

CSIT 5900

Lecture 3: Agents that Search

Department of Computer Science and Engineering
Hong Kong University of Science and Technology

Overview

- Search problems.
- Uninformed search.
- Hueristic search.
- Local search and constraint satisfaction problems.
- Game tree search.

Missionaries and cannibals problem

From wikipedia (https://en.wikipedia.org/wiki/Missionaries_and_cannibals_problem):

In the missionaries and cannibals problem, three missionaries and three cannibals must cross a river using a boat which can carry at most two people, under the constraint that, for both banks, if there are missionaries present on the bank, they cannot be outnumbered by cannibals (if they were, the cannibals would eat the missionaries). The boat cannot cross the river by itself with no people on board...

In the jealous husbands problem, the missionaries and cannibals become three married couples, with the constraint that no woman can be in the presence of another man unless her husband is also present. Under this constraint, there cannot be both women and men present on a bank with women outnumbering men...

Search in State Spaces - The 8-Puzzle

8-puzzle:

5	4	
6	1	8
7	3	2

Start State

1	2	3
8		4
7	6	5

Goal State

8-puzzle solving agent:

- Accessible environment: the agent knows exactly which state she is in.
- Four possible actions: move blank left, right, up, or down.
- Goal: find a sequence of actions that change the environment from the initial to the goal state.

Search Problem Definition

In general, a search problem consists of

- A set of states
- An initial state
- A set of deterministic actions (or operators): if an action is executable in a state, then it will produce a unique next state
- A goal test
- A path cost function

A solution is a sequence of actions that leads to a *goal state*, i.e. a state in which the goal is satisfied.

8-Puzzle - a formulation

- States: any arrangements of the blank and the numbers 1 to 8 on the board.
- Initial state: any given one.
- Goal test: the blank in the middle, and the numbers are in order clockwise starting from the left top corner.
- Actions: move the blank left, right, up, or down.
- Path cost: the length of the path.

Problem-Solving Agents

- Problem-solving agents are often goal-directed.
- Key ideas:
 - ▶ To determine what the best action is, a problem-solving agent systematically considers the expected outcomes of different possible sequences of actions that lead to some goal state.
 - ▶ A systematic search process is conducted in some representation space.
- Steps:
 - ① Goal and problem formulation
 - ② Search process
 - ③ Action execution

Searching For Solutions

Searching for a solution to a problem can be thought of as a process of building up a *search tree*:

- The root of the search tree is the node corresponding to the initial state.
- At each step, the search algorithm chooses one leaf node to expand by applying all possible actions to the state corresponding to the node.

```
function GENERAL-SEARCH(problem, strategy) returns a solution, or failure
  initialize the search tree using the initial state of problem
  loop do
    if there are no candidates for expansion then return failure
    choose a leaf node for expansion according to strategy
    if the node contains a goal state then return the corresponding solution
    else expand the node and add the resulting nodes to the search tree
  end
```


An Example Search Strategy:

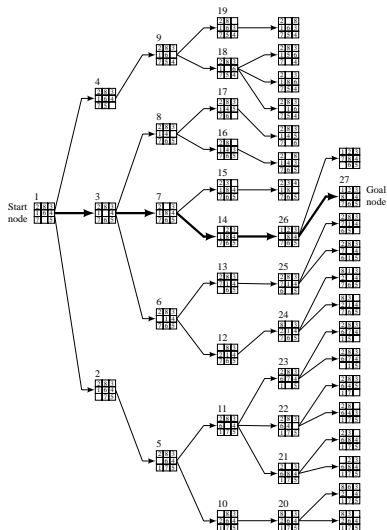
Breadth-First

- Expand first the root node.
- At each step, expand all leaf nodes.

Some properties of the breadth-first search:

- If there is a solution, then the algorithm will find it eventually.
- If there are more than one solutions, then the first one returned will be a one with the shortest length.
- The complexity (both time and space complexity) of breadth-first search is very high, making it impractical for most real world problems.

Breadth-First Search of the Eight-Puzzle

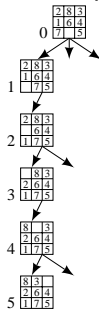


Search Strategies

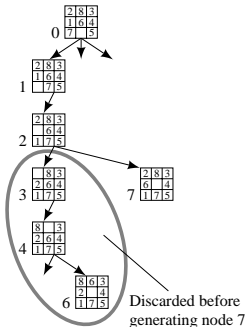
- The search strategy determines the order in which nodes in the search tree are expanded.
- Different search strategies lead to different search trees.
- Four different criteria that determine what the right search strategy is for a problem:
 - ▶ **Completeness:**
Is it guaranteed to find a solution if one exists?
 - ▶ **Time complexity:**
How long does it take to find a solution?
 - ▶ **Space complexity:**
How much memory is needed?
 - ▶ **Optimality:**
Does it find the “best” solution if several different solutions exist?
- **Types:**
 - ▶ Uninformed (or blind) search strategies
 - ▶ Informed (or heuristic) search strategies

Depth-First (Backtracking) Search

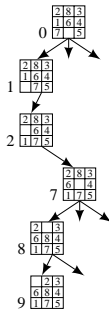
- Depth-first search generates the successors of a node one at a time;
- as soon as a successor is generated, one of its successors is generated;
- Normally, a *depth bound* is used: no successor is generated whose depth is greater than the depth bound;
- the following is a depth-first search illustration using 8 puzzle with a depth bound of 5.



