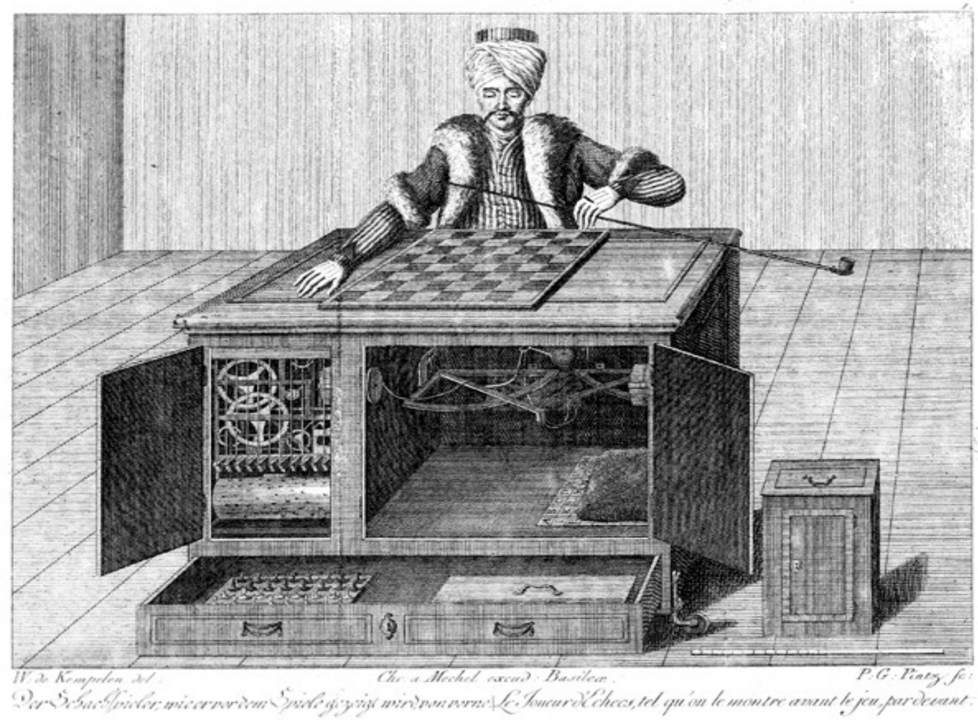
CS440/ECE448 Lecture 30: Two-Player Games

Mark Hasegawa-Johnson, 4/2022



By Karl Gottlieb von Windisch - Copper engraving from the book: Karl Gottlieb von Windisch, Briefe über den Schachspieler des Hrn. von Kempelen, nebst drei Kupferstichen die diese berühmte Maschine vorstellen. 1783.Original Uploader was Schaelss (talk) at 11:12, 7. Apr 2004., Public Domain, https://commons.wikimedia.org/w/index.php?curid=424092

Why study games?

- Games are a traditional hallmark of intelligence
- Games are easy to formalize
- Games can be a good model of real-world competitive or cooperative activities
  - Military confrontations, negotiation, auctions, etc.

Games vs. single-agent search

- We don't know how the opponent will act
- The solution is not a fixed sequence of actions from start state to goal state, but a ***strategy*** or ***policy***

Definition of **policy**: a policy is a function $\pi: \mathcal{S} \rightarrow \mathcal{A}$ that maps from world states, $s \in \mathcal{S}$, to actions, $\mathrm{a} \in \mathcal{A}$.

Outline

- Alternating two-player zero-sum games
- Minimax search
- Limited-horizon computation and heuristic evaluation functions
- Alpha-beta search
- Computational complexity of minimax and alpha-beta

Alternating two-player zero-sum games

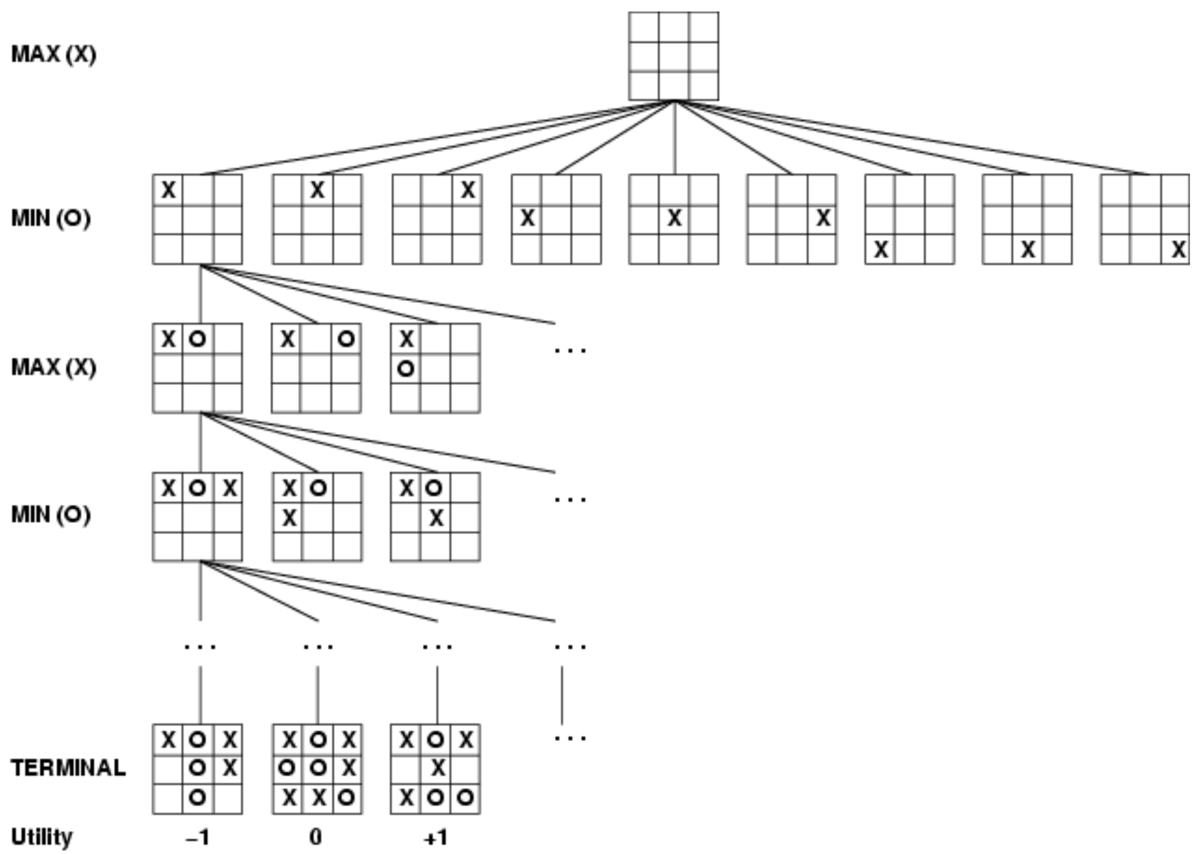
- Players take turns
- Each game outcome or **terminal state** has a **utility** for each player (e.g., 1 for win, 0 for tie, -1 for loss)
- The sum of both players' utilities is a constant, e.g.,

Utility(player 0) + Utility(player 1) = 0

- Player 0 tries to maximize Utility(player 0). Let's call this player "Max"
- Player 1 tries to minimize Utility(player 0). Let's call this player "Min"

Game tree

- A game of tic-tac-toe between two players, “max” and “min”

