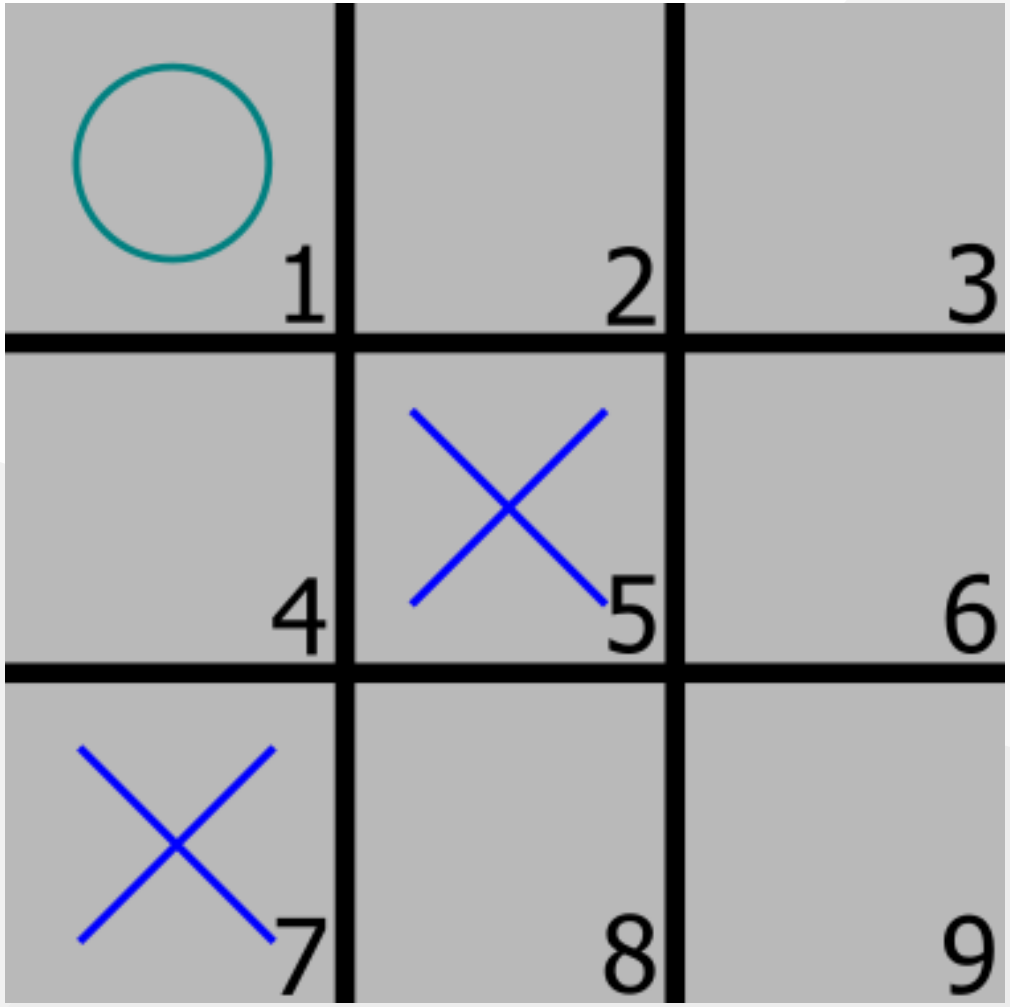
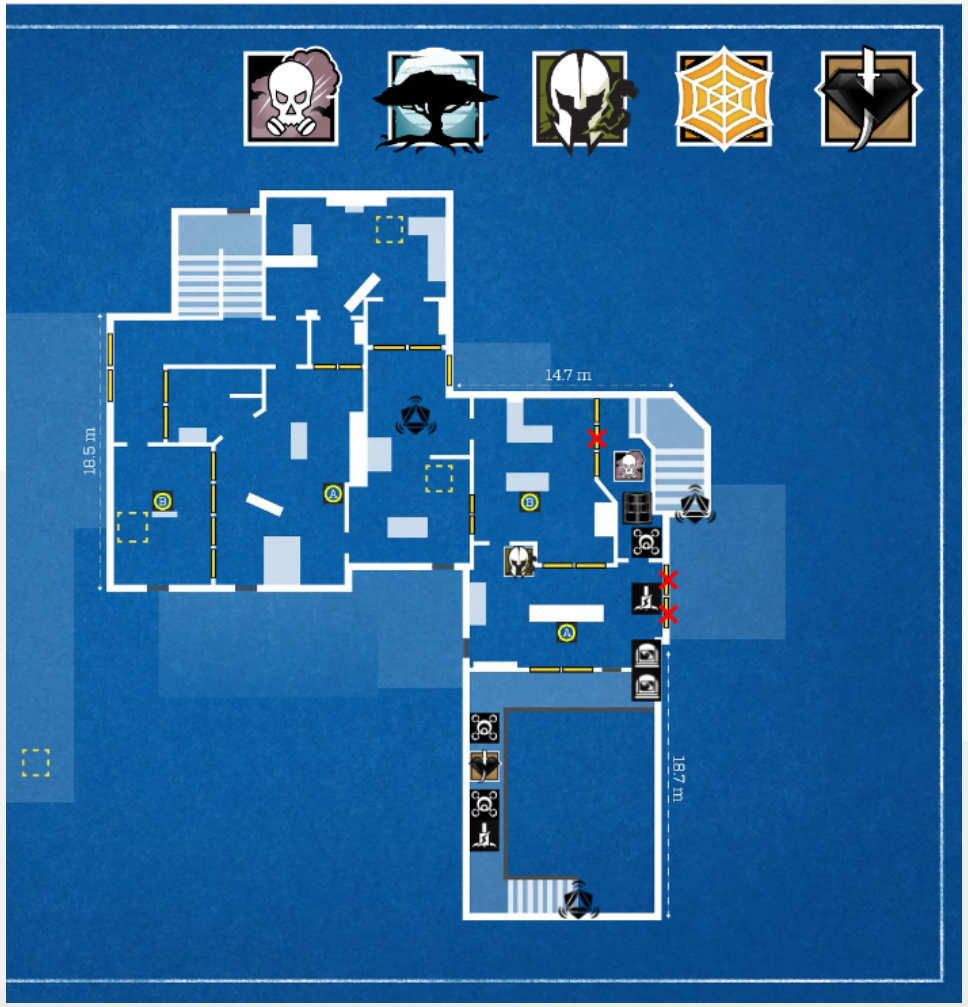
Week 1