(a)



(b)



(c)

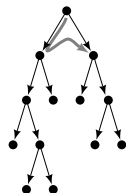
Comparing Breadth-First and Depth-First Search

In the following, b is the branching factor; d is the depth of solution; m is the depth bound of the depth first search:

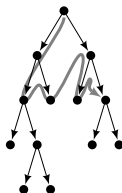
	Time	Space	Optimal?	Complete?
Breadth First	b^d	b^d	Yes	Yes
Depth First	b^m	bm	No	Yes, if $m \geq d$

Iterative Deepening Search

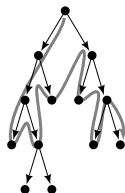
- A technique called *iterative deepening* enjoys the linear memory requirements of depth-first search while guaranteeing optimality if a goal can be found at all;
- it conducts successive depth-first search, by increasing the depth bound by 1 each time, until a goal is found:



Depth bound = 1



Depth bound = 2



Depth bound = 3



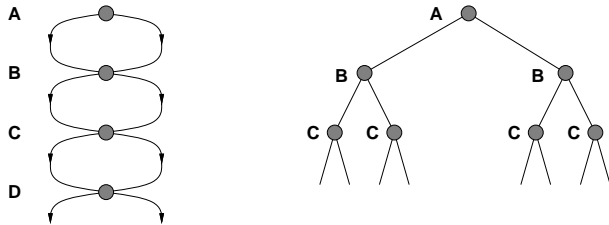
Depth bound = 4

© 1998 Morgan Kaufman Publishers

Avoiding Repeated States

- Newly expanded nodes may contain states that have already encountered before.
- There are three ways to deal with repeated states, in increasing order of effectiveness and computational overhead:
 - ▶ Do not return to the state just came from.
 - ▶ Do not create a path with cycles, that is, do not expand a node whose state has appeared in the path already.
 - ▶ Do not generate any state that was generated before.

A state space that generates an exponentially large search tree:



Heuristic Search

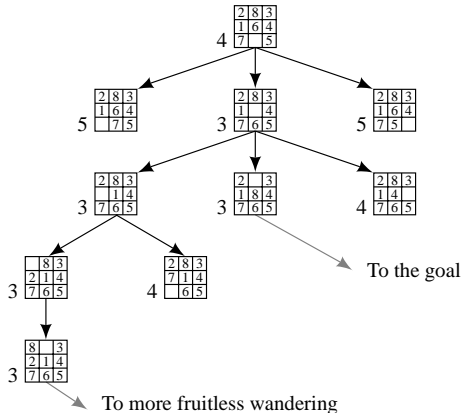
- A heuristic function is a mapping from states to numbers.
- It normally measures how far the current state is from a goal state, so the smaller the value is the closer it is from the goal.
- Heuristic or *best-first* search starts with a heuristic function and chooses nodes that have smallest value to expand.

Eight Puzzle

Consider the following heuristic function for the eight-puzzle:

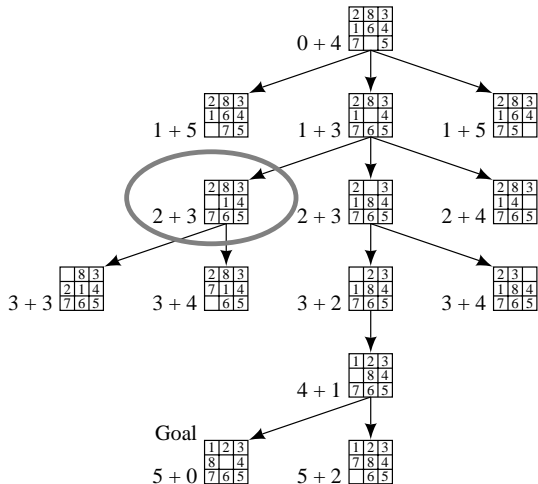
$$f(n) = \text{number of tiles out of place compared with goal}$$

Here is a search using this function (the number next to the node is its value of this function):



Eight Puzzle

The following is the search using the same heuristic function but with path cost added:



A* Search

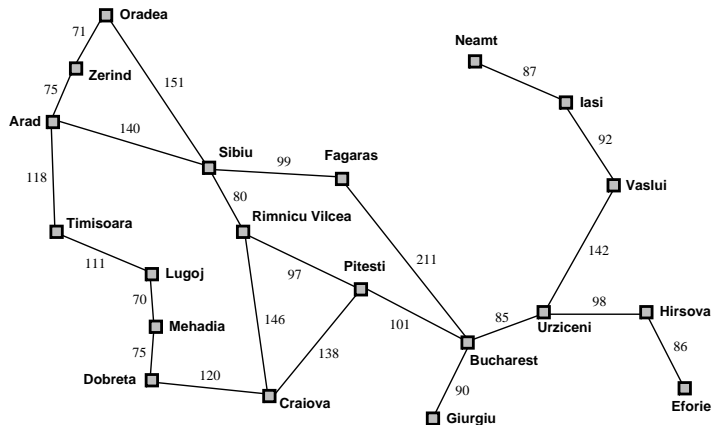
- Evaluation function:
 - ▶ Sum of:
 - ★ actual path cost $g(n)$ from the start node to node n
 - ★ estimated cost $h(n)$
 - ▶ Estimated cost of the cheapest path through node n
- Idea: Try to find the cheapest solution.

A* Search by Tree

A* search by trees - check only ancestors for repeated states

- 1 create a search tree T , consisting solely of the start node, n_0 . Put n_0 on a list called $OPEN$.
- 2 If $OPEN$ is empty, then exit with failure.
- 3 Select the first node on $OPEN$, and remove it from $OPEN$. Call this node n .
- 4 If n is a goal node, exit successfully with the solution corresponding to the path from root n_0 to this node.
- 5 Expand node n , generating the set M of its successors that are not already ancestors of n in G . Install these members of M as children of n in G , and add them to $OPEN$.
- 6 reorder the list $OPEN$ in order of increasing $g(n) + h(n)$ values. (Ties are resolved in favor of the deepest node in the search tree.)
- 7 go to step 2.

