http://algs4.cs.princeton.edu

ALGORITHM DESIGN

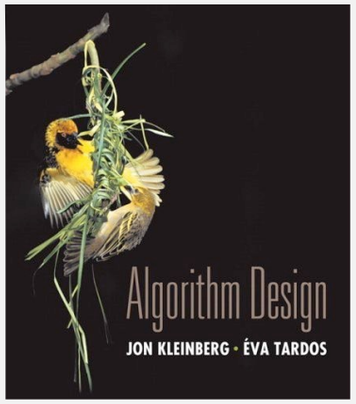
- *analysis of algorithms*
- *greedy*
- *divide-and-conquer*
- *dynamic programming*
- *network flow*
- *randomized algorithms*
- *intractability*

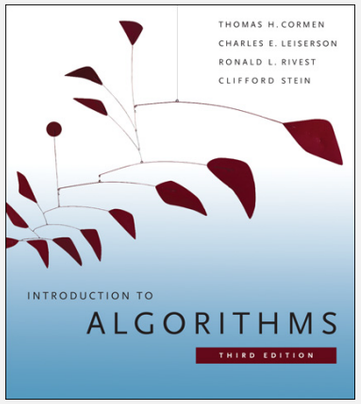
Last updated on 12/15/17 2:58 PM

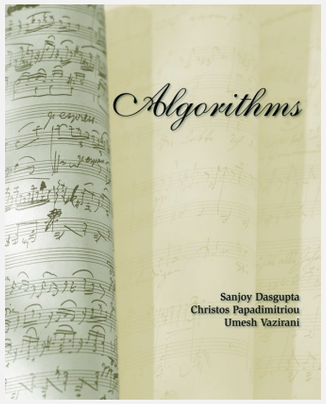
Algorithm design

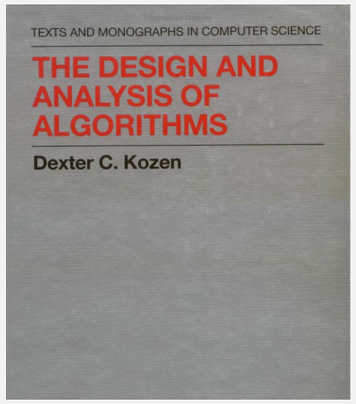
Algorithm design patterns (and antipatterns).

- Analysis of algorithms.
- Greedy.
- Divide-and-conquer.
- Dynamic programming.
- Network flow.
- Randomized algorithms.
- Intractability.



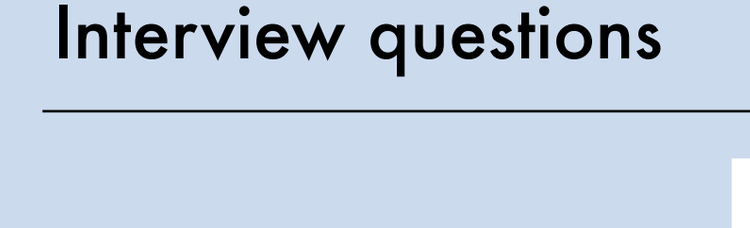






Want more? See COS 343, COS 423, COS 445, COS 451, ...

