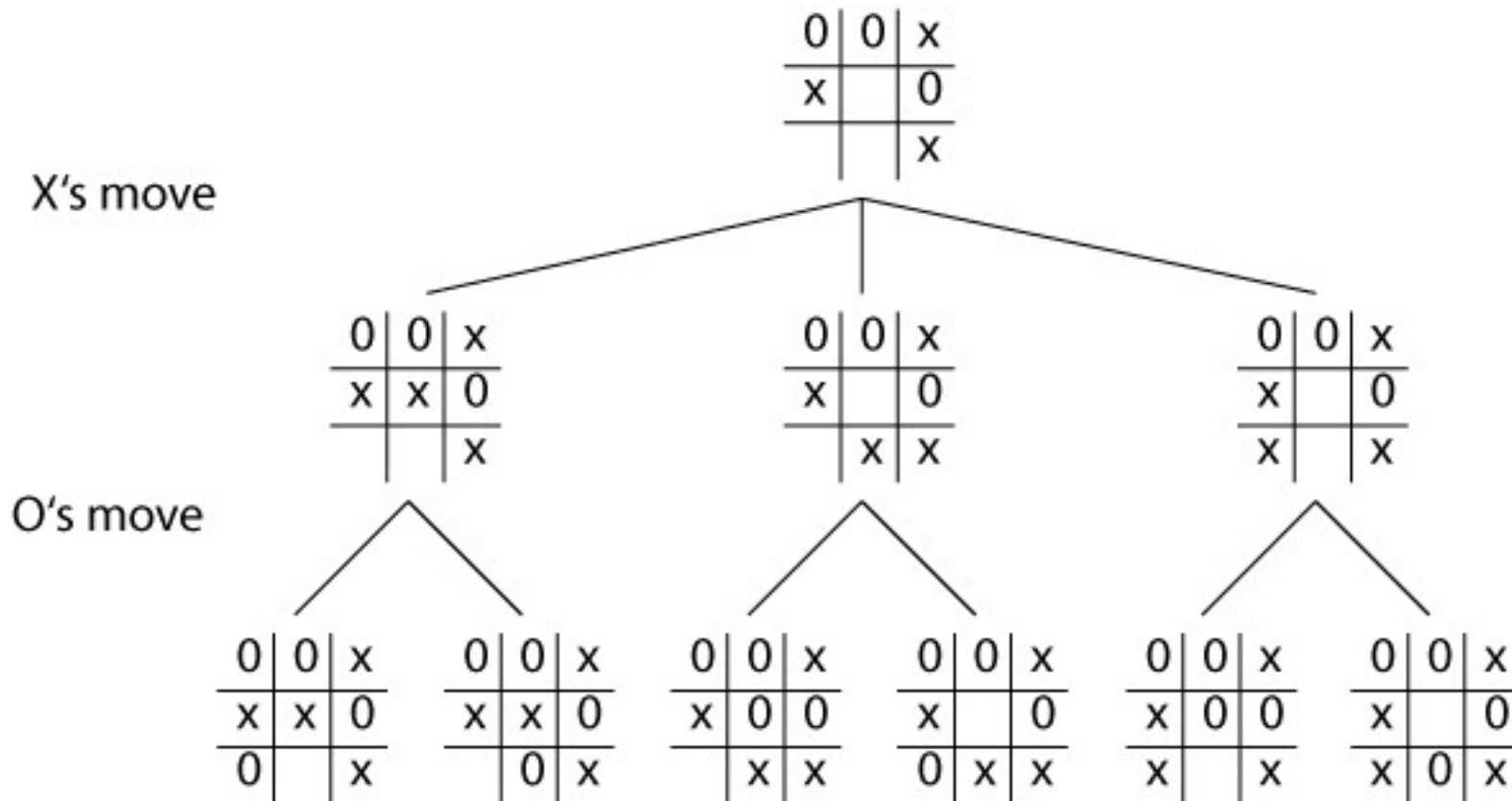# Reinforcement Learning (INF11010)