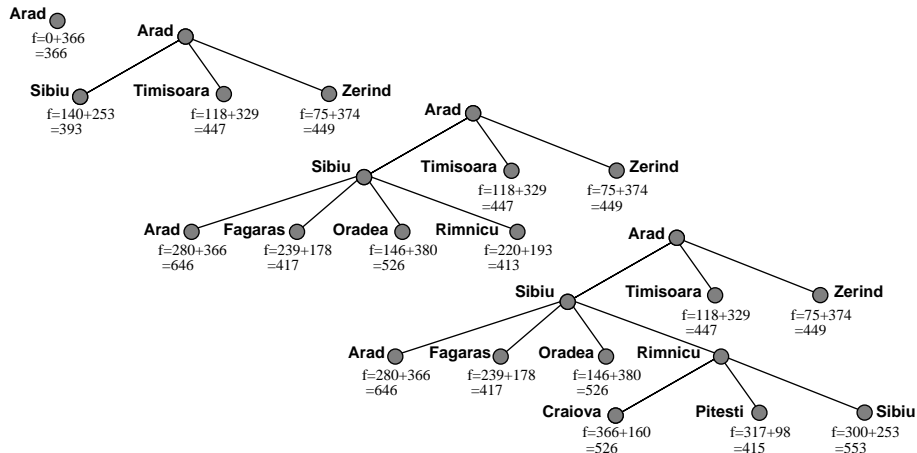
Route Finding



Straight-line distance
to Bucharest

Arad	366
Bucharest	0
Craiova	160
Dobreta	242
Eforie	161
Fagaras	178
Giurgiu	77
Hirsova	151
Iasi	226
Lugoj	244
Mehadia	241
Neamt	234
Oradea	380
Pitesti	98
Rimnicu Vilcea	193
Sibiu	253
Timisoara	329
Urziceni	80
Vaslui	199
Zerind	374

A* Search for Route Finding



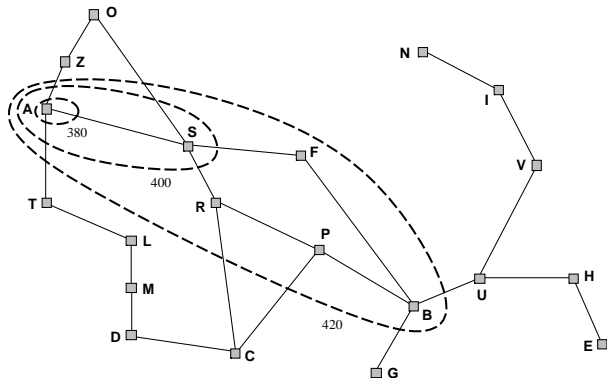
Behavior of A^*

We shall assume:

- h is *admissible*, that is, it is never larger than the actual cost.
- g is the sum of the cost of the operators along the path.
- The cost of each operator is greater than some positive amount, ϵ .
- The number of operators is finite (thus finite branching factor of search tree).

Under these conditions, we can always revise h into another admissible heuristic function so that the $f = h + g$ values along any path in the search tree is never decreasing. (*Monotonicity*.) If f^* is the cost of an optimal solution, then

- A^* expands all nodes with $f(n) < f^*$.
- A^* may expand some nodes with $f(n) = f^*$ before finding goal.

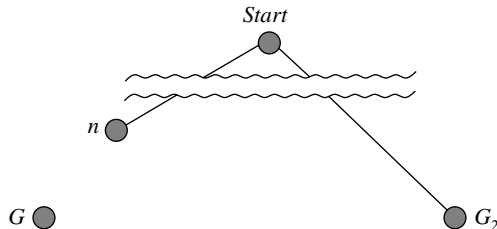


Completeness and Optimality of A^*

Under these assumptions,

- A^* is complete: A^* expands nodes in order of increasing f , there are only finite number of nodes with $f(n) \leq f^*$ is finite, so it must eventually expanded to reach a goal state.
- A^* is optimal: it will always return an optimal goal. **Proof:**
 G -optimal, G_2 -suboptimal.

It is impossible for A^* to find G_2 before G :



- It is *optimally efficient*: no other optimal algorithm is guaranteed to expand fewer nodes than A^* (Dechter and Pearl 1985).

Complexity of A^*

- Complexity:
 - ▶ Number of nodes expanded is exponential in the length of the solution.
 - ▶ All generated nodes are kept in memory. (A^* usually runs out of space long before it runs out of time.)
 - ▶ With a good heuristic, significant savings are still possible compared to uninformed search methods.
 - ▶ Admissible heuristic functions that give higher values tend to make the search more efficient.
- Memory-bounded extensions to A^* :
 - ▶ Iterative deepening A^* (IDA^*)
 - ▶ Simplified memory-bounded A^* (SMA^*)

Heuristic Functions for 8-Puzzle

- Some possible candidates (admissible heuristic functions):
 - Number of tiles that are in the wrong position (h_1)
 - Sum of the city block distances of the tiles from their goal positions (h_2)
- Comparison of iterative deepening search, A^* with h_1 , and A^* with h_2 (averaged over 100 instances of the 8-puzzle, for various solution lengths d):

d	Search Cost			Effective Branching Factor		
	IDS	$A^*(h_1)$	$A^*(h_2)$	IDS	$A^*(h_1)$	$A^*(h_2)$
2	10	6	6	2.45	1.79	1.79
4	112	13	12	2.87	1.48	1.45
6	680	20	18	2.73	1.34	1.30
8	6384	39	25	2.80	1.33	1.24
10	47127	93	39	2.79	1.38	1.22
12	364404	227	73	2.78	1.42	1.24
14	3473941	539	113	2.83	1.44	1.23
16	—	1301	211	—	1.45	1.25
18	—	3056	363	—	1.46	1.26
20	—	7276	676	—	1.47	1.27
22	—	18094	1219	—	1.48	1.28
24	—	39135	1641	—	1.48	1.26

- We see that h_2 is better than h_1 .
- In general, it is always better to use a heuristic function with higher values, as long as it is admissible, i.e. does not overestimate: h_2 *dominates* (is better than) h_1 because for all node n , $h_2(n) \geq h_1(n)$.

Inventing Heuristic Functions

Questions:

- How might one have come up with good heuristics such as h_2 for the 8-puzzle?
- Is it possible for a computer to mechanically invent such heuristics?

Relaxed Problem. Given a problem, a relaxed problem is one with less restrictions on the operators.

Strategy. use the path cost of a relaxed problem as the heuristic function for the original problem.

Inventing Heuristic Function - 8-Puzzle

Example. Suppose the 8-puzzle operators are described as:

A tile can move from square A to B if A is adjacent to B and B is blank.

We can then generate three relaxed problems by removing some of the conditions:

- (a) A tile can move from square A to B if A is adjacent to B.
- (b) A tile can move from square A to B if B is blank.
- (c) A tile can move from square A to B.

Using (a), we get h_2 . Using (c), we get h_1 .

ABSOLVER (Prieditis 1993) is a computer program that automatically generates heuristics based on this and other techniques.