Interview questions

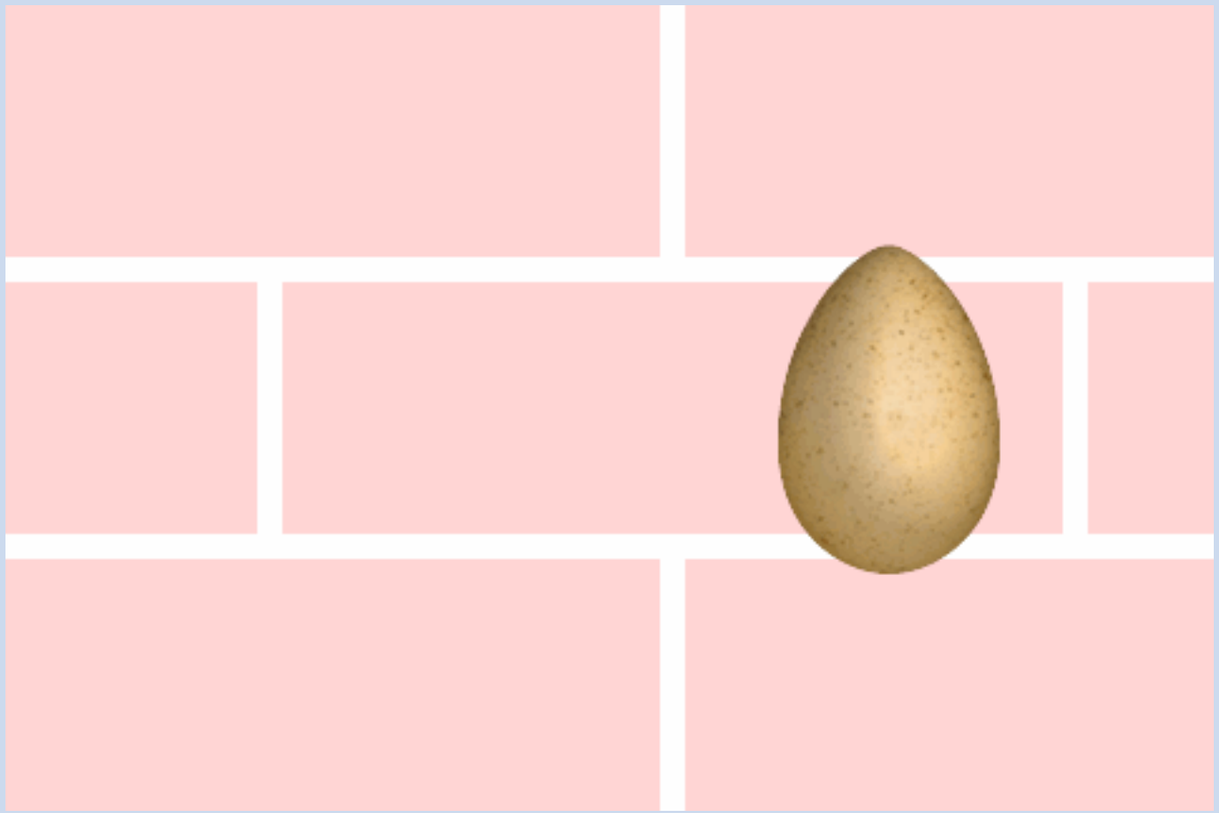


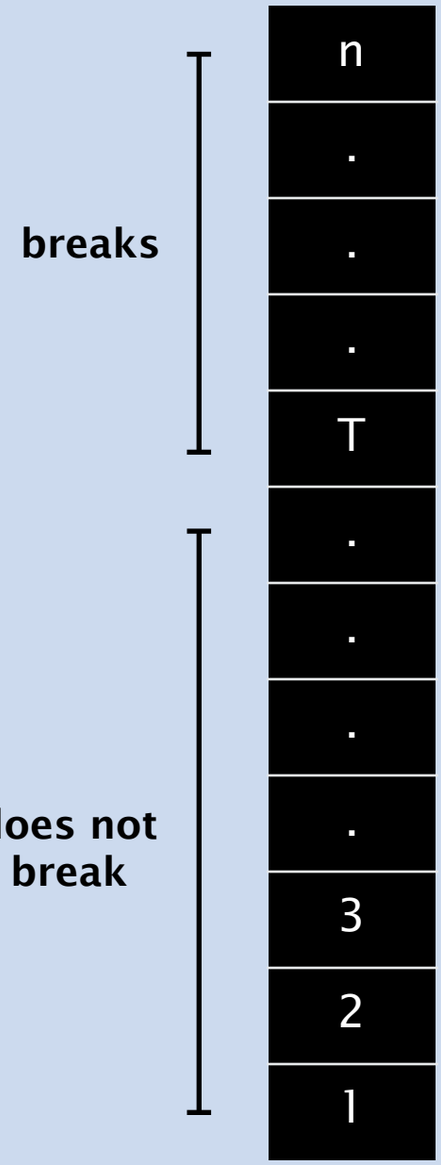
Algorithm Design

- *analysis of algorithms*
- *greedy*
- *divide-and-conquer*
- *dynamic programming*
- *network flow*
- *randomized algorithms*
- *intractability*

Egg drop

Goal. Find T using fewest number of tosses.





