

CS 5/7320
Artificial Intelligence

Adversarial Search and Games

AIMA Chapter 5

Slides by Michael Hahsler
with figures from the AIMA textbook



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).

"Reflected Chess pieces" by Adrian Askew

A stack of three white rings with black centers, and a black ring with a white center, on a light gray surface. A black L-shaped block is also visible to the right.

Games

- Games typically confront the agent with a competitive (adversarial) environment affected by an opponent (strategic environment).
- We will focus on
 - two-player zero-sum games with
 - deterministic game mechanics and
 - perfect information (i.e., fully observable environment).
- We call the two players:
 - 1) **Max** tries to maximize his utility.
 - 2) **Min** tries to minimize Max's utility since it is a zero-sum game.



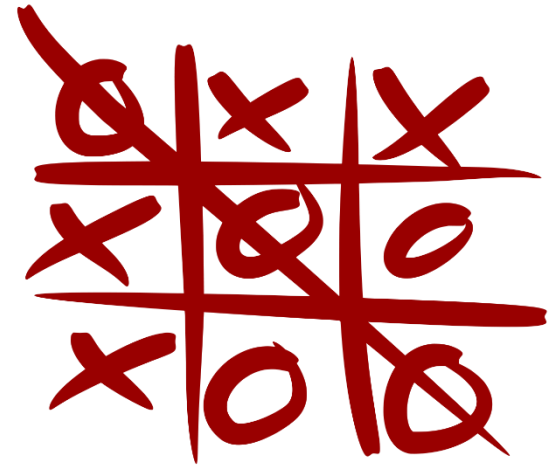
Definition of a Game

- **Definition:**

s_0	The initial state (position, board).
$Actions(s)$	Legal moves in state s .
$Result(s, a)$	Transition model.
$Terminal(s)$	Test for terminal states.
$Utility(s)$	Utility for player Max.

- **State space:** a graph defined by the initial state and the transition function containing all reachable states (e.g., chess positions).
- **Game tree:** a search tree superimposed on the state space. A complete game tree follows every sequence from the current state to the terminal state (the game ends).

Example: Tic-tac-toe



s_0

Empty board.

$Actions(s)$

Empty squares.

$Result(s, a)$

Place symbol (x/o) on empty square.

$Terminal(s)$

Did a player win or is the game a draw?

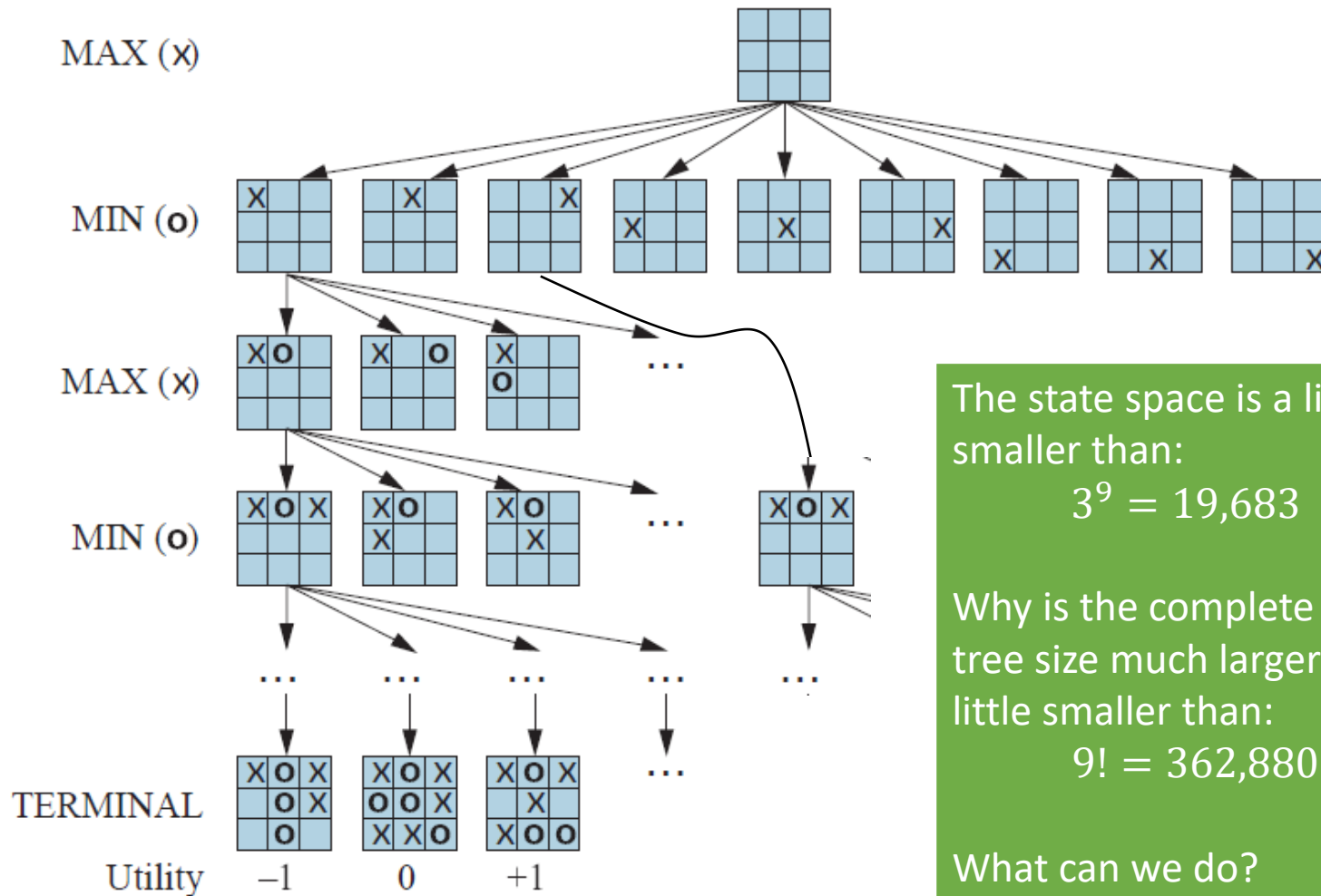
$Utility(s)$

+1 if x wins, -1 if o wins and 0 for a draw.

Utility is only defined for terminal states.

Here player x is Max
and player o is Min.

Tic-tac-toe: Partial Game Tree



The state space is a little smaller than:

$$3^9 = 19,683$$

Why is the complete game tree size much larger? A little smaller than:

$$9! = 362,880$$

What can we do?