Constraint Satisfaction Problem (Assignment Problem)

Problem definition:

- A finite set of variables and their domains.
- A finite set of conditions on these variables.
- A solution is an assignment to these variables that satisfies all the conditions.

Example: 8-queens problem:

- Variables: q_1, \dots, q_8 , the position of a queen in column $1, \dots, 8$.
- Domain: the same for all variables, $\{1, \dots, 8\}$.
- Constraints: $q_1 - q_2 \neq 0, |q_1 - q_2| \neq 1, \dots$

Constructive Methods

Constructive methods use search strategies, especially depth-first search, to try to find a solution:

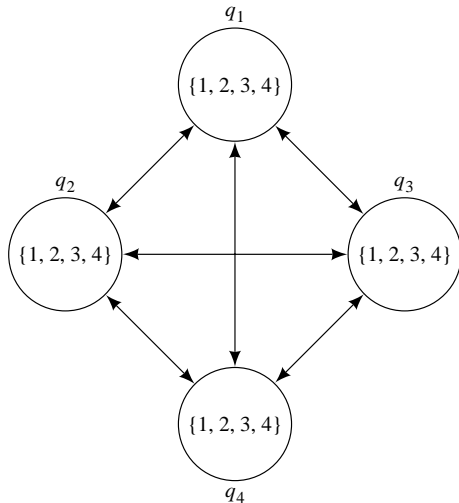
- States: any partial assignment (assign values to *some* of the variables).
- Initial state: the empty assignment.
- Operator: pick a new variable, and assign a value in its domain to it.
- Path cost: 0.
- Goal condition: when all constraints are satisfied.

Constraint Propagation

- A technique called *constraint propagation* can be used to prune the search space: given the current partial assignment, one can use the constraints to narrow down the search, i.e. eliminate some possible values of the remaining variables.
- How to do this depends on the type of constraints.
- An example using SAT (plus dependency-based backtracking and conflict-clause learning) will be given later.
- The next few slides illustrate the idea using 4-Queen Problem and constraint graphs constructed using constraints involving two variables: A *constraint graph* is one where a node is labelled by a variable and the domain that the variable can take, and there is an edge from node i to j if the variable i and j are constrained.

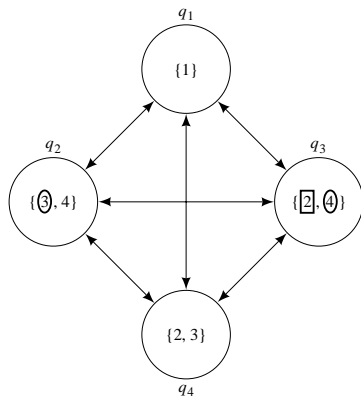
Running Example: 4-Queens Problem

Constraint graph:



Running Example: 4-Queens Problem

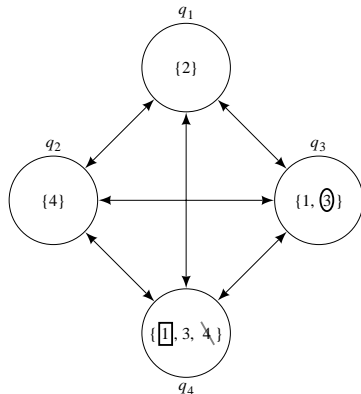
Constraint graph with $q_1 = 1$:



- ☐ Values eliminated by first making arc (q_2, q_3) and then arc (q_3, q_2) consistent
- ☐ Value eliminated by next making arc (q_3, q_4) consistent

Running Example: 4-Queens Problem

Constraint graph with $q = 2$:



- $\textcircled{}$ Value eliminated by first making arc (q_3, q_2) consistent
- $\boxed{}$ Value eliminated by next making arc (q_4, q_3) consistent
- \textbackslash Value eliminated by next making arc (q_4, q_2) consistent

Satisfiability

- A clause is a disjunction of literals - a literal is a variable or the negation of a variable.
- Satisfiability (SAT) refers to the problem of deciding if a set of clauses is satisfiable. It was the first NP-complete problem discovered by Cook, it is also one of the most intensely studied NP-complete problem.
- There is a huge literature on SAT. SAT algorithms have many applications.
- 3SAT refers to the problem of deciding if a set of clauses that have no more than 3 literals is satisfiable.
- In terms of computational complexity, SAT is equivalent to 3SAT.
- A procedure for SAT is sound if whenever it returns yes, the input is indeed satisfiable; it is complete if it will return yes whenever the input is satisfiable. Obviously, we want only sound procedures. But incomplete ones can sometimes be very useful.

DPLL Procedure

The most popular and widely studied complete method for SAT is the so called Davis-Putnam-Longemann-Loveland (DPLL) procedure. (Can you see that it's really the constructive method for CSP?)

Let α be a set of clauses. If l is a literal occurring in α , then by $\alpha(l)$, we mean the result of letting l true and simplifying: if a clause C does not mention l and $\neg l$, then do nothing; if C mention l , then eliminate it; if C mention $\neg l$, then delete $\neg l$ from C .

Procedure DPLL(CNF: α)

if α is empty, then return yes.

else if there is an empty clause in α , return no.

else if there is a pure literal l in α , then return DPLL($\alpha(l)$).

else if there is a unit clause $\{l\}$ in α , then return DPLL($\alpha(l)$).

else select a variable p in α , and do the following

if DPLL($\alpha(p)$) return yes, then return yes.

else return DPLL($\alpha(\neg p)$).

Where a literal in α is pure if it occurs only positively or negatively, and a unit clause is one whose length is 1.

Modern DPLL

The state-of-art SAT solvers use DPLL with conflict driven clause learning and dependency-based backtracking:

- ➊ Start with the empty assignment.
- ➋ If all variables are assigned then return with the assignment.
- ➌ Choose an assignment for an un-assigned variable (backtrack point).
- ➍ Propagate (deduce) the extended partial assignment.
- ➎ If propagate yields a contradiction:
 - ▶ analyze the conflict to learn some new clauses (conflict driven clause learning)
 - ▶ backtrack to the nearest backtrack point that is relevant to the conflict.
 - ▶ exit with failure (unsatisfiable) if backtrack returns no new alternatives.
- ➏ Else go back to step 2.