Reinforcement learning is how we train AI to make decisions.

Tic Tac Toe



Rainbow Six: Siege



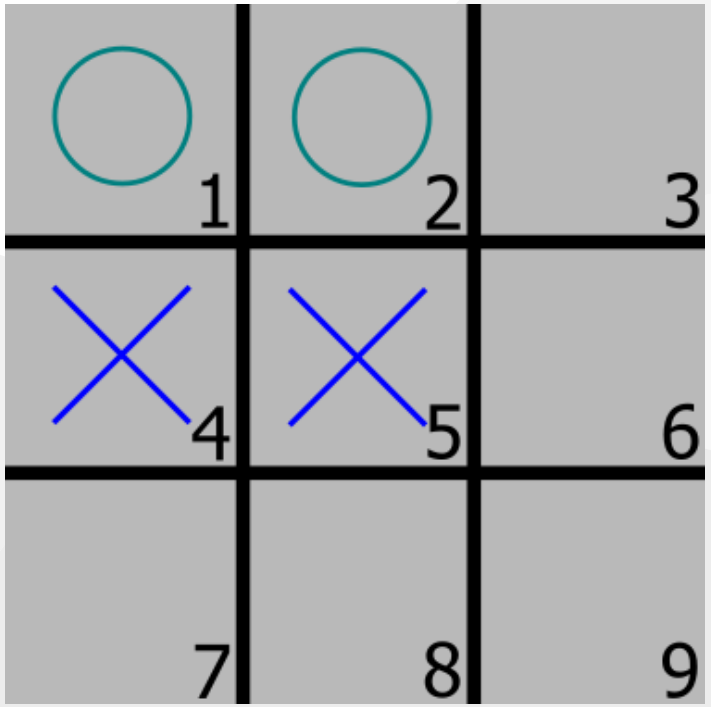
Policy

A policy comes from experience that allows you to make decisions.

You have a policy for tic tac toe because you have played many times before, but you also understand the game.