Methods for Adversarial Games

Exact Methods

- **Model as nondeterministic actions:** The opponent is seen as part of an environment with nondeterministic actions. Non-determinism is the result of the unknown moves by the opponent. We **consider all possible moves** by the opponent.
- **Find optimal decisions:** Minimax search and Alpha-Beta pruning where **each player plays optimal** to the end of the game.

Heuristic Methods

(game tree is too large)

- **Heuristic Alpha-Beta Tree Search:**
 - a. Cut off game tree and use heuristic for utility.
 - b. Forward Pruning: ignore poor moves.
- **Monte Carlo Tree search:** Estimate utility of a state by simulating complete games and average the utility.

A dynamic background image showing a bright yellow powder or smoke explosion against a black background. The particles are concentrated on the right side and spread out towards the left, creating a sense of motion and energy.

Nondeterministic Actions

Recall AND-OR Search from AIMA Chapter 4

Methods for Adversarial Games

Exact Methods

- **Model as nondeterministic actions:** The opponent is seen as part of an environment with nondeterministic actions. Non-determinism is the result of the unknown moves by the opponent. We **consider all possible moves** by the opponent.
- **Find optimal decisions:** Minimax search and Alpha-Beta pruning where **each player plays optimal** to the end of the game.

Heuristic Methods

(game tree is too large)

- **Heuristic Alpha-Beta Tree Search:**
 - a. Cut off game tree and use heuristic for utility.
 - b. Forward Pruning: ignore poor moves.
- **Monte Carlo Tree search:** Estimate utility of a state by simulating complete games and average the utility.

Nondeterministic Actions

Each action consists of the move by the player and all possible (i.e., nondeterministic) responses by the opponent.

Outcome of actions in the environment is nondeterministic = **transition model need to describe uncertainty about the opponent's behavior.**

Example transition:

$$Results(s_1, a) = \{s_2, s_4, s_5\}$$

i.e., action a in s_1 can lead to one of several states.

AND-OR DFS Search Algorithm

= nested If-then-else statements

function AND-OR-SEARCH(*problem*) **returns** a conditional plan, or *failure*
return OR-SEARCH(*problem*, *problem*.INITIAL, [])

function OR-SEARCH(*problem*, *state*, *path*) **returns** a conditional plan, or *failure*
if *problem*.IS-GOAL(*state*) **then return** the empty plan
if IS-CYCLE(*path*) **then return failure** // don't follow loops
for each *action* **in** *problem*.ACTIONS(*state*) **do** // check all possible actions
 plan \leftarrow AND-SEARCH(*problem*, RESULTS(*state*, *action*), [*state*] + *path*)
 if *plan* \neq *failure* **then return** [*action*] + *plan*
return failure

my
moves

all states that can result from
opponent's moves

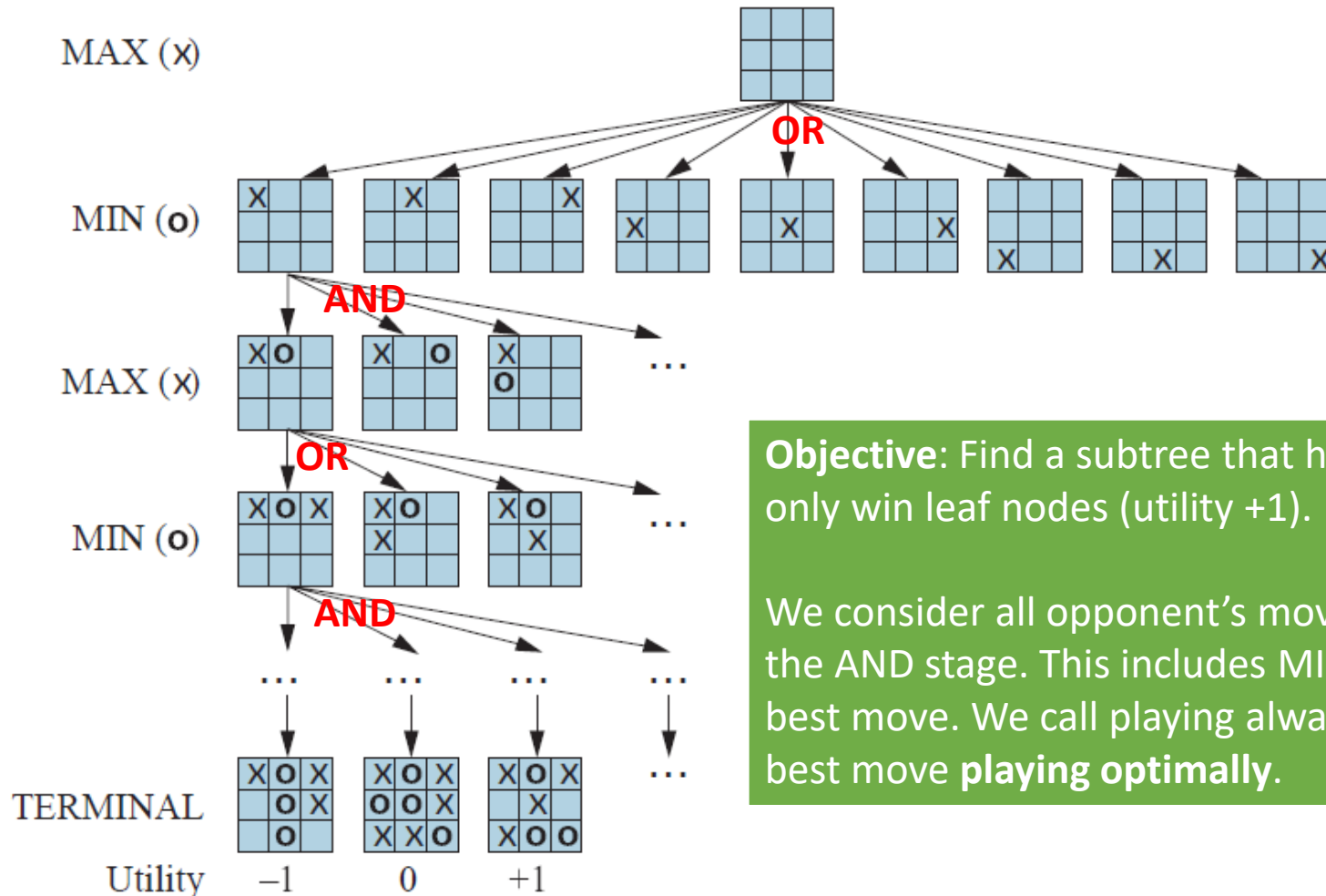
function AND-SEARCH(*problem*, *states*, *path*) **returns** a conditional plan, or *failure*
for each *s_i* **in** *states* **do** // check all possible current states
 plan_i \leftarrow OR-SEARCH(*problem*, *s_i*, *path*)
 if *plan_i* = *failure* **then return failure**
return [if *s₁* **then** *plan₁* **else if** *s₂* **then** *plan₂* **else** ... **if** *s_{n-1}* **then** *plan_{n-1}* **else** *plan_n*]

Go through
opponent
moves

- And-Or Search searches the whole tree till it finds a subtree that leads only to goal nodes.
- BFS and A* search can also be used to search an AND-OR tree.