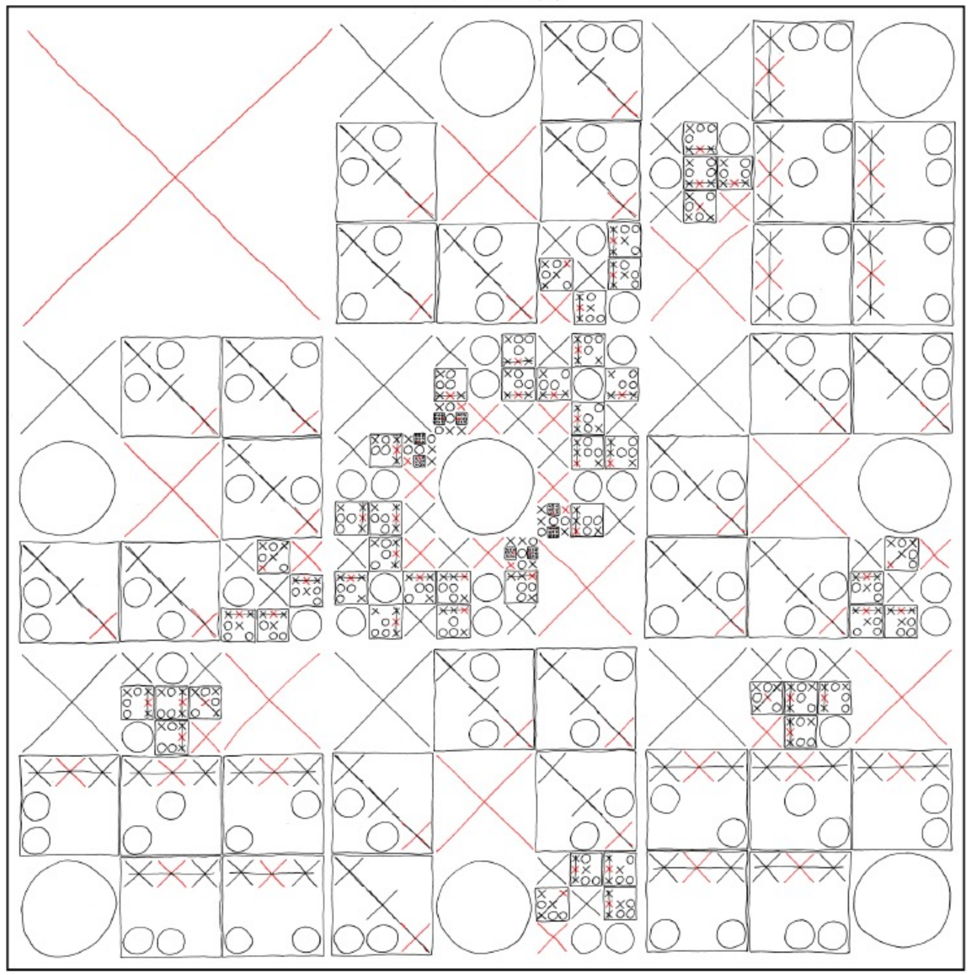


COMPLETE MAP OF OPTIMAL TIC-TAC-TOE MOVES

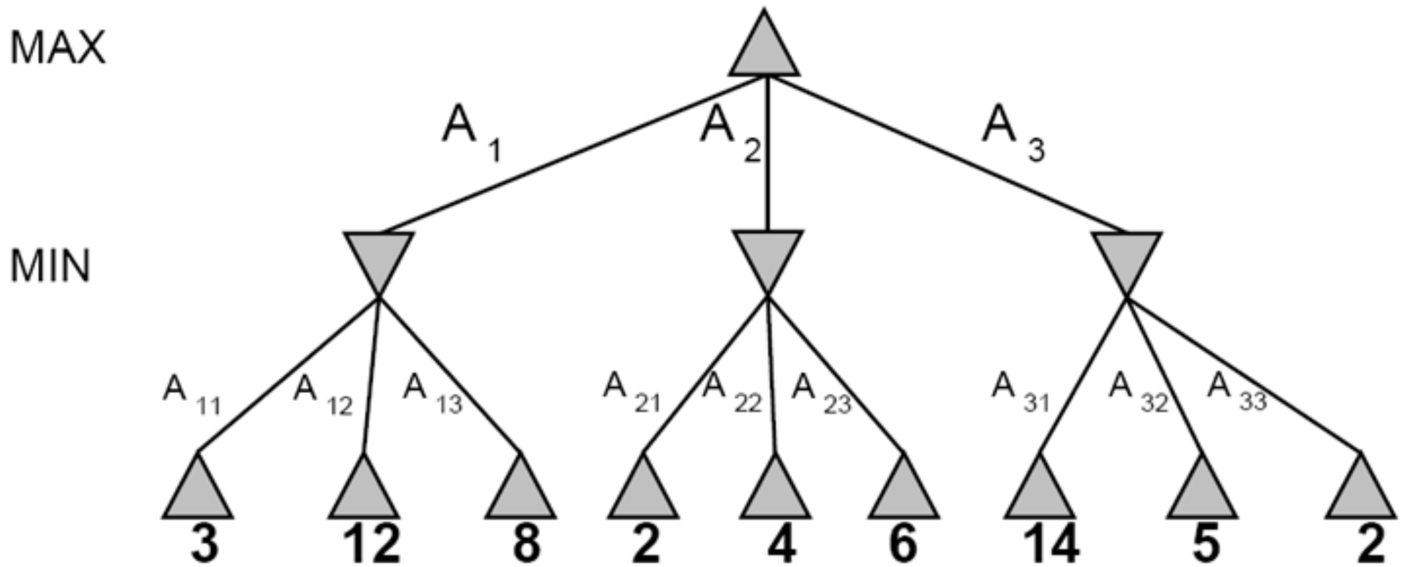
YOUR MOVE IS GIVEN BY THE POSITION OF THE LARGEST RED SYMBOL ON THE GRID. WHEN YOUR OPPONENT PICKS A MOVE, ZOOM IN ON THE REGION OF THE GRID WHERE THEY WENT. REPEAT.

http://xkcd.com/832/

MAP FOR X:



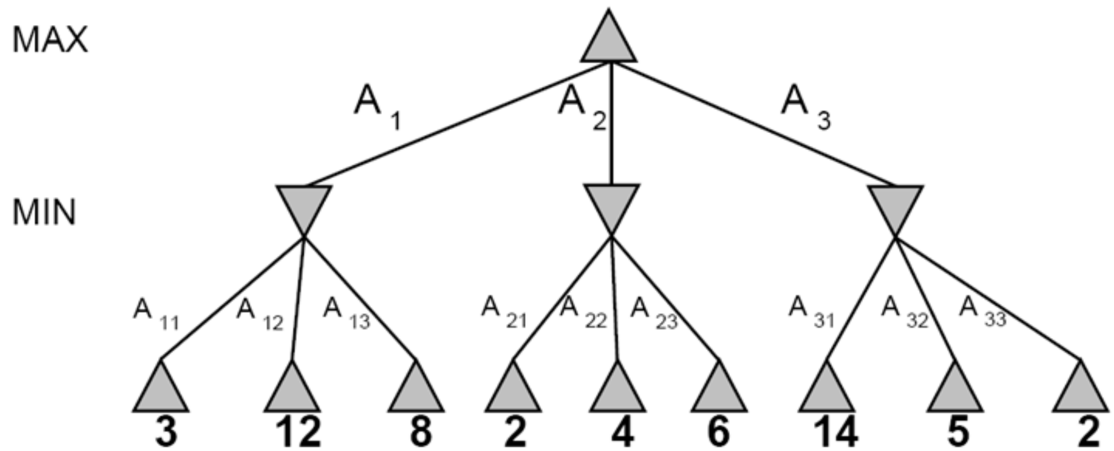
A more abstract game tree

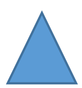


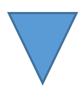
Terminal utilities (for MAX)

A *depth-two* game

Standard notation for game trees



 = game state from which MAX can play

 = game state from which MIN can play

number = value of that game state for MAX

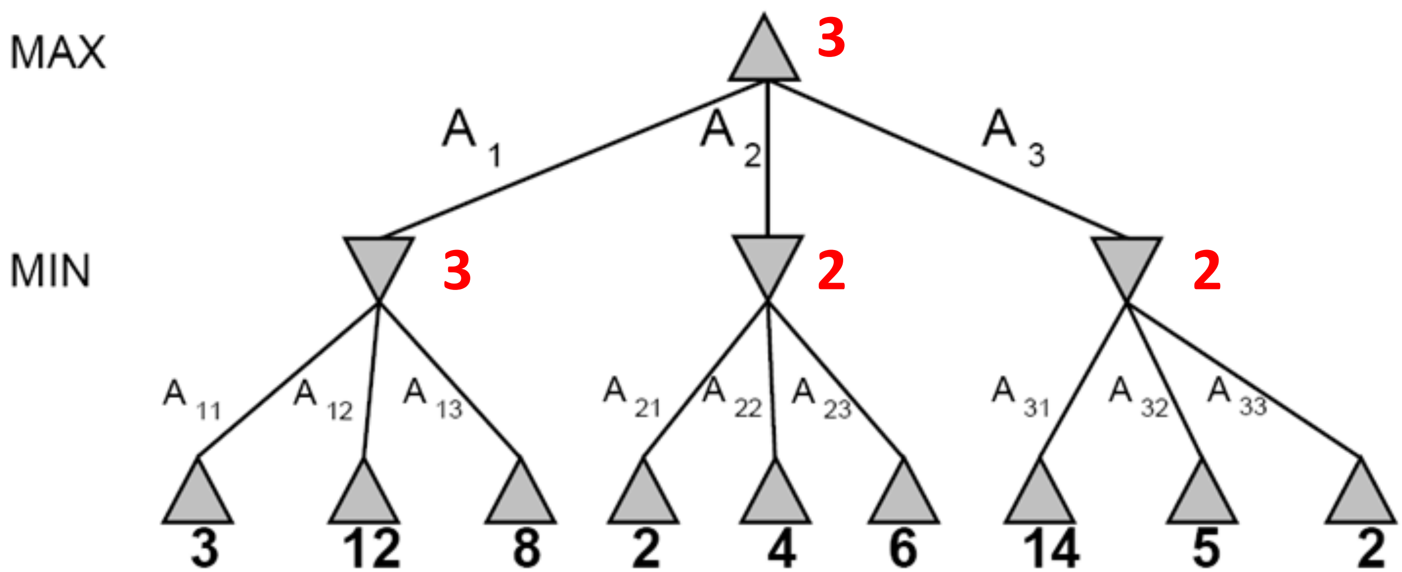
Outline

- Alternating two-player zero-sum games
- Minimax search
- Limited-horizon computation and heuristic evaluation functions
- Alpha-beta search
- Computational complexity of minimax and alpha-beta

The rules of every game

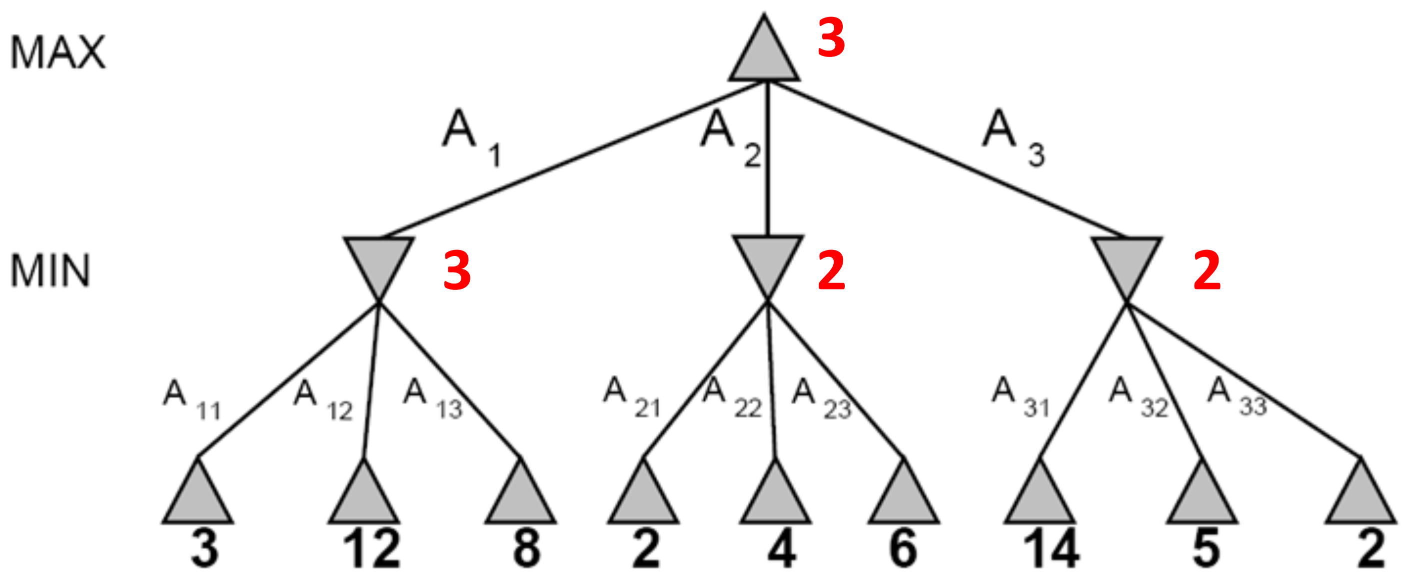
- Every possible outcome has a value (or "utility") for me.
- Zero-sum game: if the value to me is +V, then the value to my opponent is –V.
- Phrased another way:
  - My rational action, on each move, is to choose a move that will maximize the value of the outcome
  - My opponent's rational action is to choose a move that will minimize the value of the outcome
- Call me "[Max]()"
- Call my opponent "[Min]()"