Egg drop

Goal. Find T using fewest number of tosses.

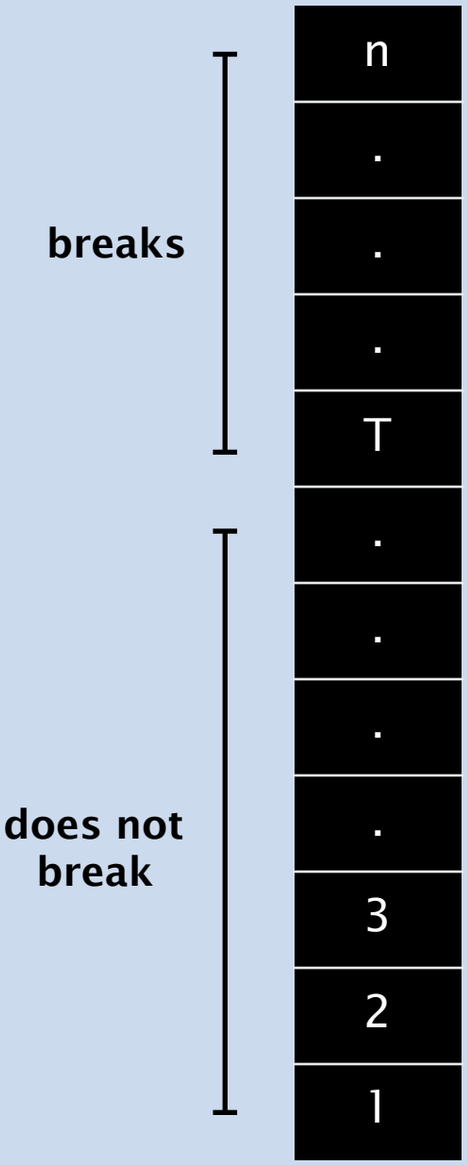
Variant 1. 1 egg.

Variant 2. ∞ eggs.

Variant 3. ∞ eggs and $\sim 2 \lg T$ tosses.

Variant 4. 2 eggs and $\leq c\ n^{1/2}$ tosses.

Variant 5. 2 eggs and $\leq c\ T^{1/2}$ tosses.



http://algs4.cs.princeton.edu

Algorithm Design

- *analysis of algorithms*
- *greedy*
- *divide-and-conquer*
- *dynamic programming*
- *network flow*
- *randomized algorithms*
- *intractability*

Greedy algorithms

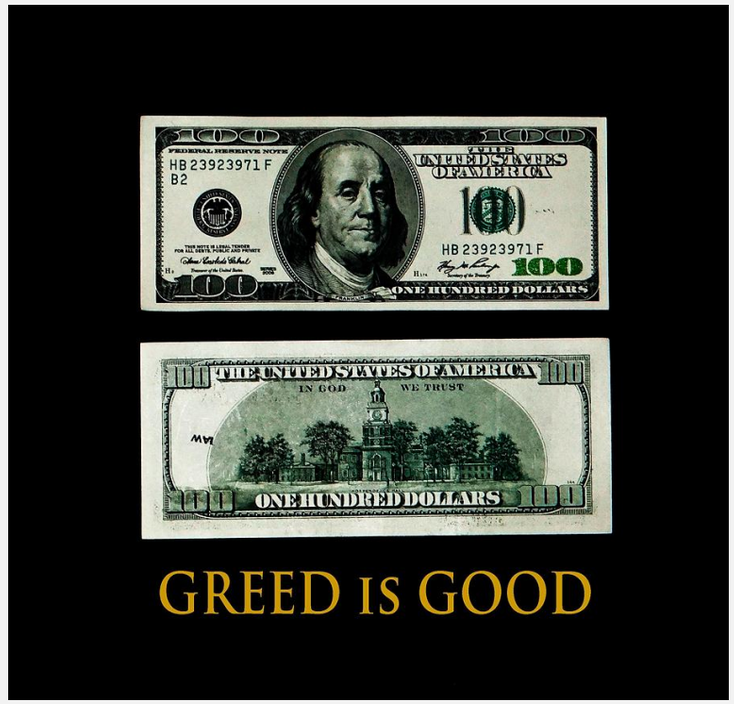
Make locally optimal choices at each step.