Tic-tac-toe: AND-OR Search

We play MAX and decide on our actions (OR).
MIN's actions introduce non-determinism (AND).



Objective: Find a subtree that has only win leaf nodes (utility +1).

We consider all opponent's moves in the AND stage. This includes MIN's best move. We call playing always the best move **playing optimally**.



Optimal Decisions

Minimax Search and Alpha-Beta Pruning

Methods for Adversarial Games

Exact Methods

- **Model as nondeterministic actions:** The opponent is seen as part of an environment with nondeterministic actions. Non-determinism is the result of the unknown moves by the opponent. We **consider all possible moves** by the opponent.
- **Find optimal decisions:** Minimax search and Alpha-Beta pruning where **each player plays optimal** to the end of the game.

Heuristic Methods

(game tree is too large)

- **Heuristic Alpha-Beta Tree Search:**
 - a. Cut off game tree and use heuristic for utility.
 - b. Forward Pruning: ignore poor moves.
- **Monte Carlo Tree search:** Estimate utility of a state by simulating complete games and average the utility.

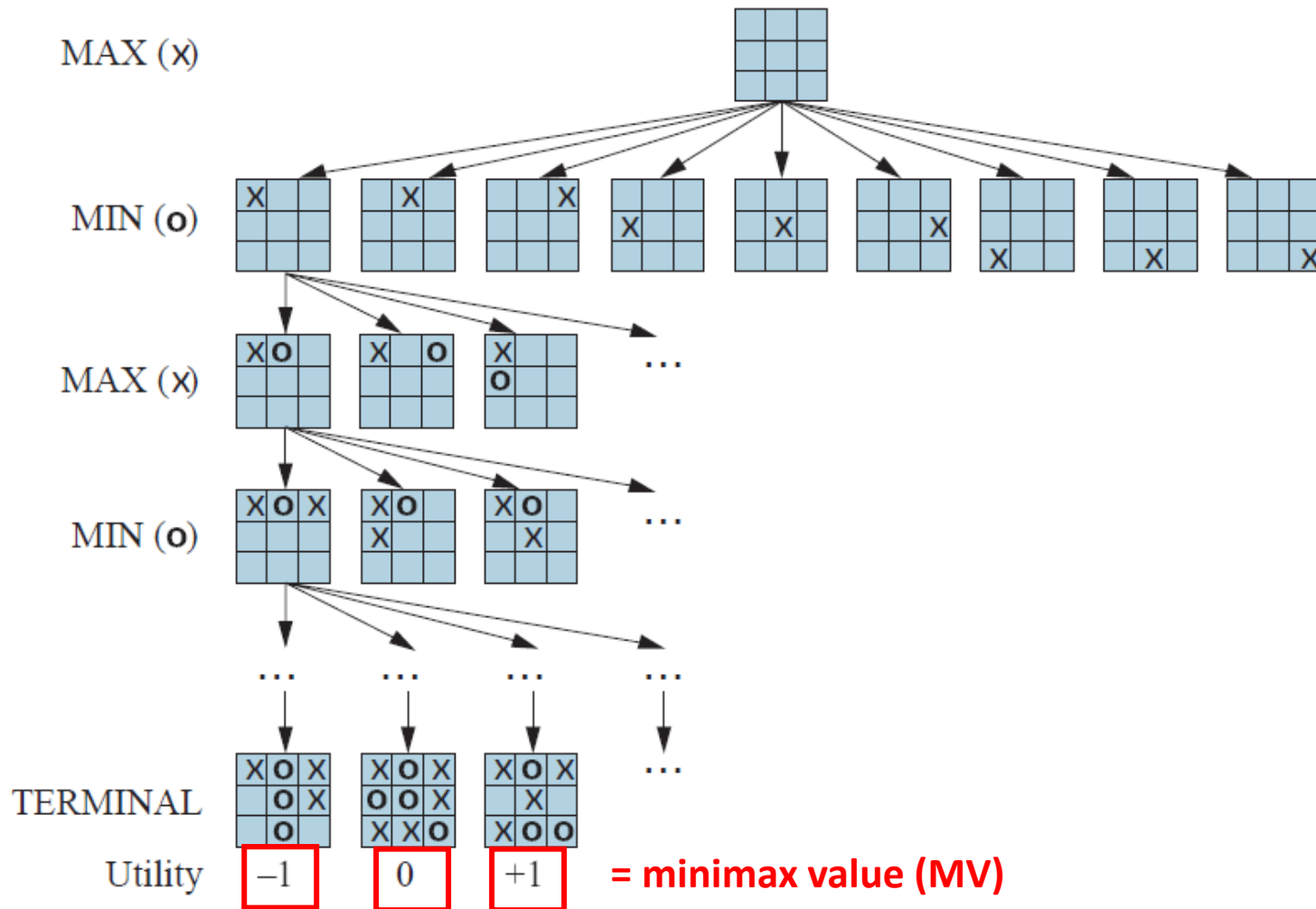
Idea: Minimax Decision

- Assign each state a **minimax value** that reflects how much Max prefers the state (= Min dislikes the state).

$$\text{Minimax}(s) = \begin{cases} \text{Utility}(s) & \text{if } \text{terminal}(s) \\ \max_{a \in \text{Actions}(s)} \text{Minimax}(\text{Result}(s, a)) & \text{if } \text{move} = \text{Max} \\ \min_{a \in \text{Actions}(s)} \text{Minimax}(\text{Result}(s, a)) & \text{if } \text{move} = \text{Min} \end{cases}$$