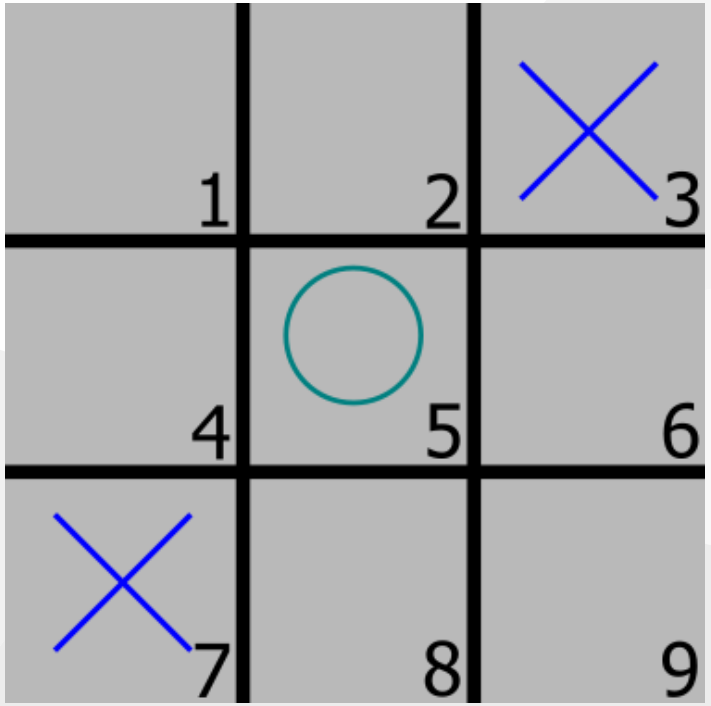
Understanding a Game

When you make a move in ttt, you automatically look ahead, because you understand the game. **Is 3 or 6 a better move for x?**



Another Example

Is 1 or 2 a better move for o?



Reward

By looking ahead, you calculate the move that will produce the most reward. By being rewarded, the AI will learn what moves are good and bad.

Reward can be positive or negative.

A move that makes you win sooner gives more reward (not applicable to ttt).

Information

What information is available (to the AI)?

State: what the board looks like

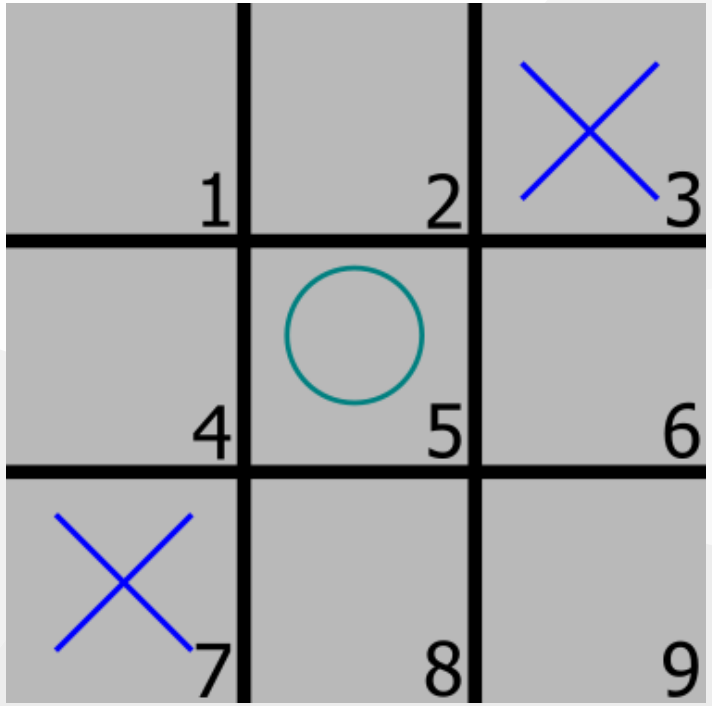
All possible actions: place a tile in squares 1-9

What can be derived from state and all actions?

All possible actions given the state.

Derive More Information

o can place a tile in 1, 2, 4, 6, 8, or 9



Why is this important?

Mao

A card game where you learn the rules by getting penalized.

You start out not knowing the rules and slowly learn.