Familiar examples.

- Huffman coding.
- Prim's algorithm.
- Kruskal's algorithm.
- Dijkstra's algorithm.

More classic examples.

- U.S. coin changing.
- Activity scheduling.
- Gale–Shapley stable marriage.
- ...



Caveat. Greedy algorithm rarely leads to globally optimal solution.
(but is often used anyway, especially for intractable problems)

Document search

Given a document that is a sequence of n words, and a query that is a sequence of m words, find the smallest range in the document that includes the m query words (in the same order).

Ex. Query = "textbook programming computer"

> This book is intended to survey the most important computer algorithms in use today, and to teach fundamental techniques to the growing number of people in need of knowing them. It is intended for use as a **textbook for a second course in computer science, after students have acquired basic programming skills and familiarity with computer** systems. The book also may be useful for self-study or as a reference for people engaged in the development of computer systems or applications programs, since it contains implementations of useful algorithms and detailed information on performance characteristics and clients.

http://algs4.cs.princeton.edu

Algorithm Design

- *analysis of algorithms*
- *greedy*
- *divide-and-conquer*
- *dynamic programming*
- *network flow*
- *randomized algorithms*
- *intractability*

Divide and conquer

- Break up problem into two or more independent subproblems.
- Solve each subproblem recursively.
- Combine solutions to subproblems to form solution to original problem.

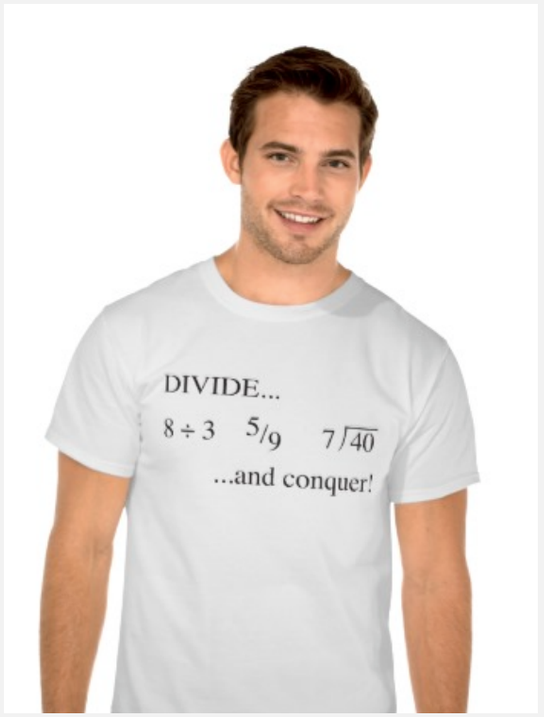
Familiar examples.

- Mergesort.
- Quicksort.

More classic examples.

- Closest pair.
- Convolution and FFT.
- Matrix multiplication.
- Integer multiplication.

  ...



needs to take COS 226?

Prototypical usage. Turn brute-force n^2 algorithm into $n \log n$ algorithm.

Personalized recommendations

Music site tries to match your song preferences with others.