See [https:](https://en.wikipedia.org/wiki/Conflict-Driven-Clause-Learning)

[//en.wikipedia.org/wiki/Conflict-Driven-Clause-Learning](https://en.wikipedia.org/wiki/Conflict-Driven-Clause-Learning) for more details.

GSAT

Incomplete methods cannot be used to check unsatisfiable CNFs. But on satisfiable CNFs, incomplete methods can sometimes be much faster than the currently known best complete methods.

The following randomized local search algorithm, often called GSAT, is typical of current best known incomplete methods:

```
Procedure GSAT(CNF:  $\alpha$ , max-restart: mr, max-climb: mc)
  for i = 1 to mr do
    get a randomly generated truth assignment  $v$ .
    for j = 1 to mc do
      if  $v$  satisfies  $\alpha$ , return yes
      else let  $v$  be a random choice of one the best successors of  $v$ 
  return failure
```

where a successor of v is an assignment with the truth value of one of the variables flipped, and a best successor is one that satisfies maximum number of clauses in α .

Heuristic Repair

- GSAT is a local search method that belongs to *Heuristic repair* class of methods in CSP.
- In CSP, constructive methods starts with empty assignment, and build up a solution gradually by assigning values to variables one by one.
- Heuristic repair starts with a proposed solution, which most probably does not satisfy all constraints, and repair it until it does.
- An often used, so-called *min-conflicts* repair method by Gu will select a variable for adjust, and find a value that minimizes conflicts - see the following example on 8-Queens.

8-Queens By Min-Conflicts

2				x		
0						
x ²			x			
1						x
3	x					
2	x					
4						x
1		x				

Move
(1,3) to (1,2)

	2			x		
x	2					
	2			x		
	2					x
	3	x				
	x ³					
	1					x
1		x				

		2			x	
x		1				
		2		x		
		2				x
		x ¹				
		1				
	x	3				x
	2	x				

Move (2,6)
to (2,7)

No change

			2		x	
x			2			
			2	x		
			4			x
	x	3				
		1				
	x	2				x
		x ¹				

Unsupervised Learning (Clustering and K-Mean via CSP)

- Unsupervised learning, by its name, can be about any learning methods that do not need explicitly labeled examples.
- However, it often refers to those clustering methods that take input \mathcal{D}_t to be a set of input x without labels y .
- Reinforcement learning, while needing no explicitly labeled data, is considered on its own (samples are generated dynamically and under the control of the agent).

Example 1: English word clustering

Brown *et al* 1992:

- Input: a (large) corpus of words (100 million words of news articles)
- Output: Clusters of words, for example:
 - 1 Friday Monday Thursday Wednesday Tuesday Saturday Sunday weekends Sundays Saturdays
 - 2 June March July April January December October November September August
 - 3 people guys folks fellows CEOs chaps doubters commies unfortunates blokes
 - 4 down backwards ashore sideways southward northward overboard aloft downwards adrift
 - 5 water gas coal liquid acid sand carbon steam shale iron
 - 6 great big vast sudden mere sheer gigantic lifelong scant colossal
 - 7 man woman boy girl lawyer doctor guy farmer teacher citizen
 - 8 American Indian European Japanese German African Catholic Israeli Italian Arab

It's often called Brown clustering or IBM clustering, and used to be a corner stone of NLP systems.

Example 2: Image clustering

Feature learning using neural networks clusters 10 million images (from YouTube stills) into 22 thousands categories:



Figure: From Percy Liang's AI note

Clustering

Given a training set of inputs $\mathcal{D}_t = \{x_1, \dots, x_n\}$, classify them into a unique cluster in $\{C_1, \dots, C_K\}$.

- A CSP problem?
- Variables: c_1, \dots, c_n .
- Domains: $\{1, \dots, K\}$ ($c_i = j$ means that the i th input x_i is assigned to cluster j).
- Constraints?

Intuition: “Similar” inputs should get the same assignment.

K-means

K-means clustering: “similar” = close in distance, and each cluster i is represented by a centroid μ_i .

K-means clustering CSP:

- Variables: $c_1, \dots, c_n, \mu_1, \dots, \mu_K$.
- Domains: $\{1, \dots, K\}$ for c_i and the feature space for μ_i (i.e. the domain for $\phi(x)$).
- Constraints: minimize the following

$$\sum_{i=1}^n \|\phi(x_i) - \mu_{c_i}\|^2.$$

K-means - example

Consider \mathcal{D}_t with four examples whose feature space is one dimensional numbers:

$$\phi(\mathcal{D}_t) = \{1, 2, 10, 11\}.$$

and $K=2$, meaning we want to have two clusters. Clearly the solution is

$$c_1 = c_2 = 1, c_3 = c_4 = 2, \mu_1 = 1.5, \mu_2 = 10.5.$$

Insight: \mathbf{c} and $\boldsymbol{\mu}$: knowing one makes the other easily computed:

- Given $\mu_1 = 1.5, \mu_2 = 10.5$,

$$c_1 = \arg \min_i (1 - \mu_i)^2 = 1.$$

- Given $c_1 = c_2 = 1, c_3 = c_4 = 2$,

$$\mu_1 = \min_{\alpha} (1 - \alpha)^2 + (2 - \alpha)^2 = 1.5.$$

K-means

Algorithm - informal:

- ① Initialize μ_1, \dots, μ_K .
- ② Iterate the following for T steps:
 - ① compute the best assignment c_1, \dots, c_n for the given μ .
 - ② compute the best centroids μ_1, \dots, μ_K for the given c .

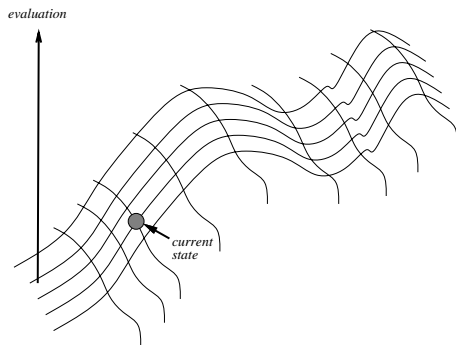
Search as Function Maximization

- Function maximization problems: find a x such that $Value(x)$ is maximal.
- Examples: 8-queens problem
 - ▶ find a state in which none of the 8 queens are attacked.
 - ▶ design a function on the states such that its value is maximal when none of the queens are attacked.
 - ▶ one such function can be: $Value(n) = 1/e^{k(n)}$, where $k(n)$ is the number of queens that are attacked.
- Another example: VLSI design - find a layout that satisfies the constraints.