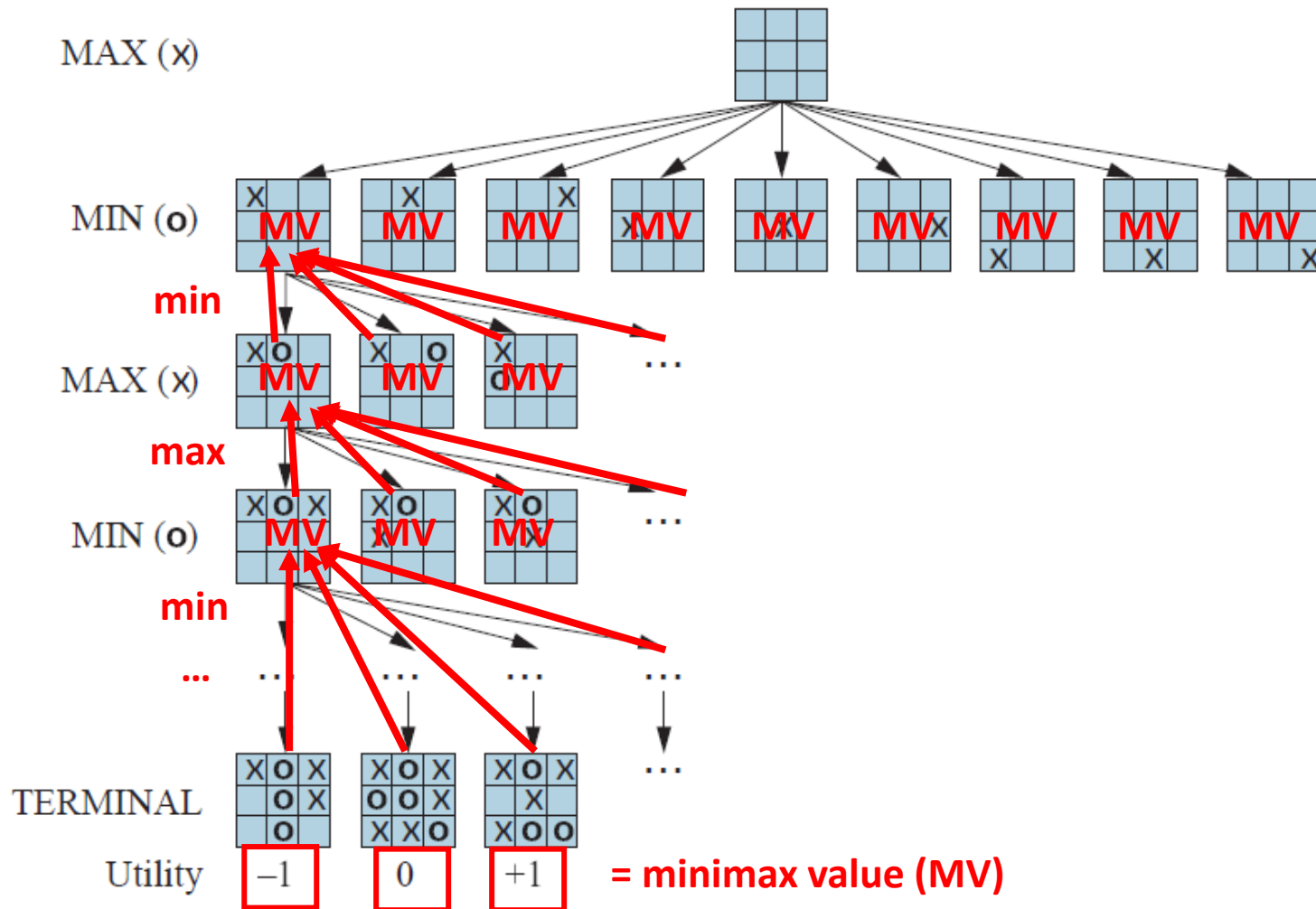
- The minimax value is the utility for Max in state s assuming that **both players play optimally** from s to the end of the game.
- The **optimal decision** for Max is the action that leads to the state with the largest minimax value.

Minimax Search

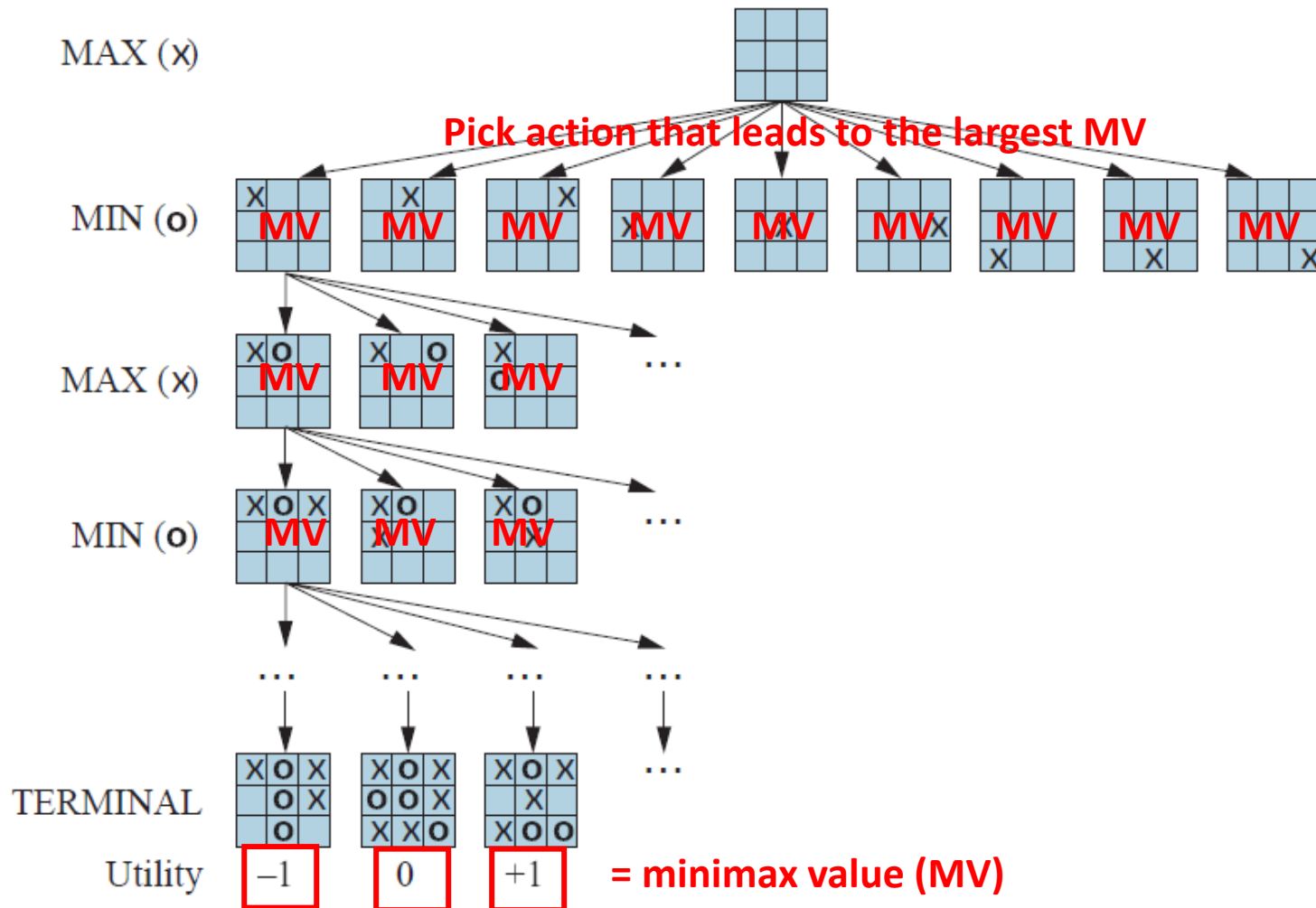


Minimax Search : Back-up

Minimax Values



Minimax Search: Decision



Approach: Follow tree to each terminal node and back up minimax value.