Game tree search



- **Minimax value of a node**: the utility (for MAX) of being in the corresponding state, assuming perfect play on both sides
- **Minimax strategy:** Choose the move that gives the best worst-case payoff

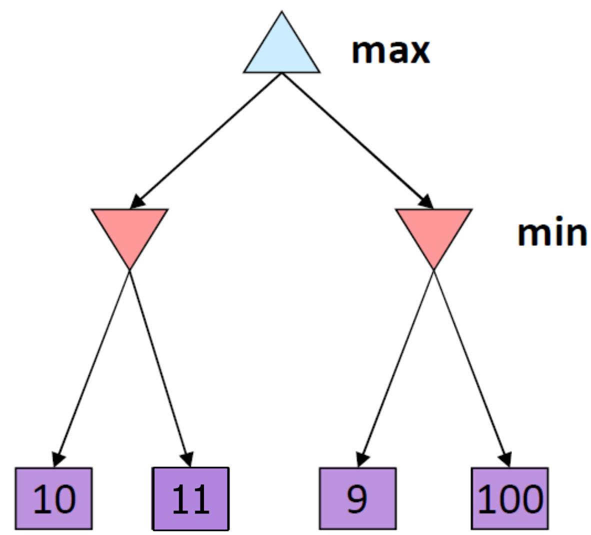
Computing the minimax value of a node



- **Minimax**(*node*) =
  - Utility(*node*) if *node* is terminal
  - $\max_{action}$ **Minimax**(Succ(*node, action*)) if *player* = MAX
  - $\min_{action}$ **Minimax**(Succ(*node, action*)) if *player* = MIN

Optimality of minimax

- The minimax strategy is optimal against an optimal opponent
- What if your opponent is suboptimal?
- If you play using the **minimax-optimal** sequence of moves, then the utility you earn will always be **greater than or equal** to the amount that you predict.

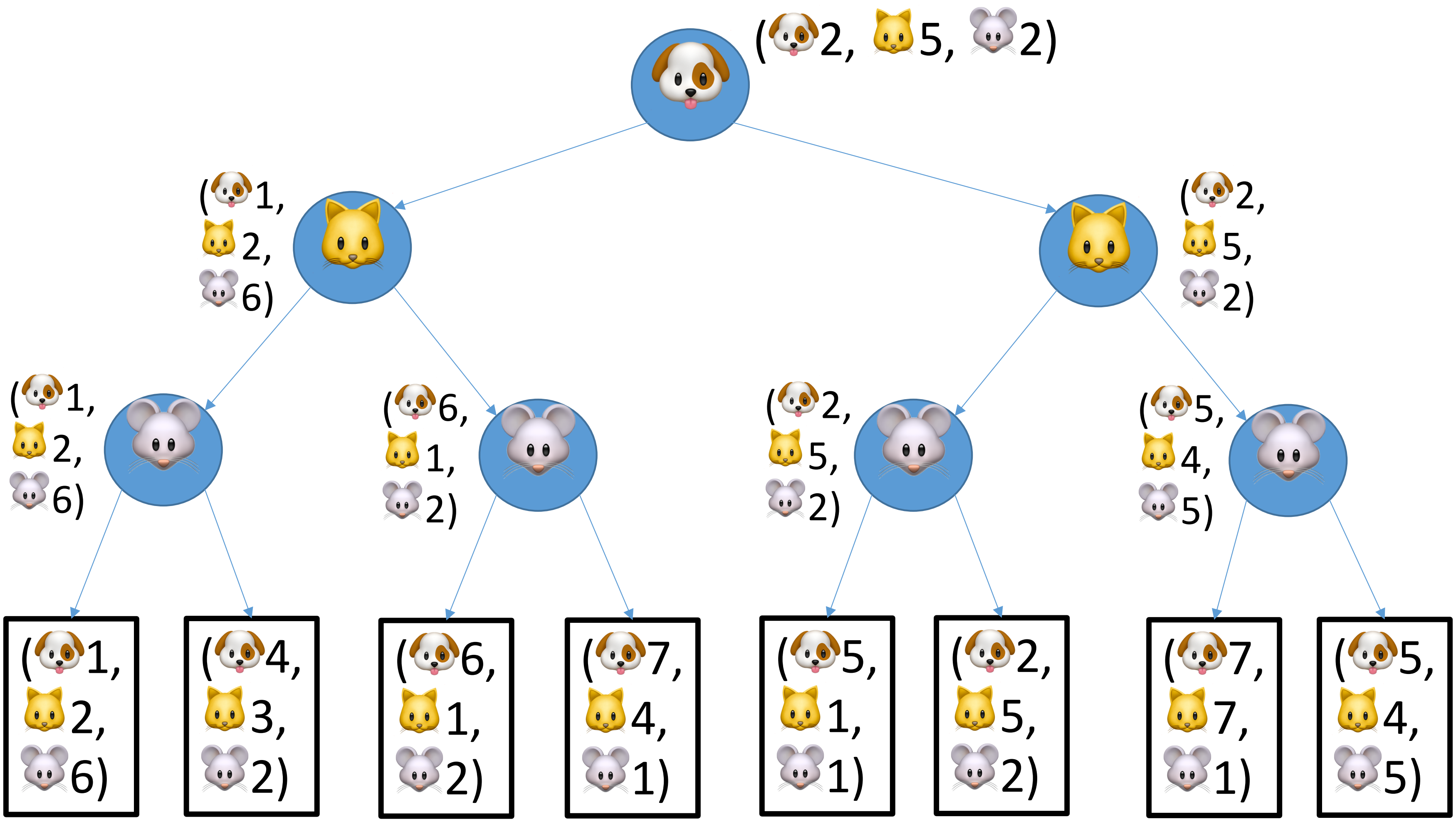


Example from D. Klein and P. Abbeel

Multi-player games; Non-zero-sum games

- More than two players. For example:
  - Dog (🐶) tries to maximize the number of doggie treats
  - Cat (🐱) tries to maximize the number of cat treats
  - Mouse (🐭) tries to maximize the number of mouse treats
- Non-zero-sum. We can't just assume that Min's score is the opposite of Max's. Instead, utilities are now tuples. For example:
  - (🐶5, 🐱8, 🐭2) = 5 doggie treats, 8 kitty treats, 2 mouse treats
- Each player maximizes their own utility at their node

Minimax in multi-player & non-zero-sum games



Outline

- Alternating two-player zero-sum games
- Minimax search
- Limited-horizon computation and heuristic evaluation functions
- Alpha-beta search
- Computational complexity of minimax and alpha-beta

Games vs. single-agent search

- We don't know how the opponent will act
  - The solution is not a fixed sequence of actions from start state to goal state, but a ***strategy*** or ***policy*** (a mapping from state to best move in that state)

Games vs. single-agent search

- We don't know how the opponent will act
  - The solution is not a fixed sequence of actions from start state to goal state, but a ***strategy*** or ***policy*** (a mapping from state to best move in that state)
- Efficiency is critical to playing well
  - The time to make a move is limited
  - The branching factor, search depth, and number of terminal configurations are huge
    - In chess, branching factor ≈ 35 and depth ≈ 100, giving a search tree of 10^{154} nodes
      - Number of atoms in the observable universe ≈ 10^{80}
  - This rules out searching all the way to the end of the game

Limited-Horizon Search

In a practical game, we compute minimax to a limited depth

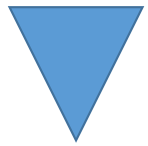
- Depth=1: evaluate every possible current move, look at the resulting game state, decide which resulting game state looks the best, and take that action.
  - Computational complexity to choose your next move: $\mathcal{O}\{b\}$, if there are b possible moves.
- Depth=2: evaluate every possible current move, and every move that your opponent might make in response, and then look at resulting game states.
  - Computational complexity to choose your next move: $\mathcal{O}\{b^2\}$.
- Depth=3: evaluate every possible sequence of three moves (mine, my opponent's, then mine), and look at the resulting game states.
  - Computational complexity to choose your next move: $\mathcal{O}\{b^3\}$.

Evaluation functions

In order to evaluate the quality of a game state $s \in \mathcal{S}$, we need to design an evaluation function $v(s)$. It should have the following properties:

- $v(s)$ should be a reasonable estimate of the outcome of the game, but
- It must be possible to compute $v(s)$ quickly, i.e., typically we desire that its computational complexity is no more than $\mathcal{O}\{b\}$. If its complexity was higher, then we might get better results by using a cheaper evaluation function in a deeper minimax search.

Example: Depth 1 search, Chess



In chess, traditionally, the black player is MIN.

What move should MIN choose, from this board position?

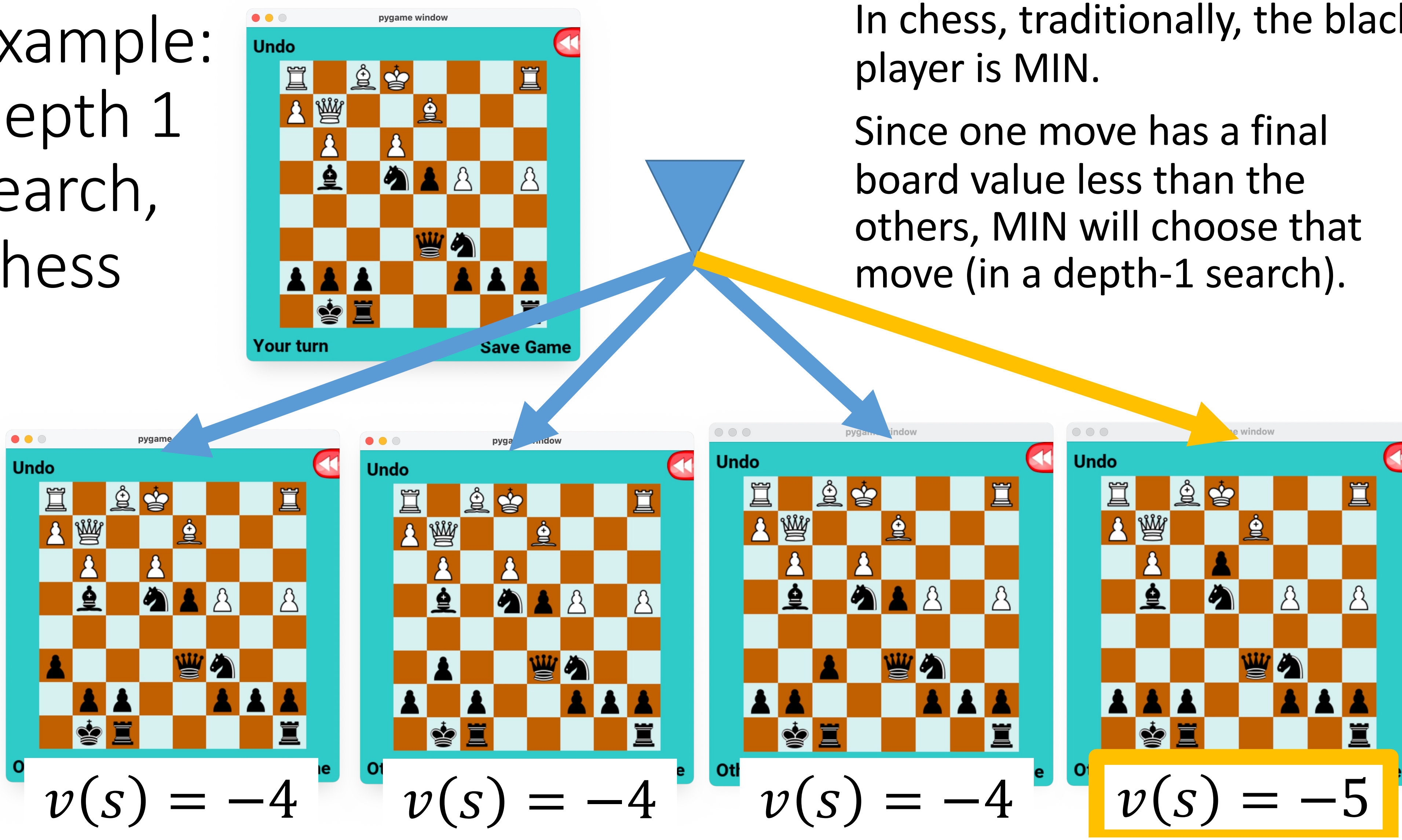
Graphics: created by the PyChess community.