Hill-Climbing

Basic ideas:

- Start at the initial state.
- At each step, move to a next state with highest value.
- It does not maintain the search tree and does not backtrack - it keeps only the current node and its evaluation.



Hill-Climbing Search Algorithm

```
function HILL-CLIMBING(problem) returns a solution state
  inputs: problem, a problem
  static: current, a node
           next, a node

  current  $\leftarrow$  MAKE-NODE(INITIAL-STATE[problem])
  loop do
    next  $\leftarrow$  a highest-valued successor of current
    if VALUE[next] < VALUE[current] then return current
    current  $\leftarrow$  next
  end
```

Simulated Annealing

- Hill-climbing can get easily stuck in a local maximum.
- One way to avoid getting stuck in a local maximum is by using *simulated annealing*.
- The main idea:
 - ▶ at each step, instead of picking the best move, it picks a random move.
 - ▶ if the move leads to a state with higher value, then execute the move.
 - ▶ otherwise, execute the move with certain probability that becomes smaller as the algorithm progresses.
 - ▶ the probability is computed according to a function (*schedule*) that maps time to probabilities.
- The idea comes from the process of annealing - cooling metal liquid gradually until it freezes.
- It has been found to be extremely effective in real applications such as factory scheduling.

Games - An Example

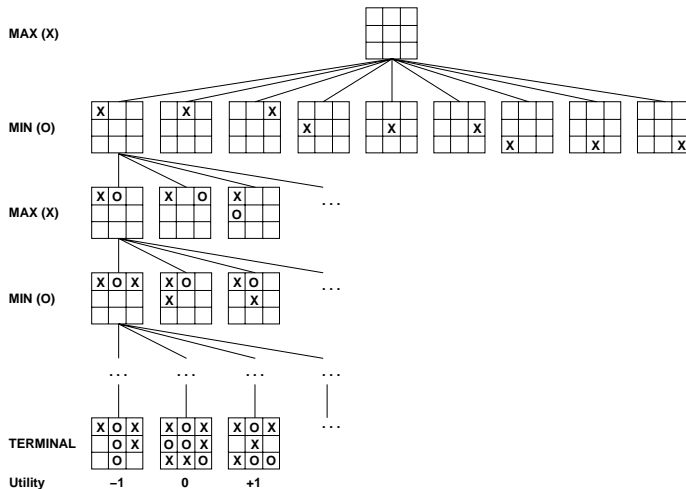
Tic-Tac-Toe:

- A board with nine squares.
- Two players: “X” and “O”; “X” moves first, and then alternate.
- At each step, the player choose an unoccupied square, and mark it with his/her name. Whoever gets three in a line wins.

Games as Search Problems

- Mainly two-player games are considered.
- Two-player game-playing problems require an agent to plan ahead in an environment in which another agent acts as its opponent.
- Two sources of uncertainty:
 - ▶ Opponent: The agent does not know in advance what action its opponent will take.
 - ▶ Time-bounded search: Time limit has to be imposed because searching for optimal decisions in large state spaces may not be practical. The “best” decisions found may be suboptimal.
- Search problem:
 - ▶ Initial state: initial board configuration and indication of who makes the first move.
 - ▶ Operators: legal moves.
 - ▶ Terminal test: determines when a terminal state is reached.
 - ▶ Utility function (payoff function): returns a numeric score to quantify the outcome of a game.

Game Tree for Tic-Tac-Toe

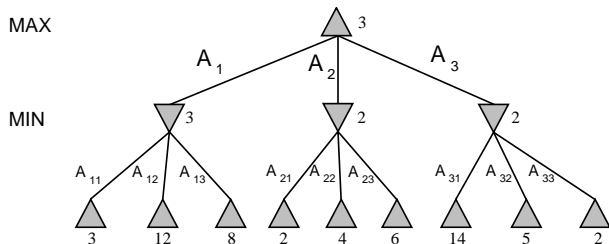


Minimax Algorithm With Perfect Decisions

Minimax Algorithm (With Perfect Decisions) Assume the two players are: MAX (self) and MIN (opponent). To evaluate a node n in a game tree:

- 1 Expand the entire tree below n .
- 2 Evaluate the terminal nodes using the given utility function.
- 3 Select a node that has not been evaluated yet, and all of its children have been evaluated. If there are no such node, then return.
- 4 If the selected node is on at which the MIN moves, assign it the minimum of the values of its children. If the selected node is on at which the MAX moves, assign it the maximum of the values of its children. Return to step 3.

MAX tries to *maximize* the utility, assuming that MIN will act to *minimize* it.



Imperfect Decisions

- Minimax algorithm with perfect decisions:
 - ▶ No time limit is imposed.
 - ▶ The complete search tree is generated.
- It is often impractical to make perfect decisions, as the time and/or space requirements for complete game tree search (to terminal states) are intractable.
- Two modifications to minimax algorithm with perfect decisions:
 - ▶ Partial tree search:
Terminal test is replaced by a *cutoff test*.
 - ▶ Evaluation function:
Utility function is replaced by a heuristic *evaluation function*.

Minimax Algorithm (With Imperfect Decisions) Assume the two players are: MAX (self) and MIN (opponent). To evaluate a node n in a game tree:

- 1 Expand the tree below n according to the partial tree search.
- 2 Evaluate the leaf nodes using the given evaluation function.
- 3 Select a node that has not been evaluated yet, and all of its children have been evaluated. If there are no such node, then return.
- 4 If the selected node is on at which the MIN moves, assign it the minimum of the values of its children. If the selected node is on at which the MAX moves, assign it the maximum of the values of its children. Return to step 3.