Note: This is just a generalization of the AND-OR Tree Search and returns first action of the conditional plan.

```
function MINIMAX-SEARCH(game, state) returns an action  
  player  $\leftarrow$  game.TO-MOVE(state)  
  value, move  $\leftarrow$  MAX-VALUE(game, state)  
  return move
```

```
function MAX-VALUE(game, state) returns a (utility, move) pair  
  if game.IS-TERMINAL(state) then return game.UTILITY(state, player), null  
  v  $\leftarrow -\infty$   
  for each a in game.ACTIONS(state) do  
    v2, a2  $\leftarrow$  MIN-VALUE(game, game.RESULT(state, a))  
    if v2 > v then  
      v, move  $\leftarrow$  v2, a  
  return v, move
```

Represents
OR Search

```
function MIN-VALUE(game, state) returns a (utility, move) pair  
  if game.IS-TERMINAL(state) then return game.UTILITY(state, player), null  
  v  $\leftarrow +\infty$   
  for each a in game.ACTIONS(state) do  
    v2, a2  $\leftarrow$  MAX-VALUE(game, game.RESULT(state, a))  
    if v2 < v then  
      v, move  $\leftarrow$  v2, a  
  return v, move
```

Represents
AND Search

b: branching factor
m: max depth of tree

Issue: Game Tree Size

- This traverses the complete game tree using DFS!

Time complexity: $O(b^m)$

- Only feasible for very simple games!
- Example: Tic-tac-toe
 $b = 9, m = 9 \rightarrow O(9^9) = O(387,420,489)$
 b decreases from 9 to 8, 7, ...
 \rightarrow we get less than $O(9!) = O(362,880)$
- We need to reduce the search space! \rightarrow **Game tree pruning**

Alpha-Beta Pruning

- **Idea:** Do not search parts of the tree if they do not make a difference to the outcome.
- **Observations:**
 - $\min(3, x, y)$ can never be more than 3
 - $\max(5, \min(3, x, y, \dots))$ does not depend on the values of x or y .
 - Minimax search applies alternating min and max.
- **Approach:** maintain for each node bounds for the minimax value $[\alpha, \beta]$ and prune subtrees that cannot be part of the solution.
 - Alpha is used by Max and means “ $\text{Minimax}(s)$ is at least α .”
 - Beta is used by Min and means “ $\text{Minimax}(s)$ is at most β .”

function ALPHA-BETA-SEARCH(*game*, *state*) **returns** an action
 player \leftarrow *game*.TO-MOVE(*state*)
value, *move* \leftarrow MAX-VALUE(*game*, *state*, $-\infty$, $+\infty$)
return *move*

= minimax search + pruning

function MAX-VALUE(*game*, *state*, α , β) **returns** a (*utility*, *move*) pair
if *game*.IS-TERMINAL(*state*) **then return** *game*.UTILITY(*state*, *player*), null
v $\leftarrow -\infty$
for each *a* **in** *game*.ACTIONS(*state*) **do**
 v2, *a2* \leftarrow MIN-VALUE(*game*, *game*.RESULT(*state*, *a*), α , β)
 if *v2* > *v* **then**
 v, *move* \leftarrow *v2*, *a*
 $\alpha \leftarrow$ MAX(α , *v*)
 if *v* \geq β **then return** *v*, *move*
return *v*, *move*

function MIN-VALUE(*game*, *state*, α , β) **returns** a (*utility*, *move*) pair
if *game*.IS-TERMINAL(*state*) **then return** *game*.UTILITY(*state*, *player*), null
v $\leftarrow +\infty$
for each *a* **in** *game*.ACTIONS(*state*) **do**
 v2, *a2* \leftarrow MAX-VALUE(*game*, *game*.RESULT(*state*, *a*), α , β)
 if *v2* < *v* **then**
 v, *move* \leftarrow *v2*, *a*
 $\beta \leftarrow$ MIN(β , *v*)
 if *v* \leq α **then return** *v*, *move*
return *v*, *move*

Notes:

- Pruning can be made more effective by **move ordering**: Check known good moves first to get a good bound early.
- Optimal decision algorithms still scale poorly!

A close-up photograph of numerous wooden Tetris blocks of various colors (purple, blue, green, orange, red, pink, grey, brown, yellow) scattered on a wooden surface. The blocks are in different orientations, some standing upright and some lying flat. The text "Heuristic Alpha-Beta Tree Search" is overlaid in the center in a white, sans-serif font.

Heuristic Alpha-Beta Tree Search

Methods for Adversarial Games

Exact Methods

- **Model as nondeterministic actions:** The opponent is seen as part of an environment with nondeterministic actions. Non-determinism is the result of the unknown moves by the opponent. We **consider all possible moves** by the opponent.
- **Find optimal decisions:** Minimax search and Alpha-Beta pruning where **each player plays optimal** to the end of the game.

Heuristic Methods

(game tree is too large)

- **Heuristic Alpha-Beta Tree Search:**
 - a. Cut off game tree and use heuristic for utility.
 - b. Forward Pruning: ignore poor moves.
- **Monte Carlo Tree search:** Estimate utility of a state by simulating complete games and average the utility.

Idea: Cutting off search

Stop search at a node before the terminal node is reached.
Use a heuristic evaluation function $Eval(s)$ to approximate the utility for that node/state.

Properties of the evaluation function:

- Fast to compute.
- $Eval(s) \in [Utility(loss), Utility(win)]$
- Correlated with the actual chance of winning (e.g., using features of the state).

Example: A weighted linear function

$$Eval(s) = w_1 f_1(s) + w_2 f_2(s) + \cdots + w_n f_n(s)$$

where f_i is a feature of the state (e.g., # of pieces captured in chess).

