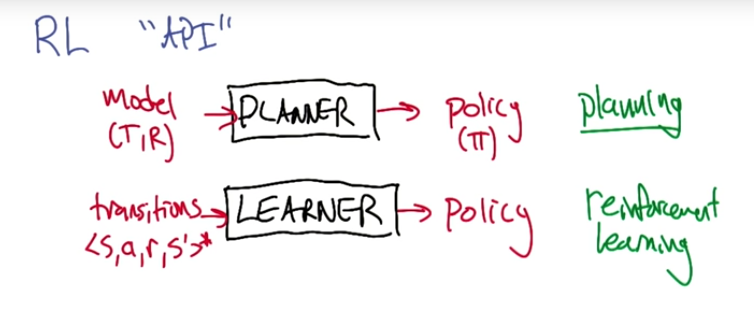
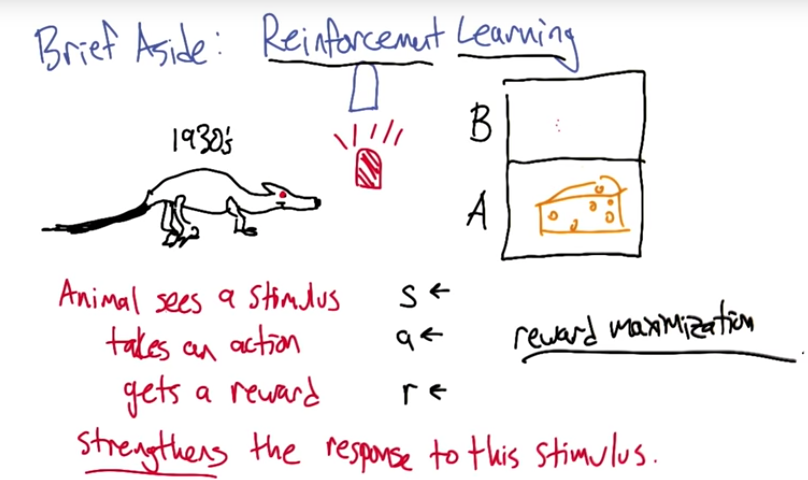
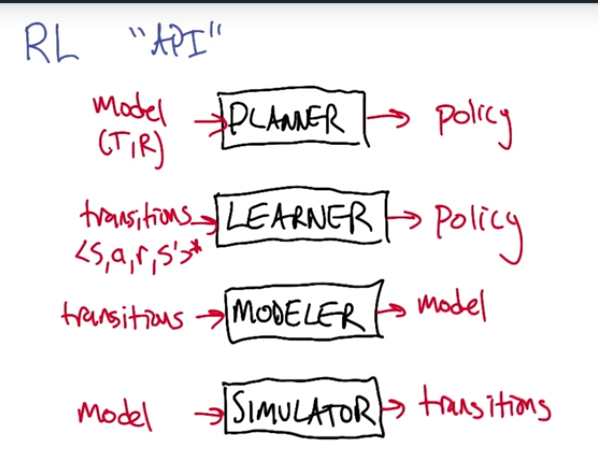
Reinforcement Learning

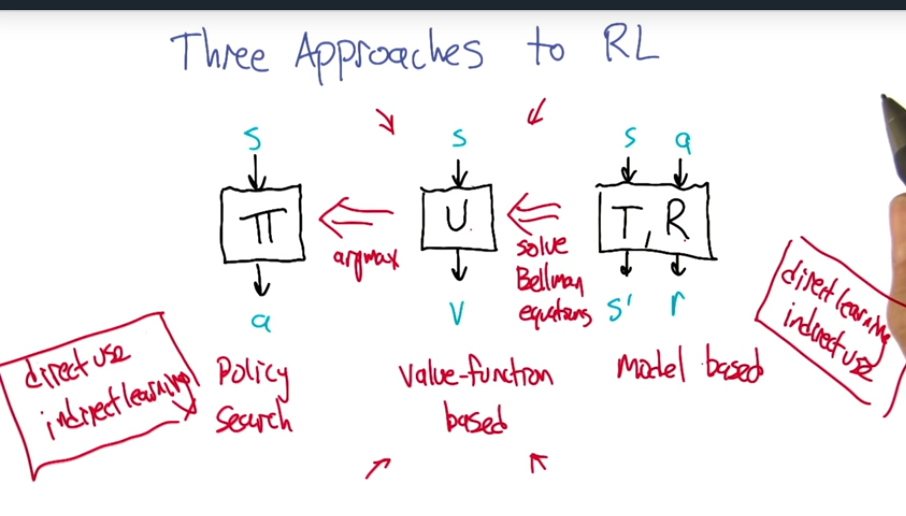


Reinforcement learning = reward maximization





* 3 approaches to RL
  + Policy search algorithm – maps states to actions
    - Pro: working on the data that you’ll be using
    - Con: function is indirect, data doesn’t tell you what action to choose
  + Value function based - Maps state to values
    - Pro: if we’re acting in the world, we can observe the values and the learning is not so indirect
    - Con: hard to use “U” directly, but can use *argmax*
  + Model based learning
    - Direct learning, can solve as a supervised learning problem
    - Indirect use



Q-learning

* Don’t have R(s) and T, so must figure out a different way to come up with optimal solution