Game board shown: game1.txt from the MP5 distribution.

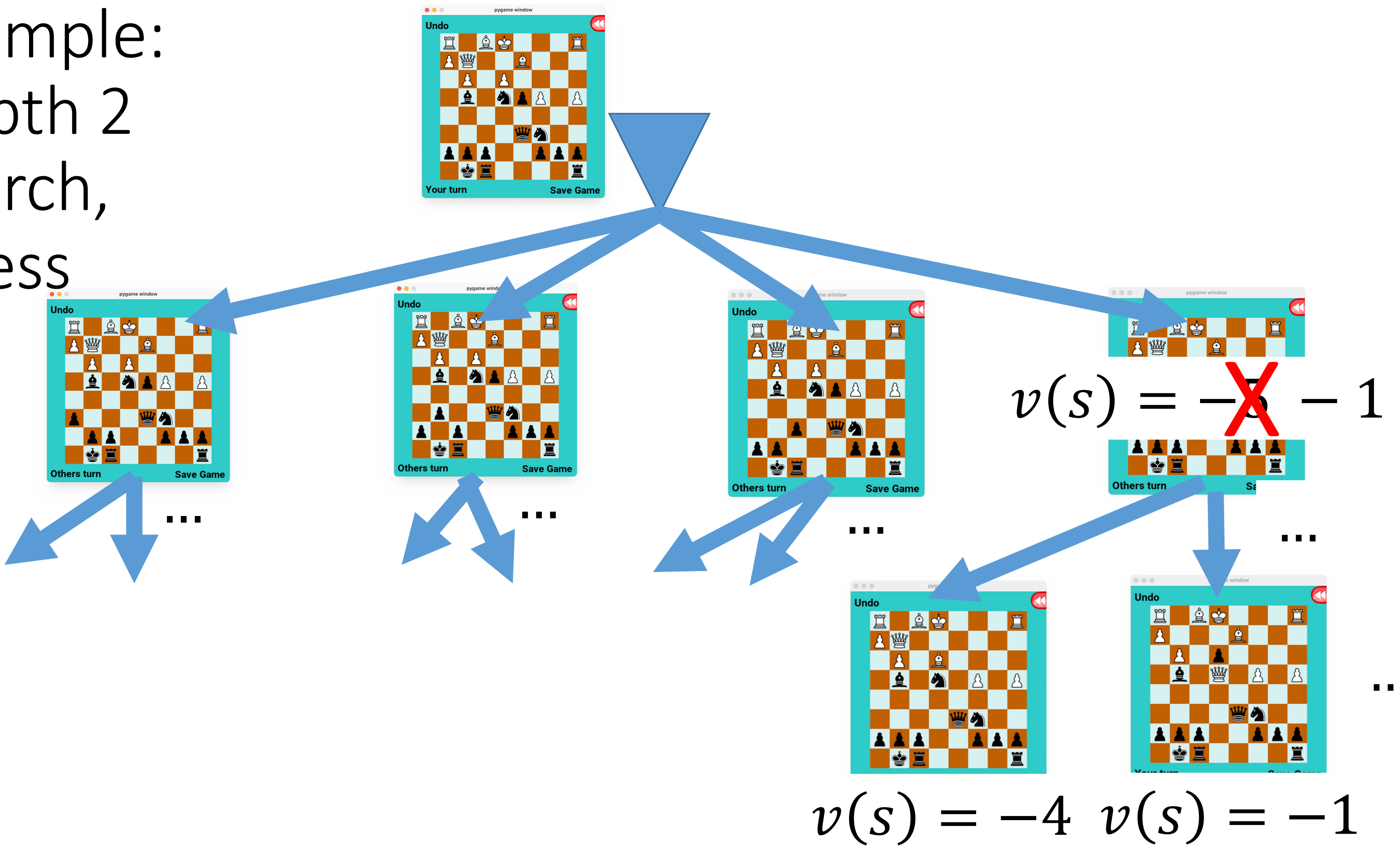
Example: Depth 1 search, Chess

In chess, traditionally, the black player is MIN.

Since one move has a final board value less than the others, MIN will choose that move (in a depth-1 search).

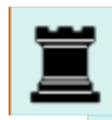
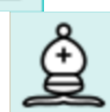
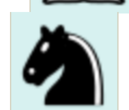
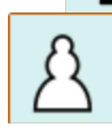


Example: Depth 2 search, Chess



Typical chess evaluation function

Each side receives:

- 9 points per remaining queen 
- 5 points per remaining rook 
- 3 points per remaining bishop 
- 3 points per remaining knight 
- 1 point per remaining pawn 

$v(s)$ = points for white - points for black

The PyChess evaluation function provides extra point depending on the location of each piece on the board.

Evaluation functions in general

Evaluation function must be reasonably accurate, but computationally simple. Often this means a linear evaluation function:

$$v(s) = w_1 f_1(s) + w_2 f_2(s) + \cdots$$

- $f_1(s), f_2(s), \ldots$ are features of the game state s
- $w_1, w_2 \ldots$ are real-valued weights.

Notice: this is just a one-layer neural net, with input vector $f(s) = [f_1(s), f_2(s), \ldots]$ and weight vector $w = [w_1, w_2, \ldots]$.

Recently, deeper neural nets are also sometimes used.

Cutting off search

- **Horizon effect:** you may incorrectly estimate the value of a state by overlooking an event that is just beyond the depth limit
  - For example, a damaging move by the opponent that can be delayed but not avoided
- Remedies: search a small number of possible extensions to depth+1.
  - **Quiescence search:** extend only “unstable” moves, e.g., moves that capture a piece.
  - **Singular extension:** extend only very strong moves.
  - **Stochastic search**: randomly sample a small number of possible future paths.

Outline

- Alternating two-player zero-sum games
- Minimax search
- Limited-horizon computation and heuristic evaluation functions
- Alpha-beta search
- Computational complexity of minimax and alpha-beta

Computational complexity of minimax

- Suppose that, at each game state, there are b possible moves
- Suppose we search to a depth of d
- Then the computational complexity is $O\{b^d\}$!

Basic idea of alpha-beta pruning

- Computational complexity of minimax is $O\{b^d\}$
- There is no known algorithm to make it polynomial time
- But… can we reduce the exponent? For example, could we make the complexity $O\{b^{d/2}\}$?
- If we could do that, then it would become possible to search twice as far, using the same amount of computation. This could be the difference between a beginner chess player vs. a grand master.

Basic idea of alpha-beta pruning

- The basic idea of alpha-beta pruning is to reduce the complexity of minimax from $O\{b^d\}$ to $O\{b^{d/2}\}$.
- We can do this by only evaluating half of the levels.
- How can we ”only evaluate half the levels” without losing accuracy?
- Why it works: It is possible to compute the exact minimax decision without expanding every node in the game tree

The pruning thresholds, alpha and beta

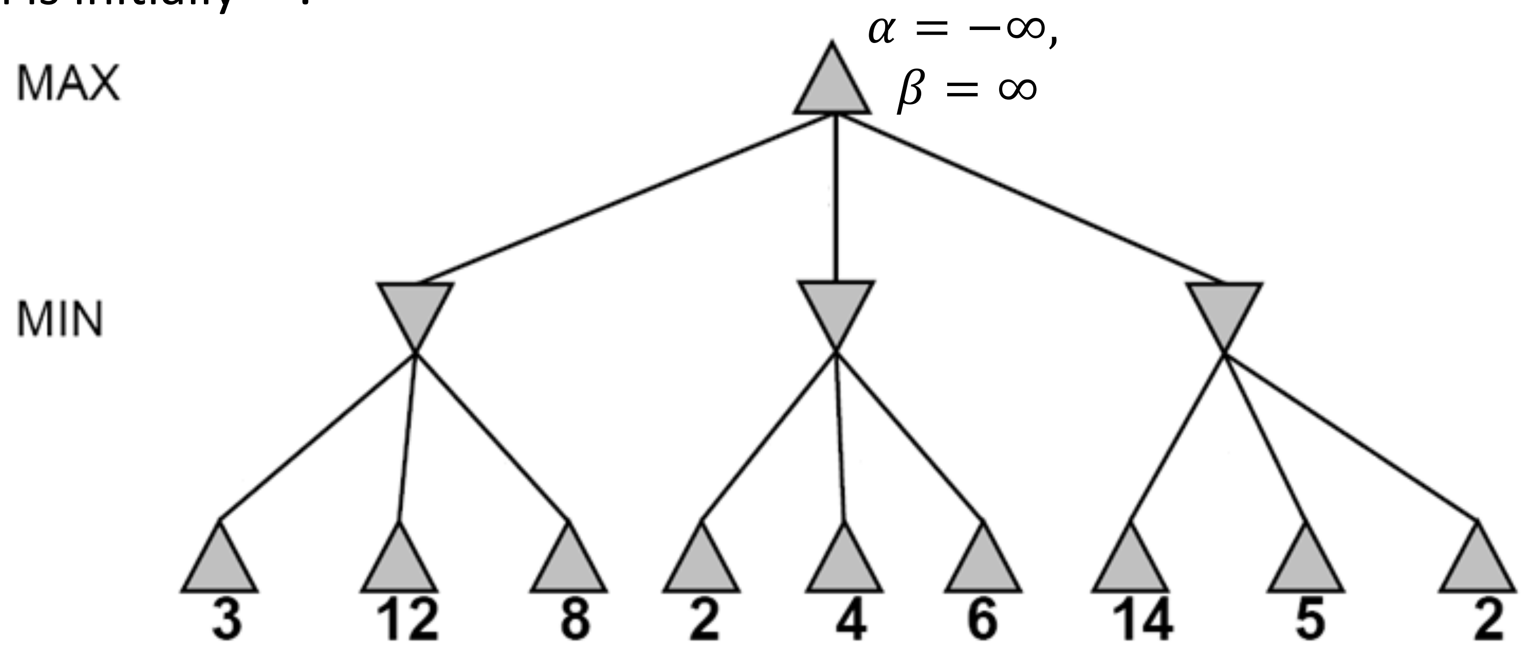
Alpha-beta pruning requires us to keep track of two pruning thresholds, alpha and beta.