Heuristic Alpha-Beta Tree Search: Cutting off search

depth

0 MAX (x)

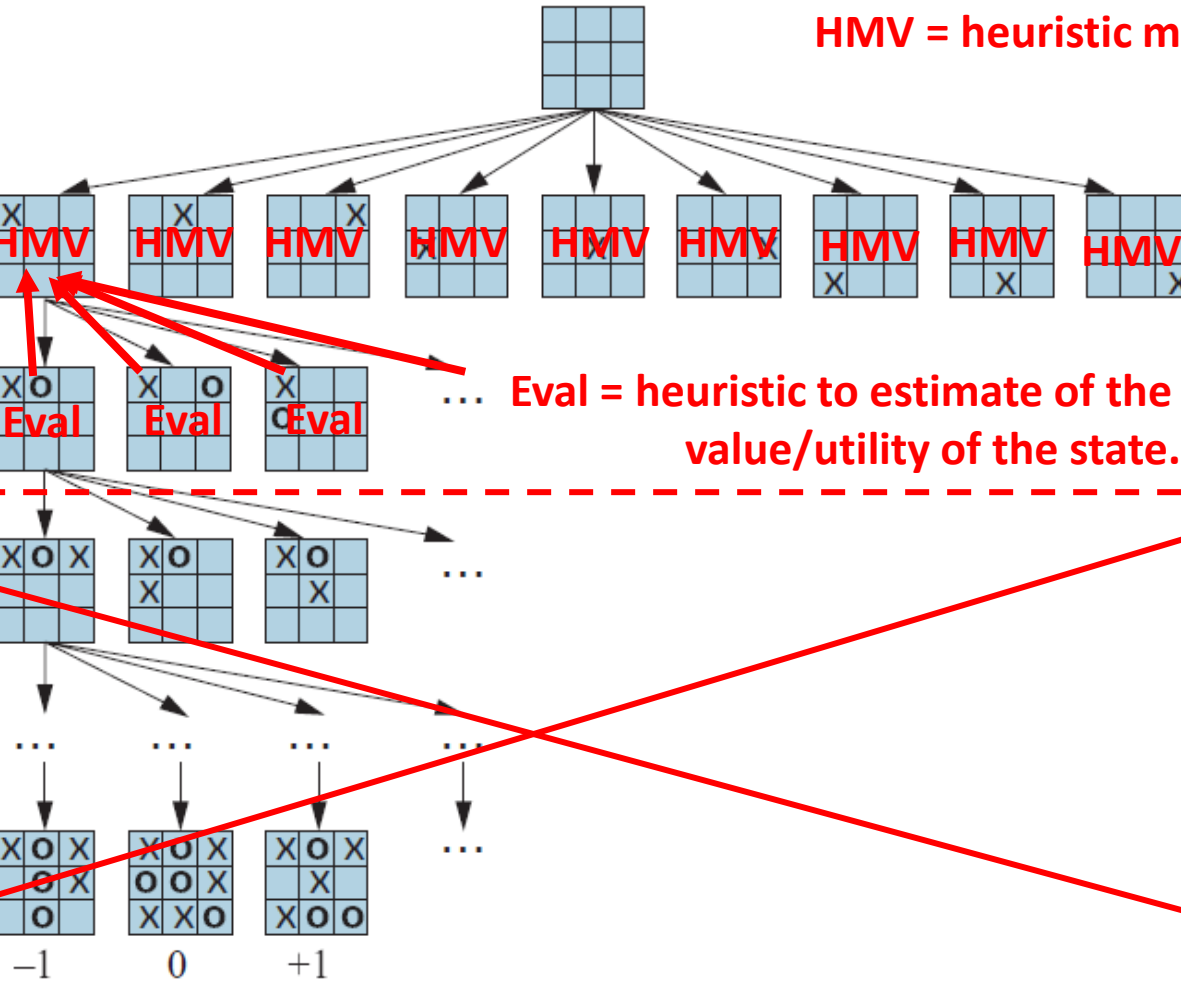
1 MIN (o)

