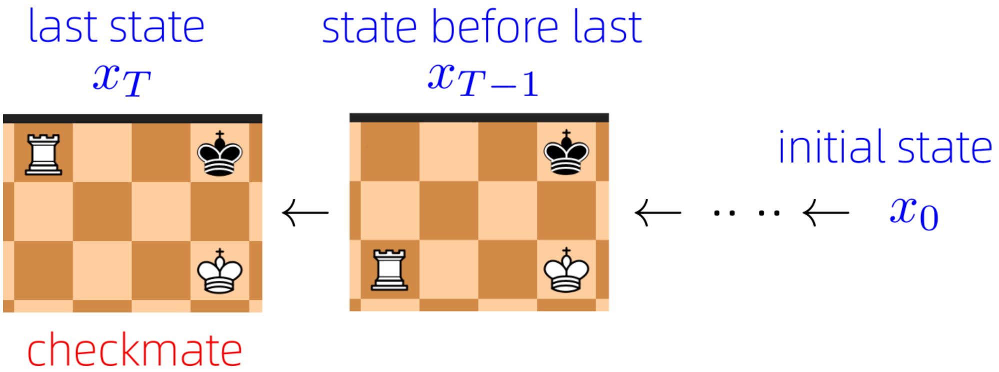


(1)

The foundation of RL is the **Bellman equation**. It can be viewed as a **recursive** formula:

$$\boxed{\text{Current state}} \quad V(x_0) = \max_a \{R + \gamma V(x_1)\} \quad \boxed{\text{Next state}} \tag{2}$$

It propagates the final state's reward back to the previous state, and the state before that... just like in a chess game... and so on until the very first move□

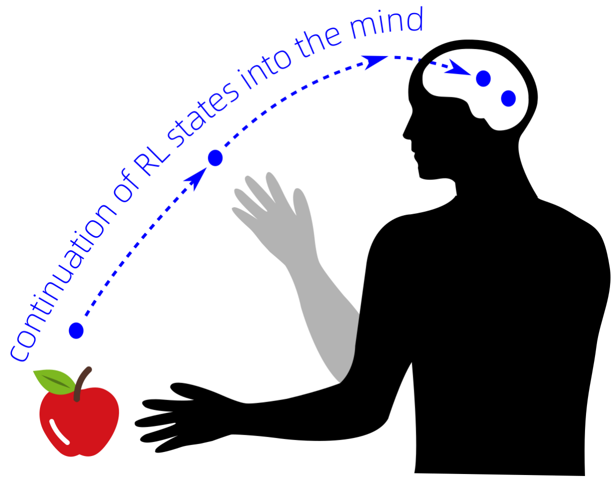


(3)

In other words, the reward of getting an apple, back-propagates to the action of reaching out an arm for the apple. So far so good. And we continue this process back to the **chain of thoughts** that decided to reach for an apple:

$$\text{□□□} \rightarrow \text{□□□□□} \rightarrow \text{□□□□□□□} \rightarrow \text{□□□□□} \rightarrow .... \tag{4}$$

In other words, we turn our **internal** mental states "inside-out", viewing them as **external** states, <u>on an equal footing</u>:



(5)

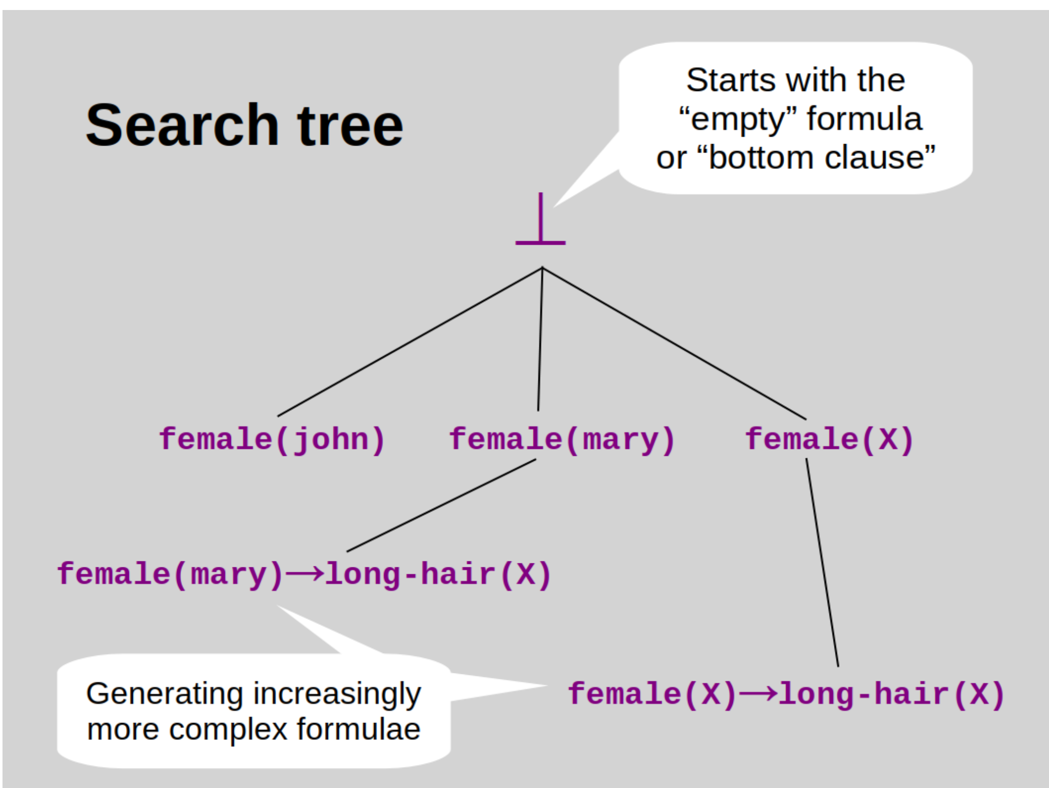
And this is <u>exactly analogous</u> to the propagation of rewards in a chess game. In other words, we can apply techniques of RL to learn the contents of mental space, thus providing a very rigorous foundation for AGI.