- alpha (α) is the highest score that MAX knows how to force MIN to accept.
- beta (β) is the lowest score that MIN knows how to force MAX to accept.
- $\alpha \leq \beta$

Alpha-beta pruning

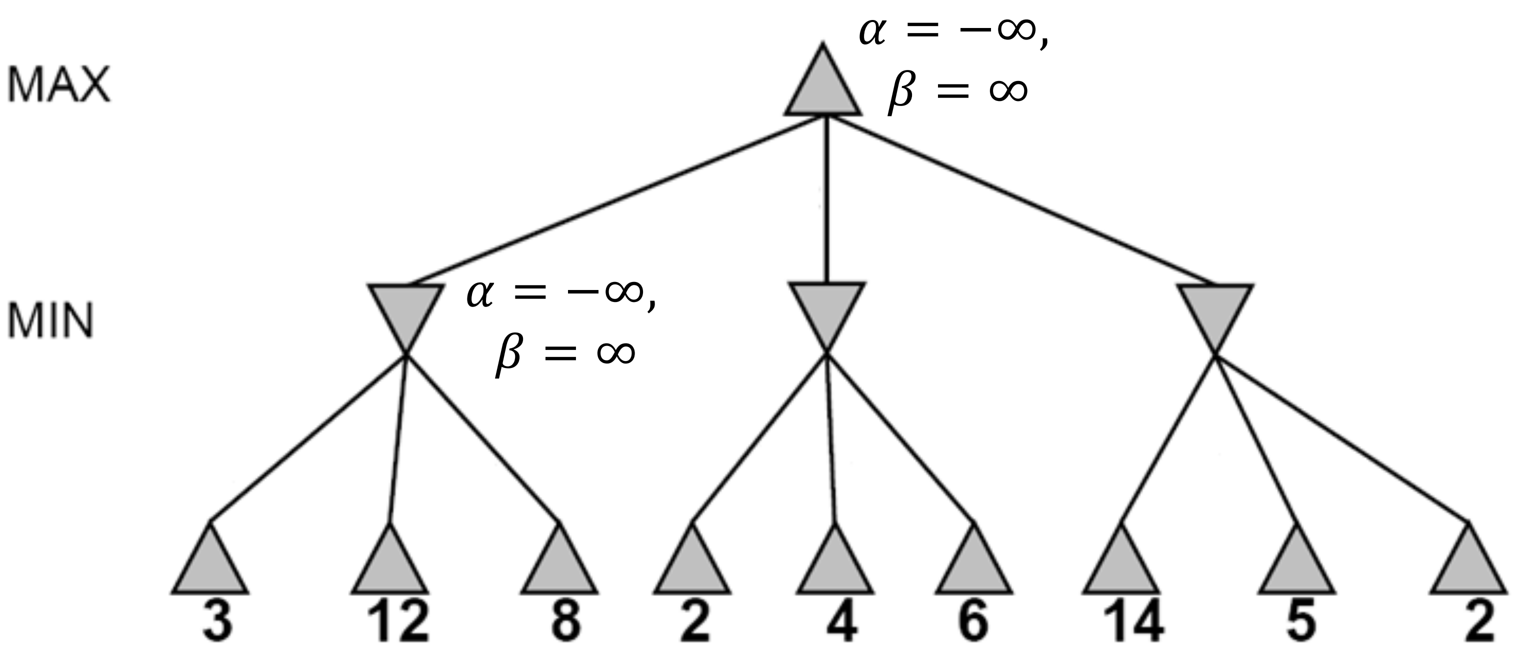
Initialize:

- alpha (α) is the highest score that MAX knows how to force MIN to accept, which is initially $-\infty$.
- beta (β) is the lowest score that MIN knows how to force MAX to accept, which is initially ∞.



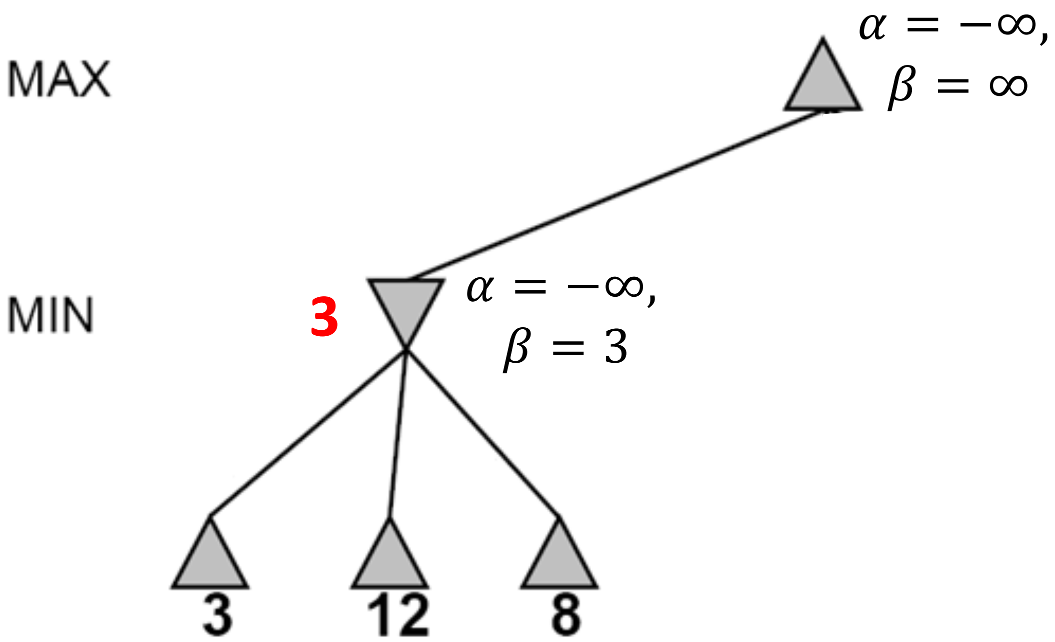
Alpha-beta pruning

Inheritance: Child inherits alpha and beta from its parent



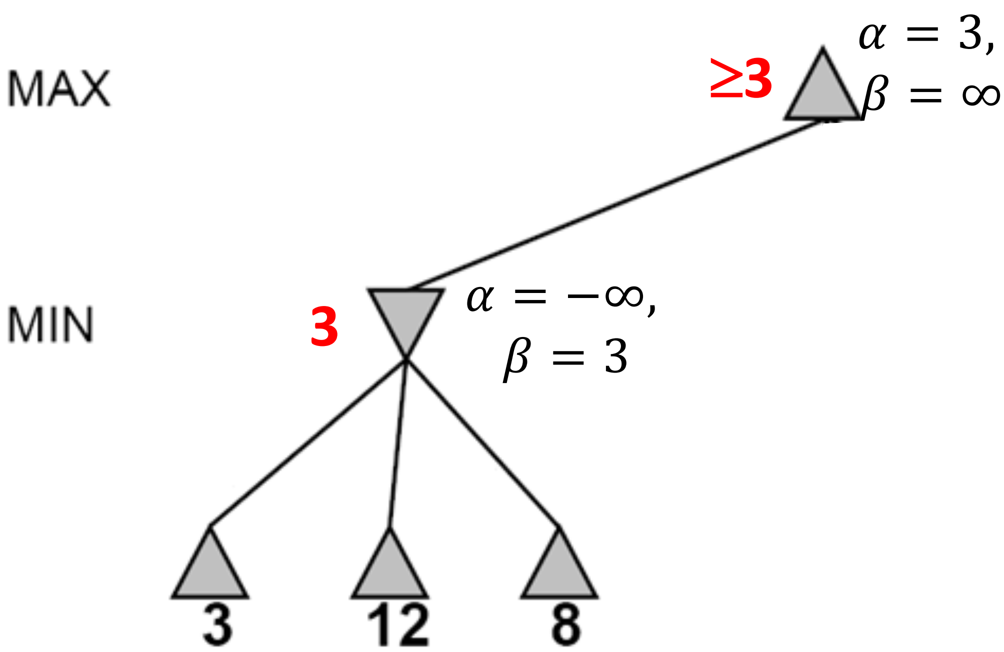
Alpha-beta pruning

Update: a **min** node can update beta. A max node can update alpha.



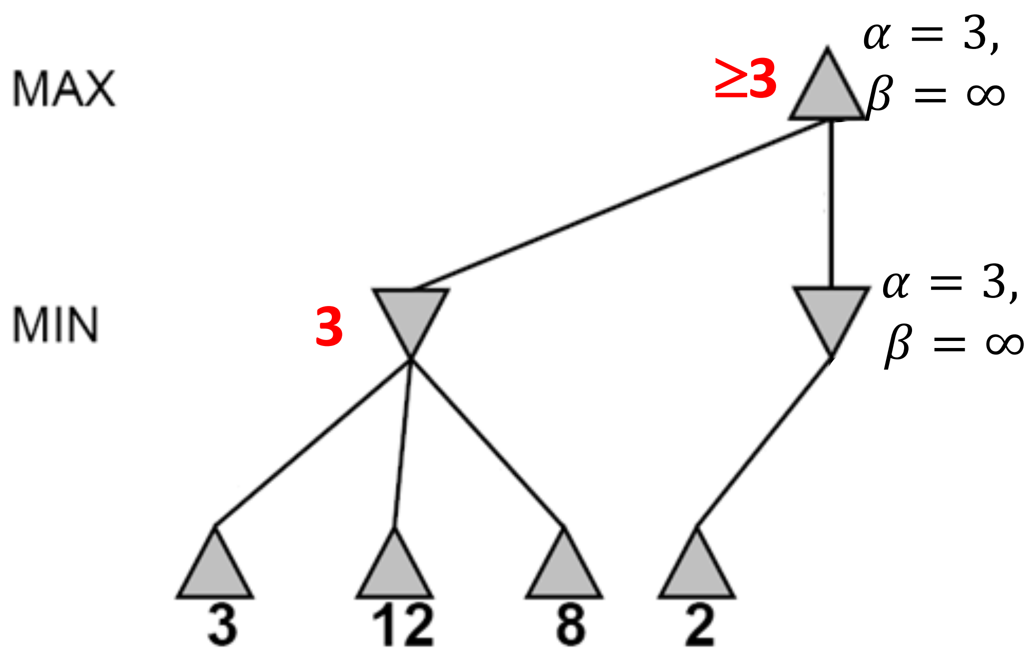
Alpha-beta pruning

Update: a min node can update beta. A **max** node can update alpha.