- Your ranking of songs: $0, 1, \ldots, n-1$.
- My ranking of songs: $a_0, a_1, \ldots, a_{n-1}$.
- Music site consults database to find people with similar tastes.

Kendall-tau distance. Number of inversions between two rankings.

Inversion. Songs i and j are inverted if $i < j$, but $a_i > a_j$.

	A	B	C	D	E	F	G	H
you	0	1	2	3	4	5	6	7
me	0	2	3	1	4	5	7	6

3 inversions: 2-1, 3-1, 7-6

Counting inversions

Problem. Given a permutation of length n, count the number of inversions.

0	2	3	1	4	5	7	6

3 inversions: 2-1, 3-1, 7-6

Brute-force n^2 algorithm. For each $i < j$, check if $a_i > a_j$.

A bit better. Run insertion sort; return number of exchanges.

Goal. $n \log n$ time (or better).

http://algs4.cs.princeton.edu

Algorithm Design

- *analysis of algorithms*
- *greedy*
- *divide-and-conquer*
- *dynamic programming*
- *network flow*
- *randomized algorithms*
- *intractability*

Dynamic programming

- Break up problem into a series of overlapping subproblems.
- Build up solutions to larger and larger subproblems.
  (caching solutions to subproblems in a table for later reuse)

Familiar examples.

- Shortest paths in DAGs.
- Seam carving.
- Bellman–Ford.

THE THEORY OF DYNAMIC PROGRAMMING

RICHARD BELLMAN

More classic examples.

- Unix diff.
- Viterbi algorithm for hidden Markov models.
- Smith–Waterman for DNA sequence alignment.
- CKY algorithm for parsing context-free grammars.

  ...

House coloring problem

Goal. Paint a row of n houses red, green, or blue so that

- No two adjacent houses have the same color.
- Minimize total cost, where $cost(i, color)$ is cost to paint i given color.

	A	B	C	D	E	F
	7	6	7	8	9	20
	3	8	9	22	12	8
	16	10	4	2	5	7

cost to paint house i the given color

http://algs4.cs.princeton.edu

Algorithm Design

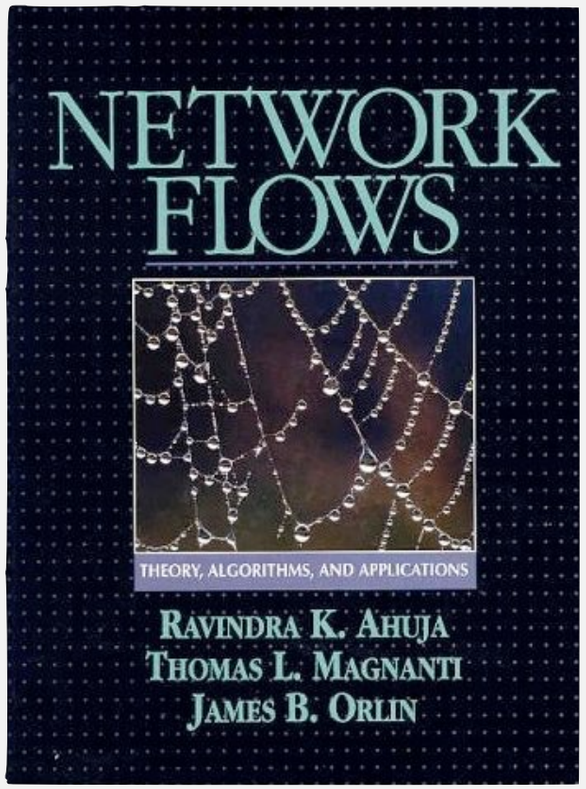
- *analysis of algorithms*
- *greedy*
- *divide-and-conquer*
- *dynamic programming*
- *network flow*
- *randomized algorithms*
- *intractability*

Network flow

Classic problems on edge-weighted graphs.

Familiar examples.

- Shortest paths.
- Bipartite matching.
- Maxflow and mincut.
- Minimum spanning tree.