2 MAX (x)

3 MIN (o)

TERMINAL

Utility



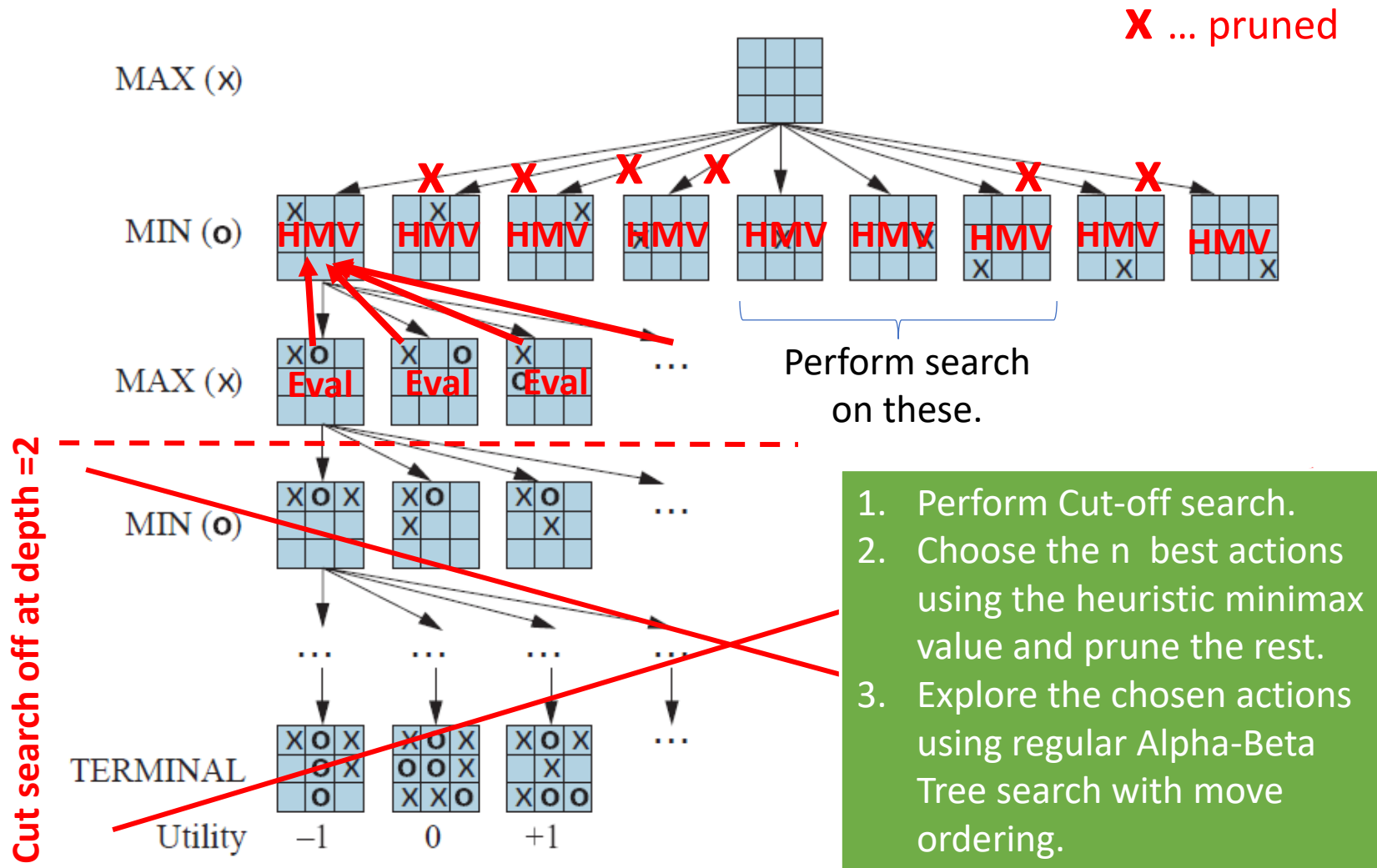
Idea: Forward pruning

Prune moves that appear poor. Poor can be evaluated in several ways:

- Low evaluation value after shallow search.
- Past experience.

Issue: May prune important moves.

Heuristic Alpha-Beta Tree Search: Forward Pruning Example



A close-up, slightly blurred image of a roulette wheel. The wheel is red with black numbers and green pockets. The text "Monte Carlo Tree Search (MCTS)" is overlaid in white, centered on the wheel. The wheel is tilted, and the numbers are visible in a circular pattern. The text is in a clean, sans-serif font.

Monte Carlo Tree Search (MCTS)

Methods for Adversarial Games

Exact Methods

- **Model as nondeterministic actions:** The opponent is seen as part of an environment with nondeterministic actions. Non-determinism is the result of the unknown moves by the opponent. We **consider all possible moves** by the opponent.
- **Find optimal decisions:** Minimax search and Alpha-Beta pruning where **each player plays optimal** to the end of the game.

Heuristic Methods

(game tree is too large)

- **Heuristic Alpha-Beta Tree Search:**
 - a. Cut off game tree and use heuristic for utility.
 - b. Forward Pruning: ignore poor moves.
- **Monte Carlo Tree search:** Estimate utility of a state by simulating complete games and average the utility.

Idea

- **Approximate $Eval(s)$** as the average utility of several simulation runs to the terminal state (called playouts).
- **Playout policy:** How to choose moves during the simulation runs? Example policies:
 - Random.
 - Heuristics for good moves developed by experts.
 - Learn good moves from self-play (e.g., with deep neural networks).
- Typically used for problems with
 - High branching factor (many possible moves).
 - Unknown or hard to define good evaluation functions.

Pure Monte Carlo Search

Find the next best move.

- Method
 1. Simulate N playouts from **current state**.
 2. Select the move with the highest win percentage.
- Converges to optimal play for stochastic games as N increases.
- **Do as many playouts as you can** given the available time.

Monte Carlo Tree Search

- Plan ahead and build a game tree using simulation.
- **Select the starting state** for playouts to focus on important parts of the game tree. It is a tradeoff between:
 - a) **Exploration**: search from states that currently have few playouts.
 - b) **Exploitation**: more playouts for states that have done well to get more accurate estimates.

Selection using Upper Confidence Bounds applied to Trees (UCT)

Tradeoff constant ($\approx \sqrt{2}$) can be optimized using experiments

$$UCB1(n) = \frac{U(n)}{N(n)} + C \sqrt{\frac{\log(N(\text{Parent}(n)))}{N(n)}}$$

Average utility
(=exploitation)

High for nodes with few
playouts (=exploration)

$U(n)$... total utility of all playouts going through node n
 $N(n)$... number of playouts through n

Policy: Select leaf with highest UCB1 score.

function MONTE-CARLO-TREE-SEARCH(*state*) *returns an action*

tree \leftarrow NODE(*state*)

while IS-TIME-REMAINING() **do**

leaf \leftarrow SELECT(*tree*)

Highest UCB1 score

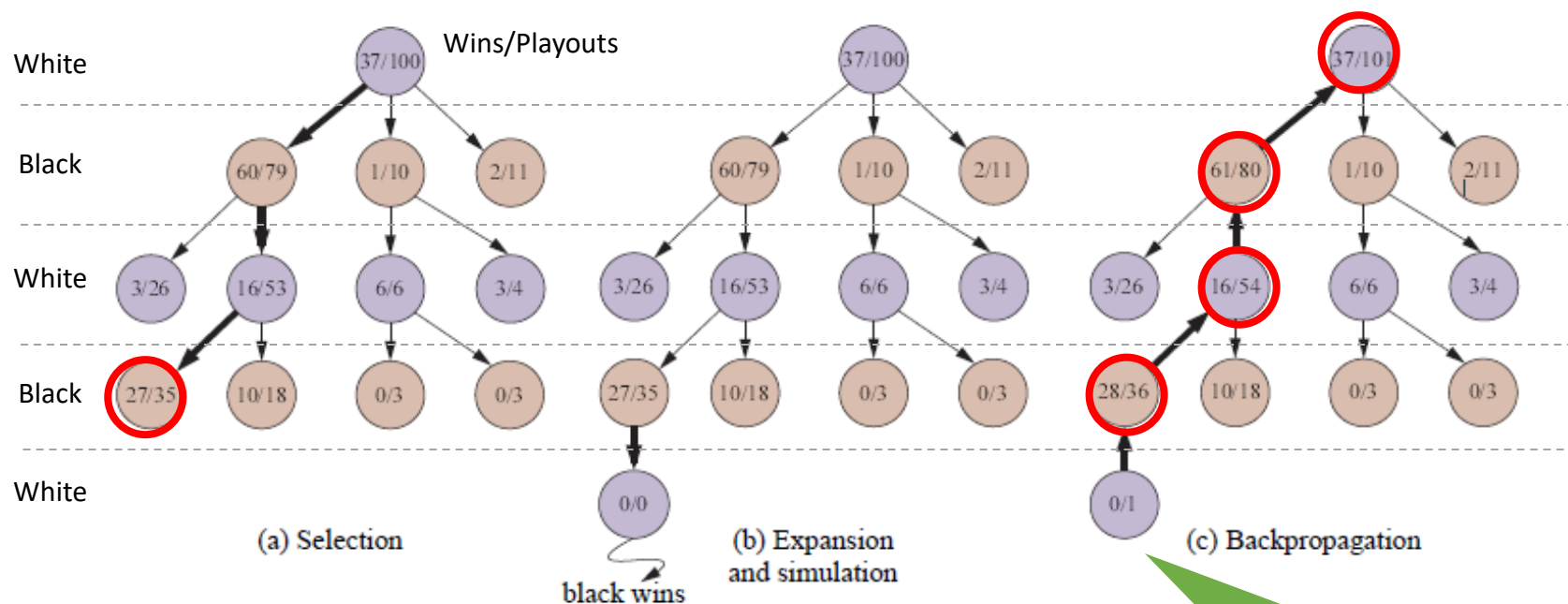
child \leftarrow EXPAND(*leaf*)

result \leftarrow SIMULATE(*child*)

BACK-PROPAGATE(*result*, *child*)

UCB1 selection favors win percentage more and more.

return the move in ACTIONS(*state*) whose node has highest number of playouts



Note: the simulation path is not recorded!