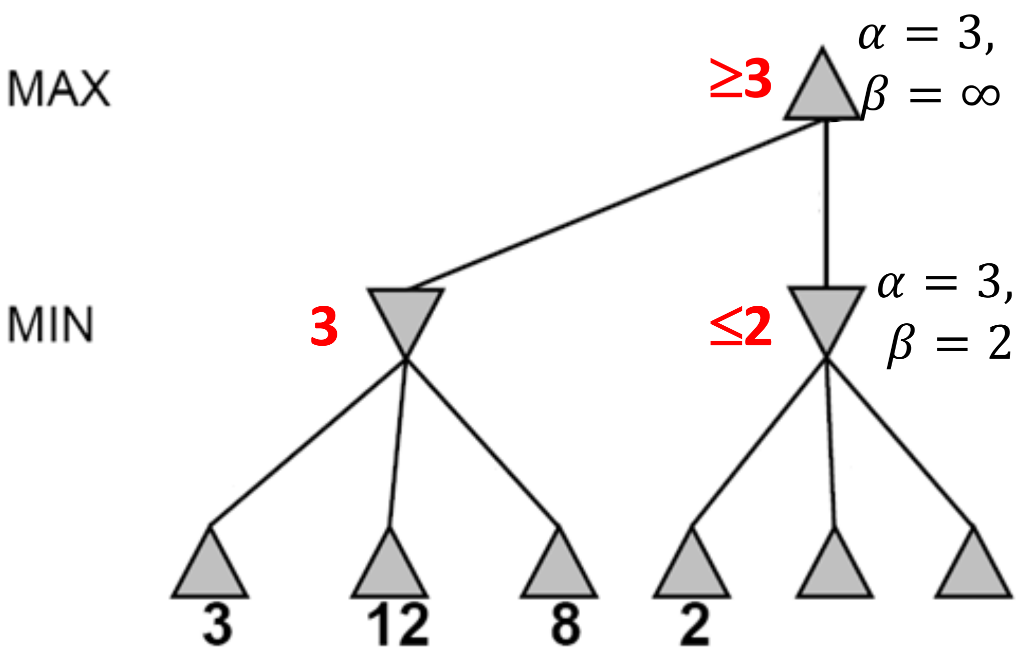
Alpha-beta pruning

Inheritance: Child inherits alpha and beta from its parent



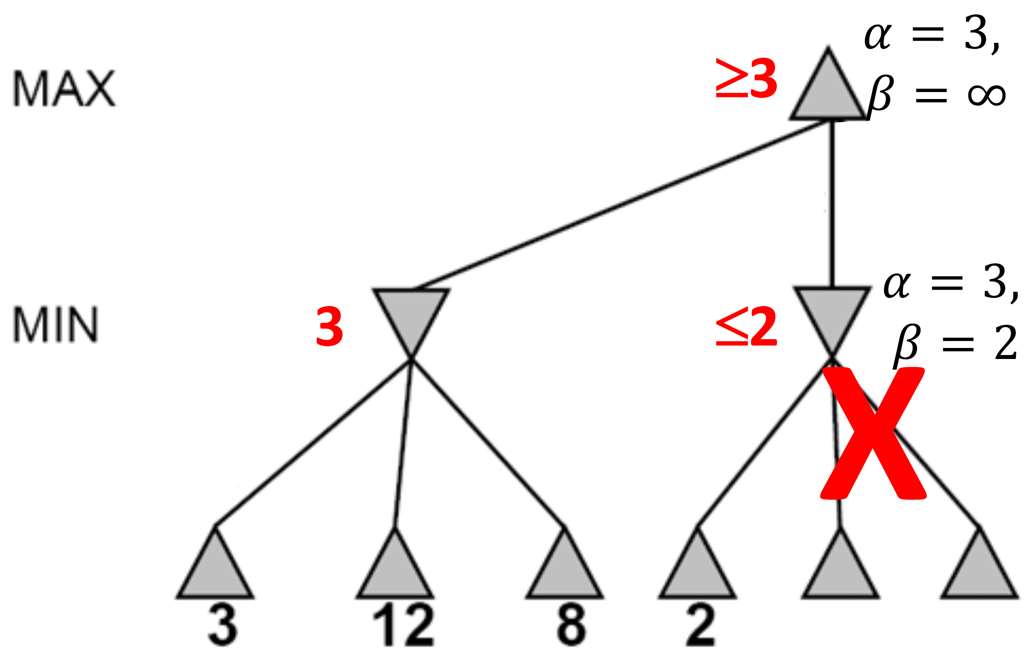
Alpha-beta pruning

Update: a min node can update beta. A max node can update alpha.



Alpha-beta pruning

Pruning: If beta ever falls below alpha, prune any remaining children, and return.

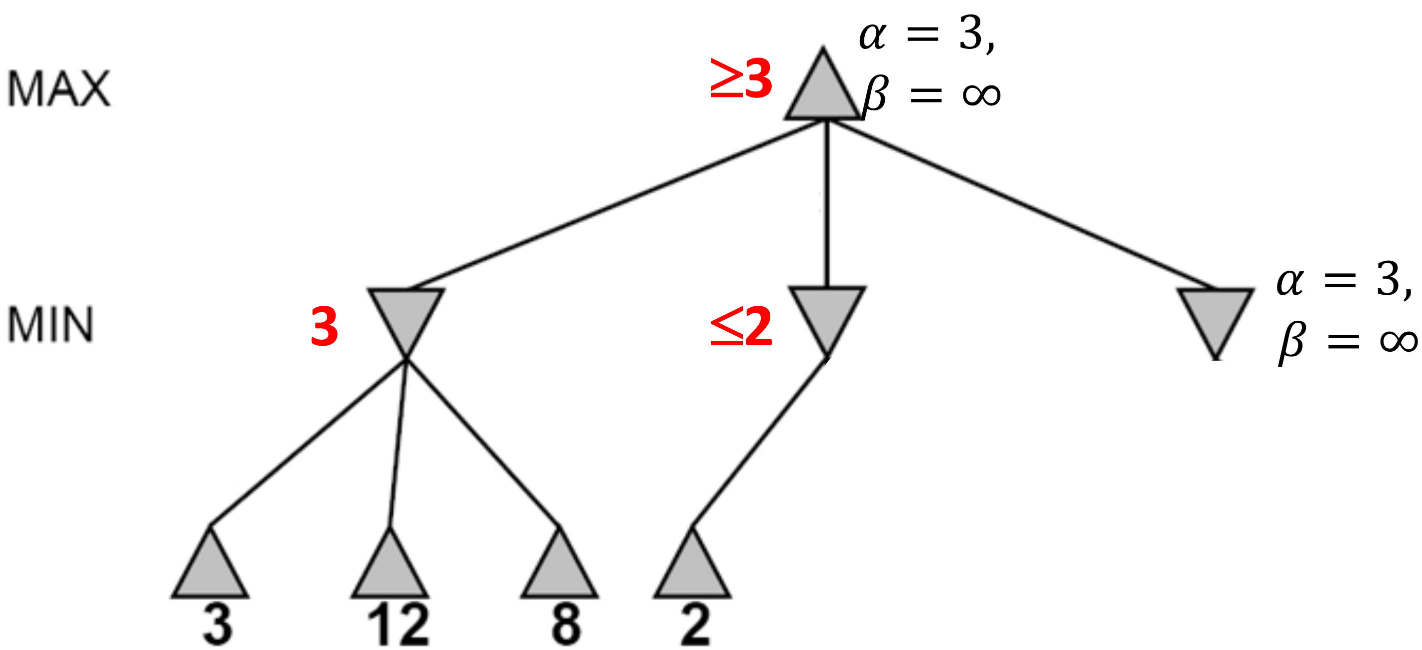


PRUNE!

- If MAX lets us get to this state, then MIN would achieve a final score <=2
- Therefore MAX will never let us get to this state!
- Therefore there’s no need to score the remaning children of this node.