Lecture 15b: Minimax Search
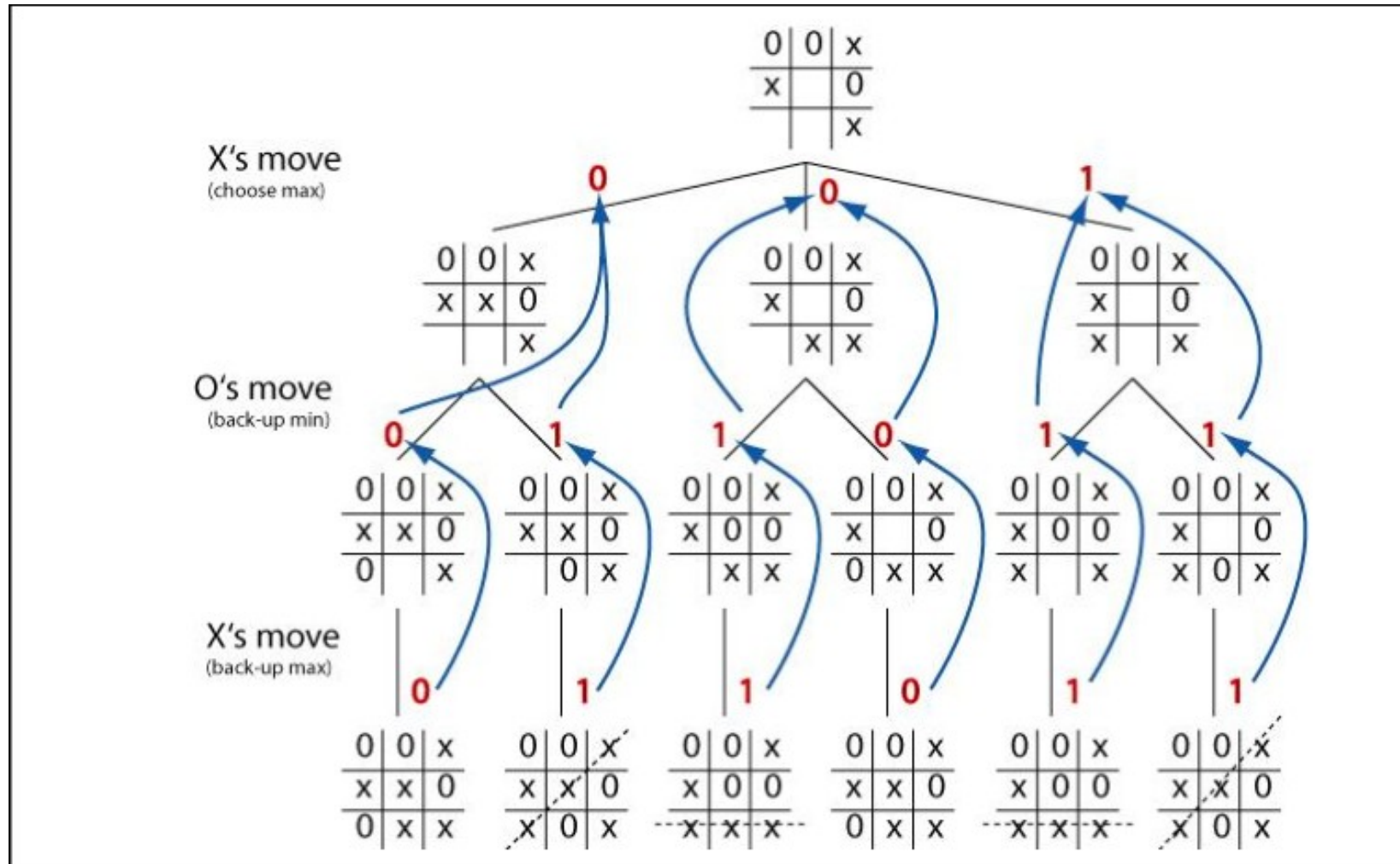
Pavlos Andreadis, April 3rd 2018

Minimax Search

- Shannon, Claude E. (1950) "XXII. Programming a computer for playing chess."

Minimax Search

Combining Minimax search with RL

- If the tree becomes intractable, we can stop branching out early, and using some heuristic evaluation of the leaf nodes, compute an optimal policy up till then.

- There are good examples of applications that did well using only minimax search and handcrafted features and weights for a linear approximation of the state value (no RL methods).

- For example *Deep Blue [Newborn, Monty. (2000) "Deep Blue's contribution to AI."]*.

Combining Minimax search with RL

- The evaluation at the leaves can also be computed by using Monte Carlo or TD methods over simulations, typically with function approximation.

- There exist different methods for updating the value function, by propagating the error to the root or leaf nodes of the branched out trees. (e.g. TD leaf, TD root).

- Such problems can also be solved by ignoring the Minimax solution, and just applying, e.g., TD with the other agent(s) as part of the environment, or by *self-play.*

- *For example, TD-Gammon [Tesauro, Gerald (1995) "Temporal difference learning and TD-Gammon."].*

- Alternatively, we can compute the value function with RL offline, and then use minimax search online, using the evaluations at the root nodes.