The above approach of unifying internal and external states is philosophically entirely sound, because our "brain states" are actually physical states; they are electrical activations of neuronal populations, and their transitions are governed by synaptic weights between neurons, which weights are updated through Hebbian learning, at least according to our current best understanding.

□ □□□□ □□□□□□ □□□□□ □□ □□□□□□□□□ □□□□□□ □□□□□□□□□□□ □□□□□ □□□ □□□□□□□□□□ □□ rule. □□□□□□ □□□□□ rules□□□□□□□□□ □□□ □□□□□□□□□□□□□□□□□ rules □□□□□□□□□□□□ rule(s). □□□□□□□→□□□□□□□□□□□□ rule□□□□□□→□□□□□□□□□□ 😂

□ □□□□ □□□□□ □□□□□□□□□□□□□□□ □□□□□□ rules□□□□□□□□□□□□□□□□□□□□□□□□□

□□□ AI □□□□□□ □□ rules □ combinatorial □□□□□□□□□□□□□□□□□□□□□□□□ RL □□□□□□□[2]□

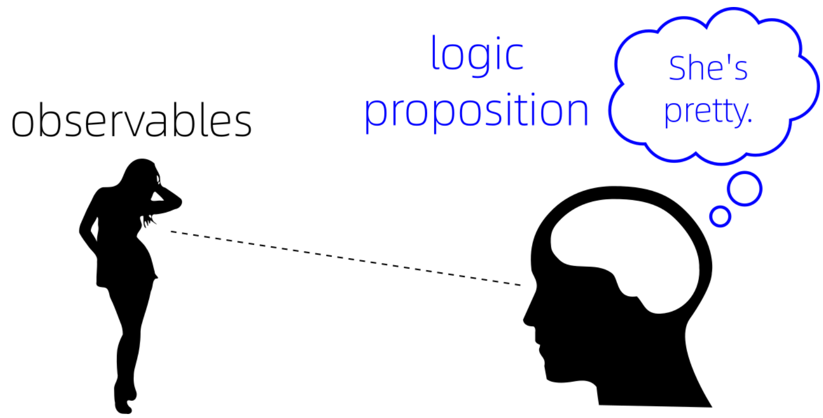


(6)

[2]□□□□□□□□□□□ RL □□□□□ □□□□□□□□□□□ □□□□□□ □□□□□/□□□

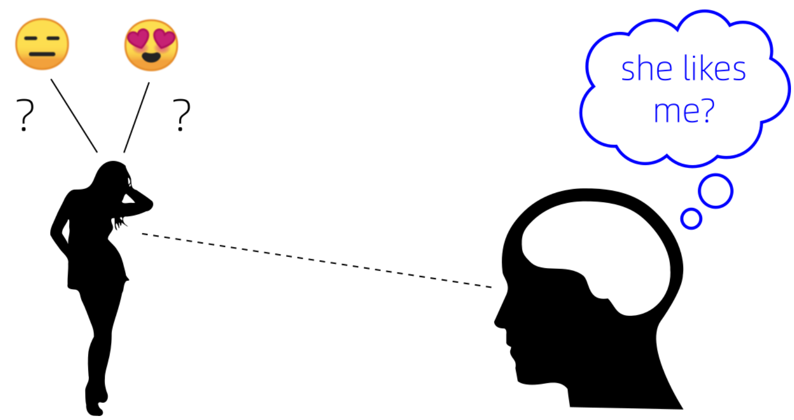
In traditional RL, there are **model-based** methods, and in our brain we build **mental models** of the world. These are actually one and the same concept, there is no conflict between the two.

□□□□□□□□□□□ (□□□) □□□□□
□□□□□□□□ □□□□□



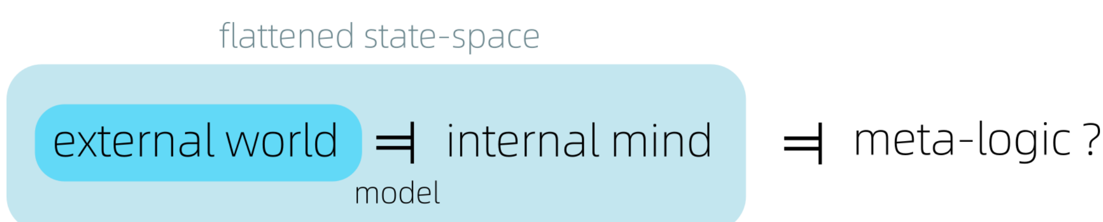
(7)

□□□□□□□□□□□□□□□□□□ "other minds" **□□□□**



(8)

□□□□□□□□□□□□□ flattened view **□□□□□□□□□ □ □□□□ □□□□□□□□□□□□□□ □□□□ □ □□□□ □** model **□** "theory":[3]



(9)

□□□□□□□□□ □□□□□□ theory **□□ □□□□□□ □□□□**

Picture credits:
Human figure from www.onlinewebfonts.com licensed by CC BY 3.0
Thought bubble created by Catherine Please from the Noun Project

[3] □□ $T \models M$ □□□□□ M is a model of T; T is a theory of M. □□ □□□□□ □□□□□□□□□□□