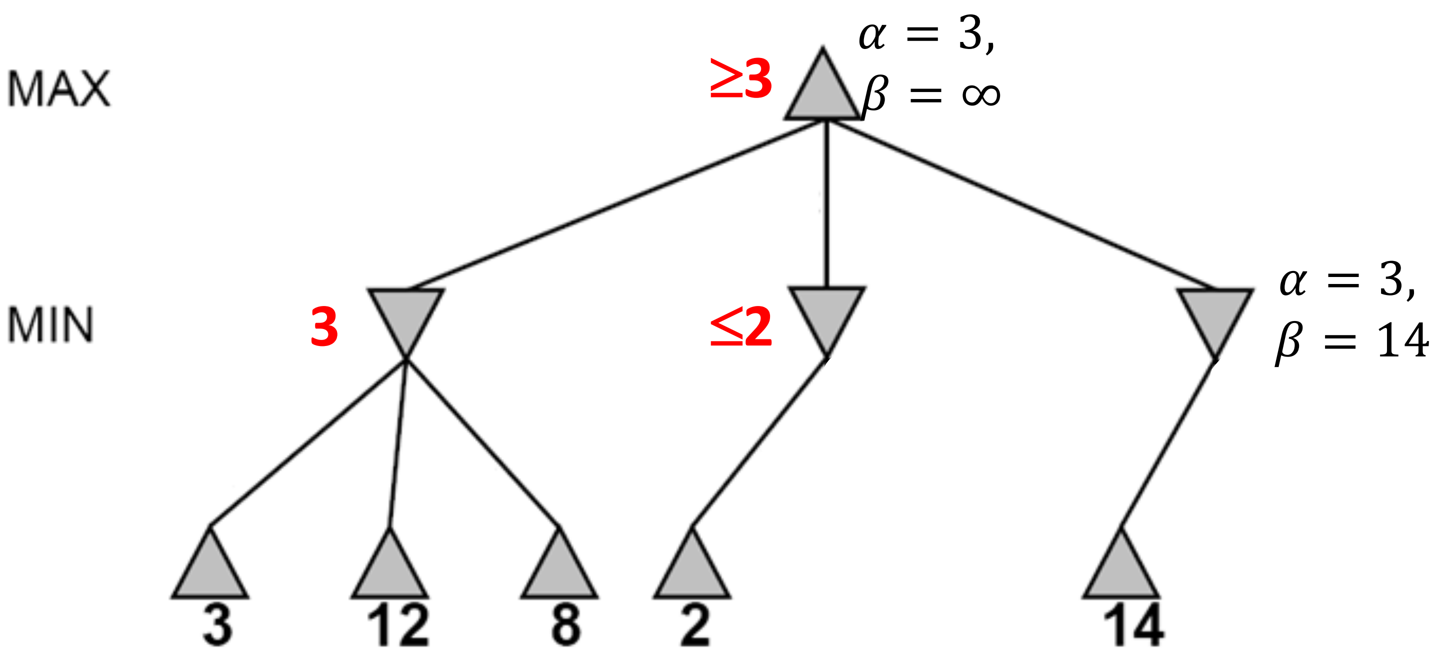
Alpha-beta pruning

Inheritance: Child inherits alpha and beta from its parent



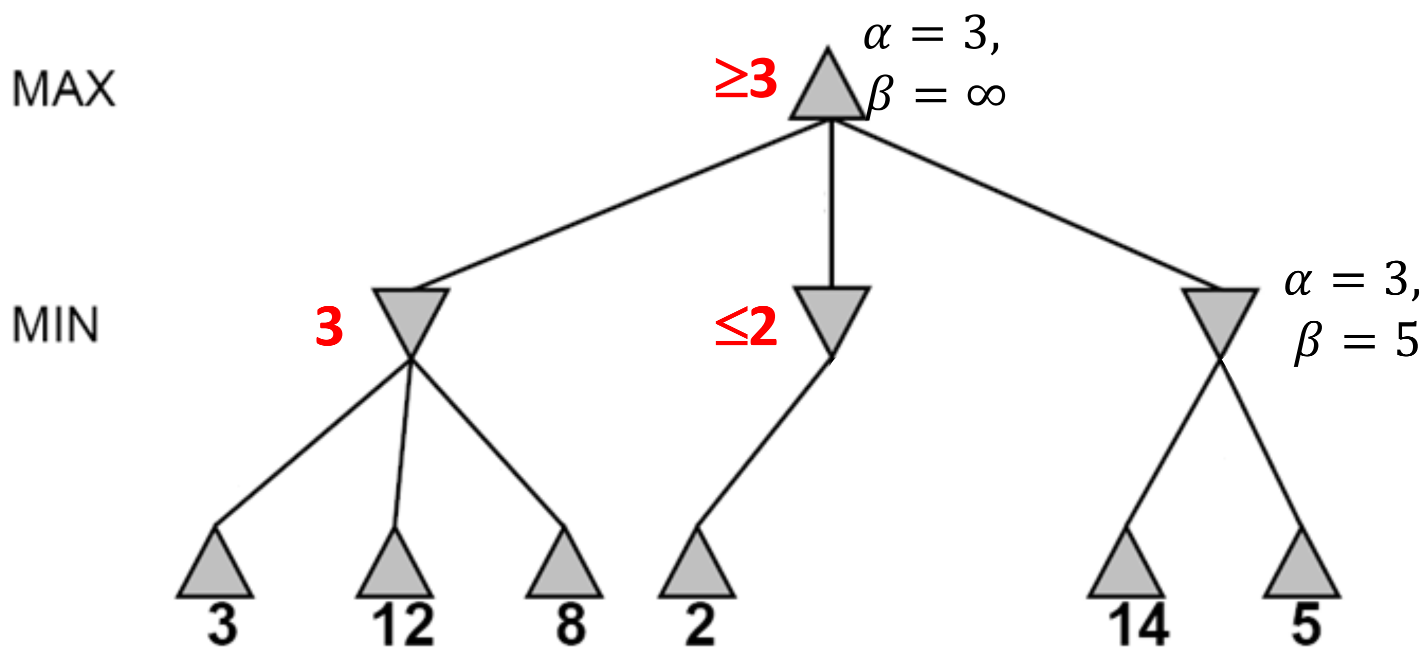
Alpha-beta pruning

<u>Update</u>: a min node can update beta. A max node can update alpha.



Alpha-beta pruning

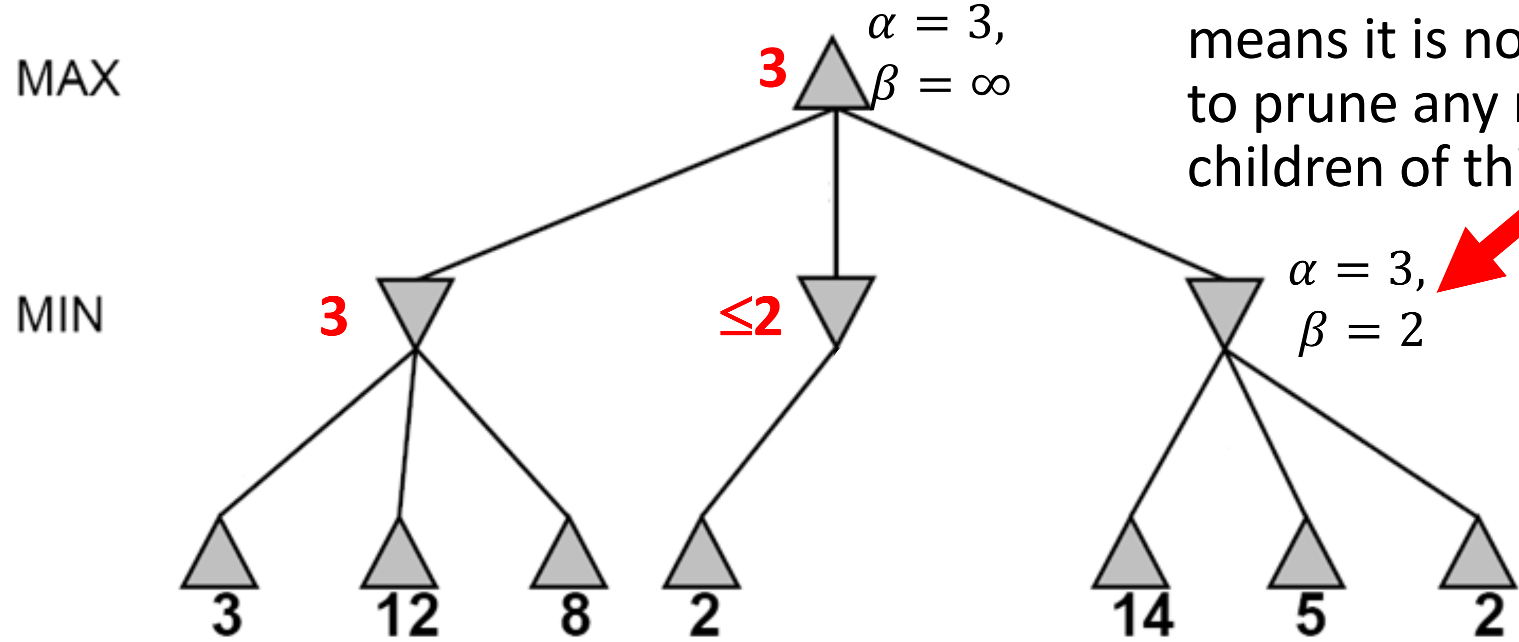
Update: a min node can update beta. A max node can update alpha.



Alpha-beta pruning

Pruning: If beta ever falls below alpha, prune any remaining children, and return.

Violation of $\alpha < \beta$ means it is now possible to prune any remaining children of this node.



Outline

- Alternating two-player zero-sum games
- Minimax search
- Limited-horizon computation and heuristic evaluation functions
- Alpha-beta search
- Computational complexity of minimax and alpha-beta

Computational complexity of alpha-beta pruning

- The basic idea of alpha-beta pruning is to reduce the complexity of minimax from $O\{b^d\}$ to $O\{b^{d/2}\}$.
- That can be done with no loss of accuracy,
- … but only if the children of any given node are optimally ordered.

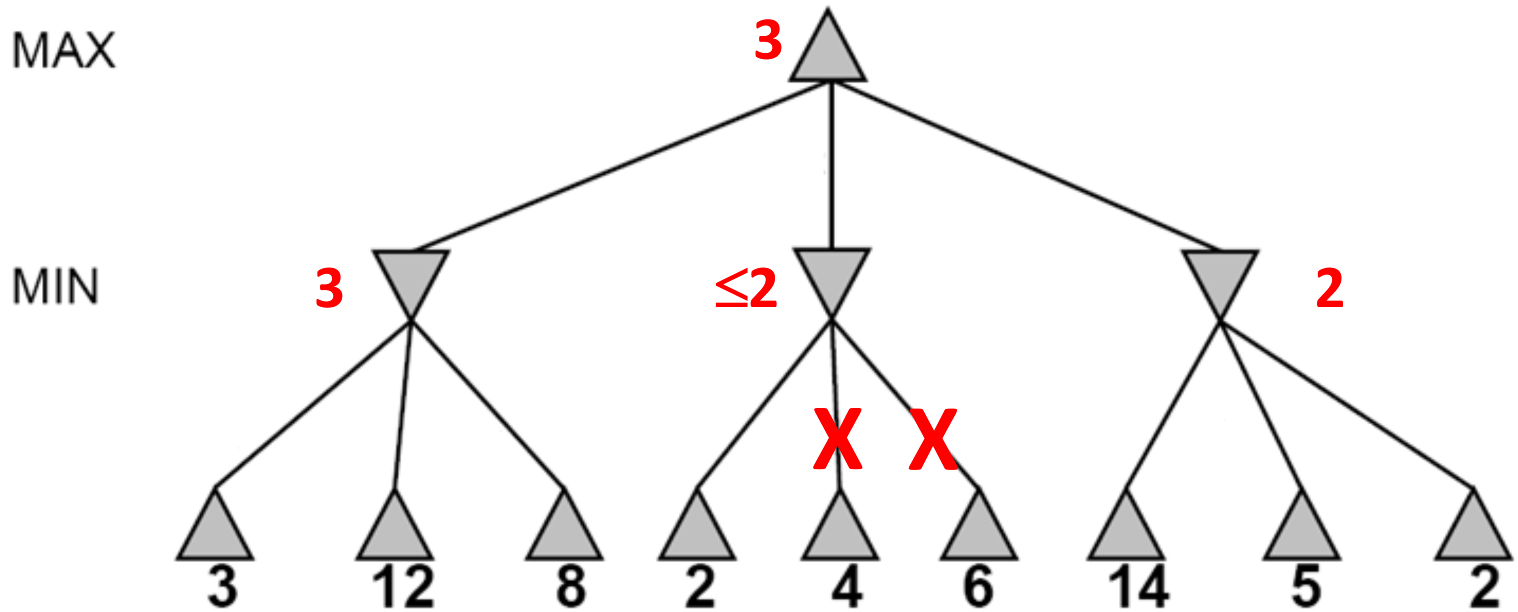
Optimal ordering

Minimum computational complexity ($O\{b^{d/2}\}$) is only achieved if:

- The children of a MAX node are evaluated, in order, starting with the highest-value child.
- The children of a MIN node are evaluated, in order, starting with the lowest-value child.