Evaluation Functions

- An evaluation function returns an *estimate* of the *expected utility* of the game from a given position.
- Requirements:
 - ▶ Computation of evaluation function values is efficient.
 - ▶ Evaluation function agrees with utility function on terminal states.
 - ▶ Evaluation function accurately reflects the chances of winning.
- Most game-playing programs use a *weighted linear function*:

$$w_1 f_1 + w_2 f_2 + \cdots w_n f_n$$

where the f 's are the features (e.g. number of queens in chess) of the game position, and w 's are the weights that measure the importance of the corresponding features.

- *Learning* good evaluation functions *automatically* from past experience is a promising new direction.

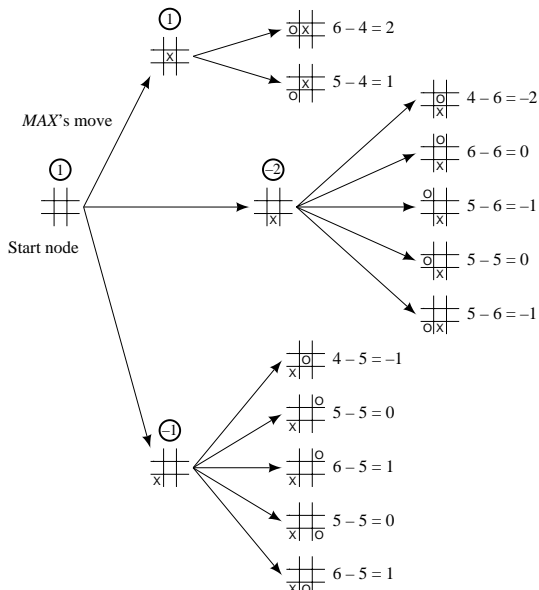
Tic-Tac-Toe

Evaluation function, $e(p)$, for Tic-Tac-Toe

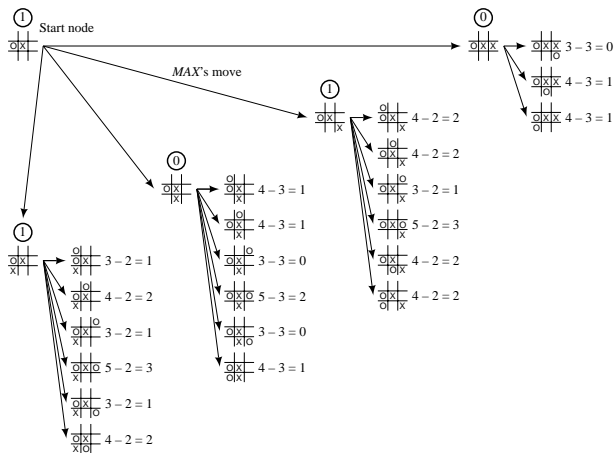
- If p is not a winning position for either player, then $e(p) = (\text{the number of complete rows, columns, or diagonals that are still open for MAX}) - (\text{the number of complete rows, columns, or diagonals that are still open for MIN})$.
- If p is a win for MAX, then $e(p) = \infty$; if p is a win for MIN, then $e(p) = -\infty$.

The next three slides illustrate Minimax search with imperfect decisions. Notice that we have eliminated symmetric positions.

The First Stage of Search

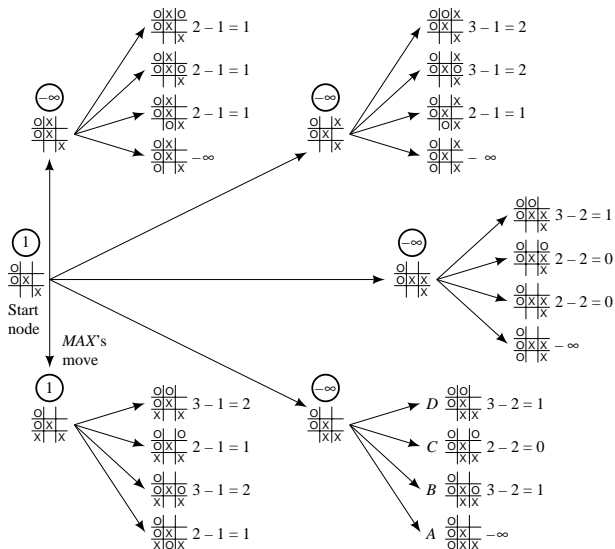


The Second Stage of Search



© 1998 Morgan Kaufman Publishers

The Last Stage of Search

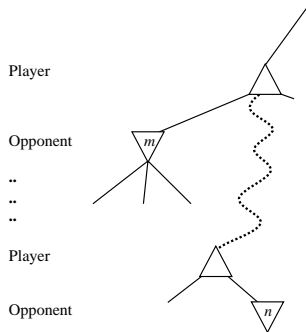


Partial Tree Search

- Different approaches:
 - ▶ Depth-limited search
 - ▶ Iterative deepening search
 - ▶ Quiescent search
- Quiescent positions are positions that are not likely to have large variations in evaluation in the near future.
- Quiescent search:
 - ▶ Expansion of nonquiescent positions until quiescent positions are reached.

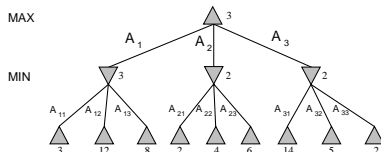
Pruning

- Pruning: The process of eliminating a branch of the search tree from consideration without actually examining it.
- General idea: If m is better than n for Player, then n will never be reached in actual play and hence can be pruned away.

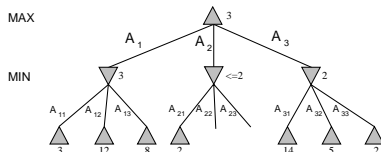


Alpha-Beta Pruning

- Minimax search without pruning:



- Minimax search with alpha-beta pruning:



- Effectiveness of alpha-beta pruning depends on the order in which successor nodes are examined.