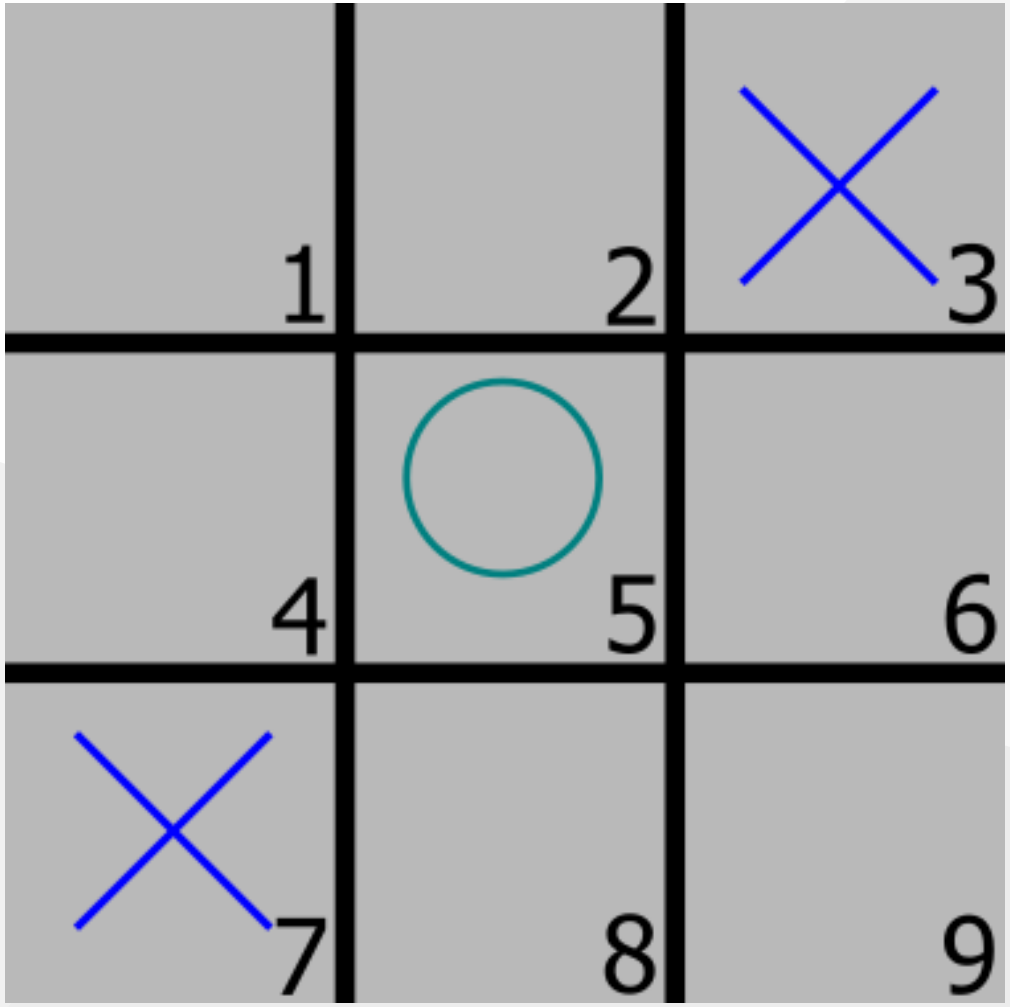
A lot harder to learn compared to if you knew the rules to begin with.

Remembering Moves

The AI needs to remember the moves it made to get to the current state because if it wins, it needs to know what moves it made to get there.

It will positively change the reward of those moves with more recent moves getting more rewar.

State



How the AI makes a decisions

```
pred = model.predict(...)
print(pred)
>>>
[
     .1, .9, .0,
     .9, .0, .9,
     .0, .9, .1
]
idx = np.argmax(pred)
print(idx)
>>> 1
```

State represented in code

```
state = [
    # x spaces
    0,0,1,
    0,0,0,
    1,0,0,
    # o spaces
    0,0,0,
    0,1,0,
    0,0,0,
    # legal moves
    1,1,0,
    1,0,1,
    0,1,1
]
```

Snake state

What state is available?

- Snake position
- Body position
- Apple position
- Current direction

Actions

What actions are available?

- Move up
- Move down
- Move left
- Move right

or

- turn left
- turn right
- continue straight