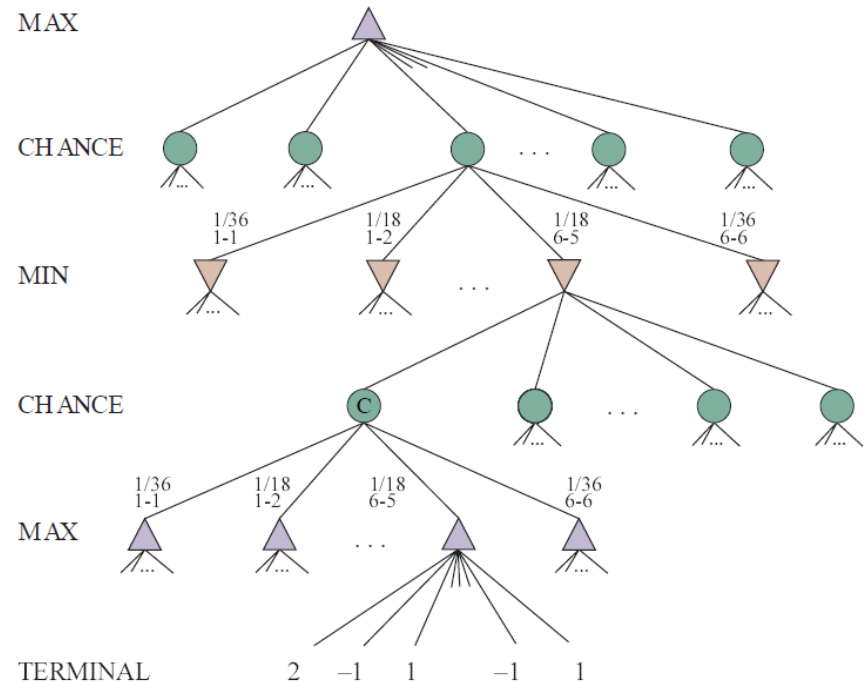
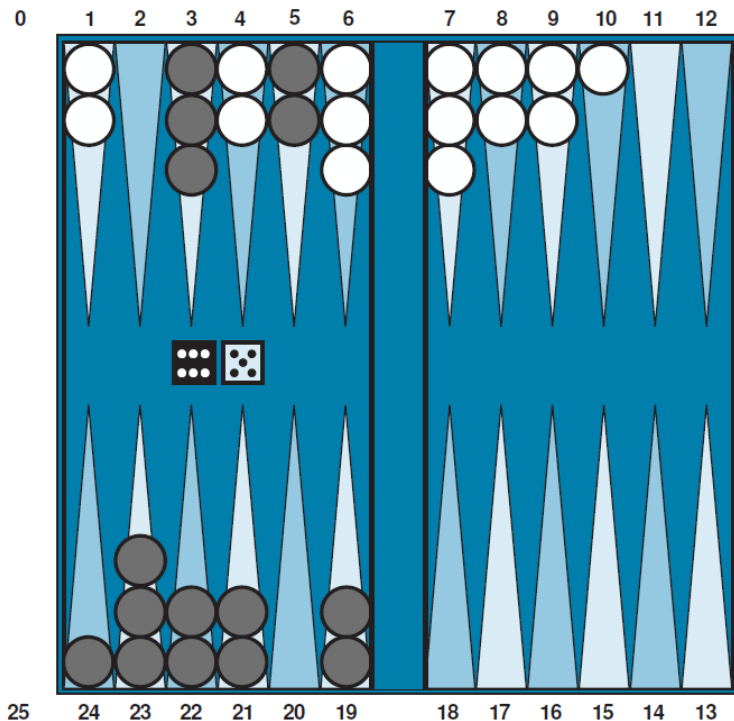
The background of the slide is a photograph of four dice on a wooden surface. There are two black dice and two white dice. The dice are scattered across the frame, with some in sharp focus and others blurred. The wooden surface has a warm, brown tone with visible grain patterns. The overall lighting is soft and even.

Stochastic Games

Games With Random Events

Stochastic Games

- Game includes a “random action” r (e.g., dice, dealt cards)
- Add **chance nodes** that calculate the expected value.



Backgammon

Expectiminimax

- Game includes a “random action” r (e.g., dice, dealt cards)
- For **chance nodes** we calculate the expected minimax value.

$Expectiminimax(s) =$

$$\left\{ \begin{array}{ll} Utility(s) & \text{if } terminal(s) \\ \max_{a \in Actions(s)} Expectiminimax(Result(s, a)) & \text{if } move = Max \\ \min_{a \in Actions(s)} Expectiminimax(Result(s, a)) & \text{if } move = Min \\ \sum_r P(r) Expectiminimax(Result(s, r)) & \text{if } move = Chance \end{array} \right.$$

- Options:
 - Use Minimax algorithm. Issue: Search tree size explodes if the number of “random actions” is large. Think of drawing cards for poker!
 - Approximate Expectiminimax with an evaluation function.
 - Perform Monte Carlo Tree Search.

Conclusion

Nondeterministic actions:

- The opponent is seen as part of an environment with nondeterministic actions. Non-determinism is the result of the unknown moves by the opponent. *All possible moves are considered.*

Optimal decisions:

- Minimax search and Alpha-Beta pruning where *each player plays optimal* to the end of the game.
- Choice nodes and Expectiminimax for stochastic games.

Heuristic Alpha-Beta Tree Search:

- Cut off game tree and use *heuristic evaluation function* for utility (based on state features).
- Forward Pruning: ignore poor moves.

Monte Carlo Tree search:

- Simulate complete games and calculate proportion of wins.
- Learn playout policy using self-play and deep learning.
- Use modified UCB1 scores to expand the game tree.

Scale only for tiny problems!

State of the Art