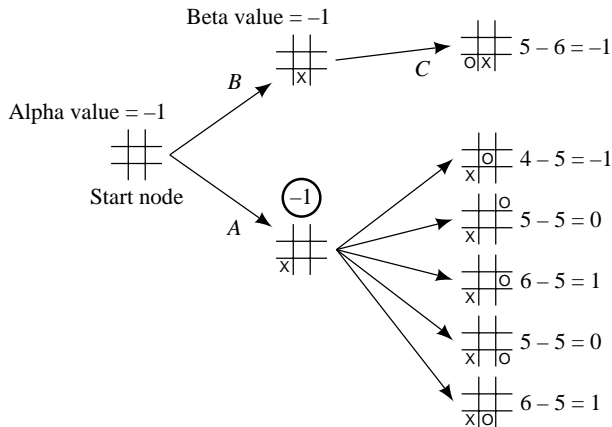
Alpha-Beta Search Algorithm - Informal Description

To evaluate a MAX node in a game tree (we shall call the value assigned to a MAX node its alpha value, and that to a MIN node its beta value):

- ➊ Expand the node depth-first until a node that satisfies the cutoff test is reached.
- ➋ Evaluate the cutoff node.
- ➌ Update the values of all the nodes that have so far been expanded according to Minimax algorithm, and using the following pruning strategy:
 - ▶ prune all children of any MIN node whose beta value is \leq the alpha value of any of its MAX ancestors.
 - ▶ prune all children of any MAX node whose alpha value is \geq the beta value of any of its MIN ancestors.
- ➍ Backtrack to a node that has not been pruned, and go back to step 1. If there are no such node to backtrack to, then return with the value assigned to the original node.

Example

Part of the first state of search in Tic-Tac-Toe using Alpha-Beta search.



Alpha-Beta Search Algorithm

The above informal description corresponds to calling the following function

MAX-VALUE(*node*, *game*, $-\infty$, ∞):

function MAX-VALUE(*state*, *game*, α , β) **returns** the minimax value of *state*

inputs: *state*, current state in game

game, game description

α , the best score for MAX along the path to *state*

β , the best score for MIN along the path to *state*

if CUTOFF-TEST(*state*) **then return** EVAL(*state*)

for each *s* **in** SUCCESSORS(*state*) **do**

$\alpha \leftarrow \text{MAX}(\alpha, \text{MIN-VALUE}(s, \text{game}, \alpha, \beta))$

if $\alpha \geq \beta$ **then return** β

end

return α

function MIN-VALUE(*state*, *game*, α , β) **returns** the minimax value of *state*

if CUTOFF-TEST(*state*) **then return** EVAL(*state*)

for each *s* **in** SUCCESSORS(*state*) **do**

$\beta \leftarrow \text{MIN}(\beta, \text{MAX-VALUE}(s, \text{game}, \alpha, \beta))$

if $\beta \leq \alpha$ **then return** α

end

return β

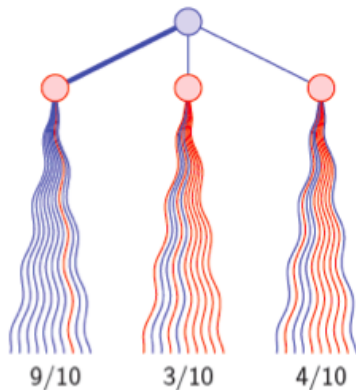
Analyses of Game-Playing Algorithm

- The effectiveness of alpha-beta pruning depends on the ordering in which successors are generated.
- Assuming optimal ordering, alpha-beta pruning can reduce the branching factor from b to \sqrt{b} (Knuth and Moore 1975).
- The result also assumes that all nodes have the same branching factor, that all paths reach the same fixed depth limit, and that the leaf (cutoff node) evaluations are randomly distributed.
- All the game-playing algorithms assume that the opponent plays optimally.

Monte-Carlo Search

Pure Monte-Carlo Search:

- Expand the current node;
- For each child, play a large number of random games to the finish and compute the average payoffs (values).
- Play the child move that has the largest average value.



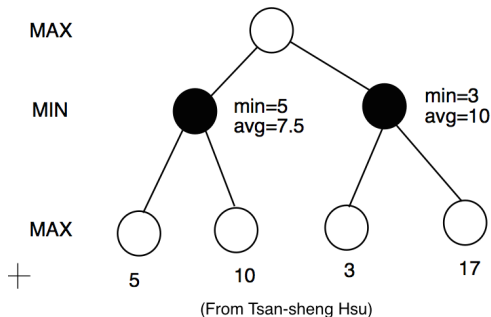
Monte-Carlo Tree Search

Making use of MCS in game playing:

- Use formulas other than average to summarize the sampling results.
- Combine it with alpha-beta pruning search: start with alpha-beta pruning and go to Monte-Carlo when the level is too deep to continue or when there is no good evaluation function on the states.
- Make use of domain knowledge, and combine it with learning techniques.

An Example

Averages may be misleading: minimax would choose the left branch while averaging would lead to the right branch.



Monte-Carlo Search with UCB

MCS with UCB (Upper Confidence Bound):

- Expand the current node v ;
- For each child M_i , play a large number of random games to the finish and compute its UCB. A common formula is:

$$UCB(M_i) = \mu_i + c \sqrt{\frac{\log N}{N_i}}$$

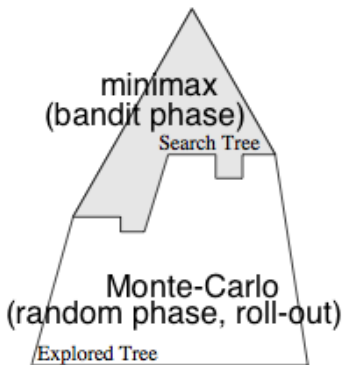
where

- ▶ μ_i is the expected value of the games (tryouts) for the child node M_i . For example, in zero-sum game, it's $\frac{W_i}{N_i}$, where W_i is the number of win's for M_i , and N_i is the total number of plays for M_i ;
 - ▶ N is the total number of plays for v ;
 - ▶ c is a constant called the *exploration* parameter used to balance branches.
- Play the child move that has the largest average value.

Monte-Carlo Tree Search (MCTS)

MCTS:

- Start building the tree using alpha-beta pruning.
- Continue on with Monte-Carlo search with UCB scoring rule.
- Propagate the values up the alpha-beta tree.



State of Art

- Chess: Deep Blue (Benjamin, Tan, *et al.*) defeated world chess champion Gary Kasparov on May 11, 1997.
- Checker: Chinook (Schaeffer *et al.*) became the world champion in 1994.
- Go: AlphaGo (Silver *et al.*) defeated a top-level player in 2016.