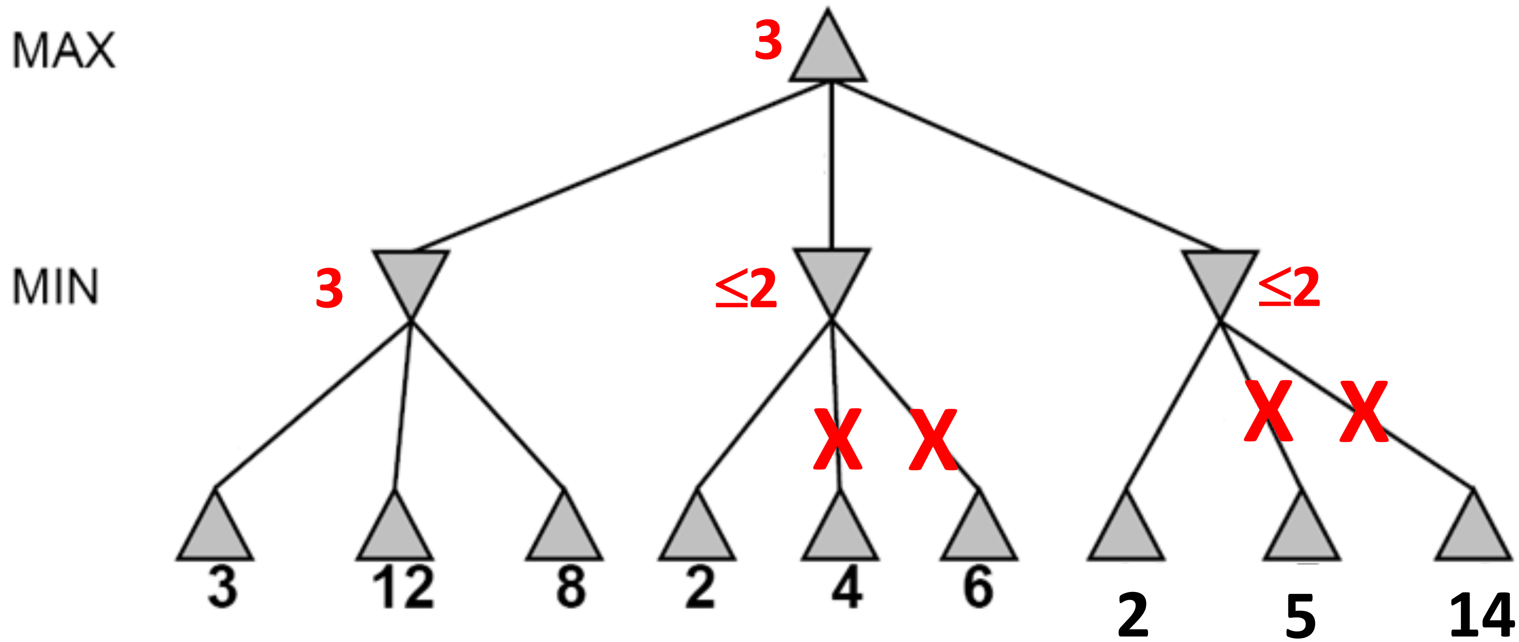
Non-optimal ordering

In this tree, the moves are not optimally ordered, so we were only able to prune two nodes.



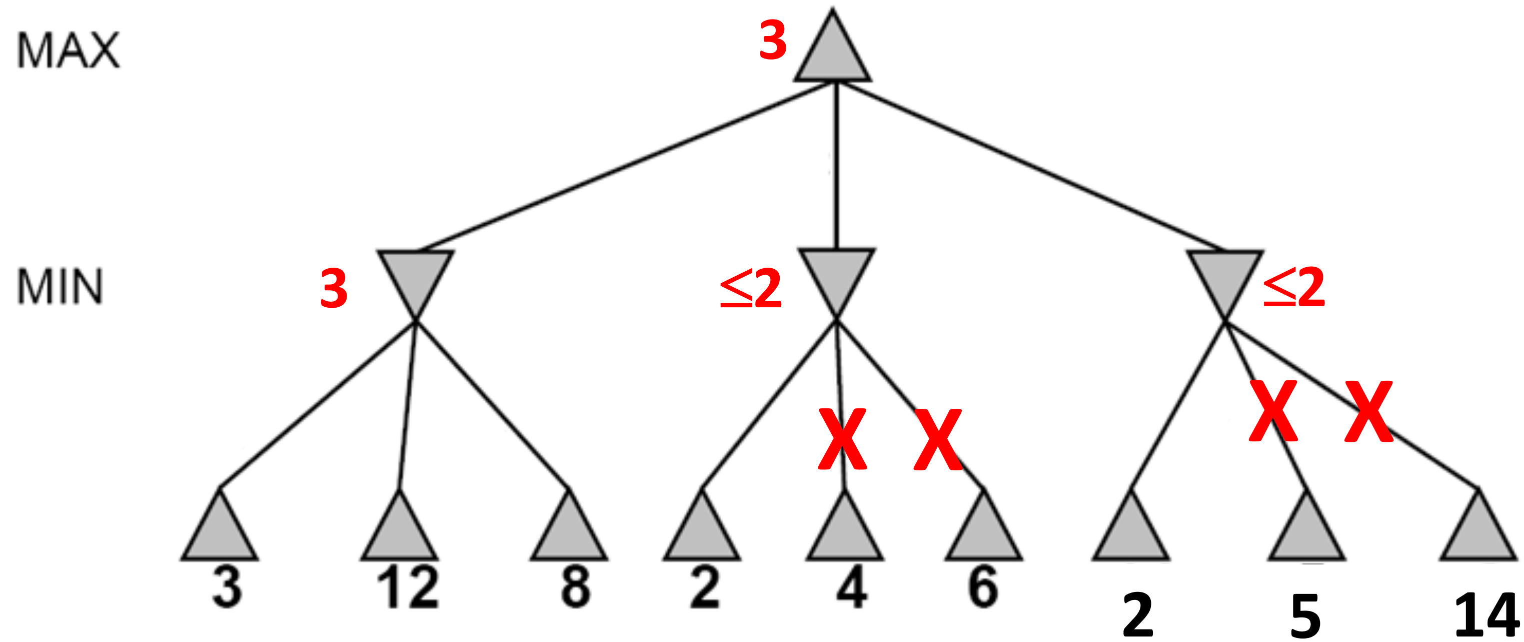
Optimal ordering

In this tree, the moves ARE optimally ordered, so we are able to prune four nodes (out of nine).

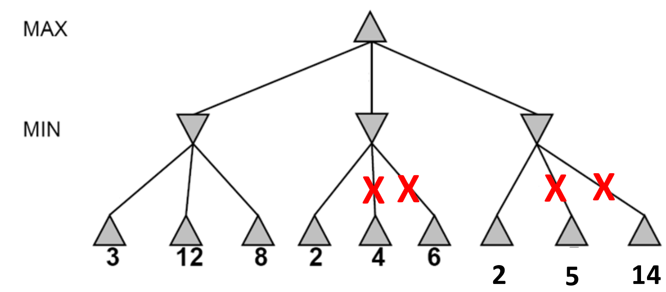


Optimal ordering

In this tree, the moves ARE optimally ordered, so we are able to prune four nodes (out of nine).



Computational Complexity

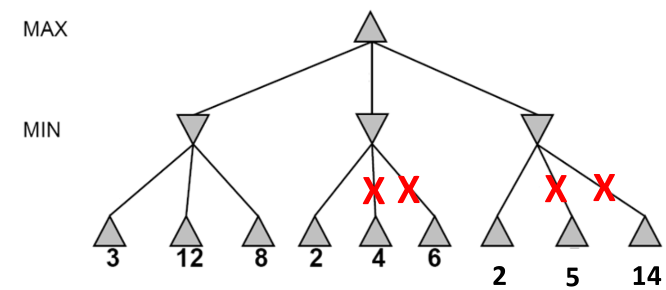


Consider a sequence of two levels, with b moves per level, and with optimal ordering.

- There are b^2 terminal nodes.
- Alpha-beta will evaluate <u>all the children of the first child</u>: b nodes.
- Alpha-beta will also evaluate <u>the first child of each non-first child</u>: $b - 1$ nodes.
- In total, alpha-beta will evaluate $2b - 1$ out of every b^2 nodes.
- For a tree of depth d, the number of nodes evaluated by alpha-beta is

$$(2b-1)^{d/2} = O\{b^{d/2}\}$$

Computational Complexity



…but wait… this means we need to know, IN ADVANCE, which move has the highest value, and which move has the lowest value!!

- Obviously, it is not possible to know the true value of a move without evaluating it.
- However, heuristics often are pretty good.
- We use the heuristic to decide which move to evaluate first.
- For games like chess, with good heuristics, complexity of alpha-beta is closer to $O\{b^{d/2}\}$ than to $O\{b^{d}\}$.

Conclusions

- Alternating two-player zero-sum games
  - $\wedge$ = a max node, $\vee$ = a min node
- Minimax search
  - $v(s) = \max_a v(\text{child}(s,a))$ or $v(s) = \min_a v(\text{child}(s,a))$
- Limited-horizon computation and heuristic evaluation functions

$$v(s) = w_1 f_1(s) + w_2 f_2(s) + \cdots$$

- Alpha-beta search
  - Min node can update beta, Max node can update alpha
  - If beta ever falls below alpha, prune the rest of the children
- Computational complexity of minimax and alpha-beta
  - Minimax is $O\{b^d\}$. With optimal move ordering, alpha-beta is $O\{b^{d/2}\}$.