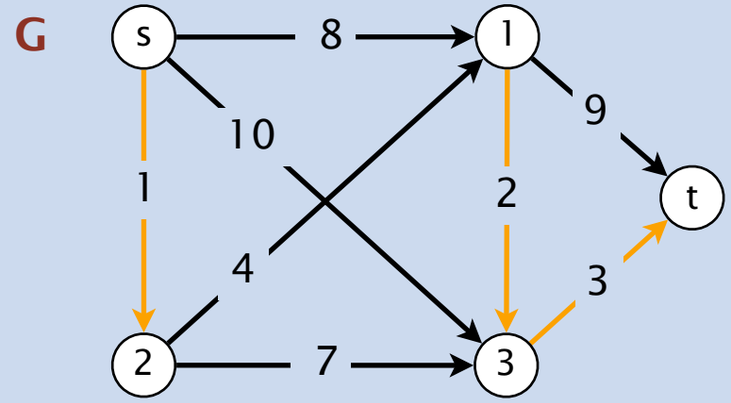
Other classic examples.

- Minimum-cost arborescence.
- Non-bipartite matching.
- Assignment problem.
- Minimum-cost flow.
- ...

Applications. Many many problems can be modeled using network flow.

Shortest path with orange and black edges

Goal. Given a digraph, where each edge has a positive weight and is orange or black, find shortest path from s to t that uses at most k orange edges.



k = 0: s→1→t (17)

k = 1: s→3→t (13)

k = 2: s→2→3→t (11)

k = 3: s→2→1→3→t (10)

http://algs4.cs.princeton.edu

Algorithm Design

- *analysis of algorithms*
- *greedy*
- *divide-and-conquer*
- *dynamic programming*
- *network flow*
- *randomized algorithms*
- *intractability*

Randomized algorithms

Algorithm that uses random coin flips to guide its behavior.

Familiar examples.

- Quicksort.
- Quickselect.

More classic examples.

- Rabin–Karp substring search.
- Miller–Rabin primality testing.
- Polynomial identity testing.
- Volume of convex body.
- Universal hashing.
- Global min cut.

...

Nuts and bolts

Problem. A disorganized carpenter has a mixed pile of n nuts and n bolts.

- The goal is to find the corresponding pairs of nuts and bolts.
- Each nut fits exactly one bolt and each bolt fits exactly one nut.
- By fitting a nut and a bolt together, the carpenter can see which one is bigger (but cannot directly compare either two nuts or two bolts).

Brute-force n^2 solution. Compare each bolt to each nut.

Challenge. Design an $n \log n$ algorithm.

http://algs4.cs.princeton.edu

Algorithm Design

- *analysis of algorithms*
- *greedy*
- *divide-and-conquer*
- *dynamic programming*
- *network flow*
- *randomized algorithms*
- *intractability*

NP-completeness

Fundamental barrier to designing efficient algorithms.

Familiar examples.

- 3-SATISFIABILITY.
- INTEGER-LINEAR-PROGRAMMING.
- TRAVELING-SALESPERSON-PROBLEM.

More classic examples.

- PLANAR-MAP-3-COLOR.
- REGISTER-ALLOCATION.
- PROTEIN-FOLDING.
- HAMILTON-PATH.
- KNAPSACK.
- 3D-ISING.
- ...

Exhaustive search

Exhaustive search. Iterate over all elements of a search space.

Applicability. Huge range of problems (including intractable ones).

Caveat. Search space is typically exponential in size of input.

Backtracking. Systematic method to iterate over elements of a search space, pruning useless ones to save time.

Sudoku

Goal. Fill 9-by-9 grid so that every row, column, and box contains each of the digits 1 through 9.

7		8				3		
			2		1			
5								
	4						2	6
3				8				
			1				9	
	9		6					4
				7		5		

Sudoku

Goal. Fill 9-by-9 grid so that every row, column, and box contains each of the digits 1 through 9.

7	2	8	9	4	6	3	1	5
9	3	4	2	5	1	6	7	8
5	1	6	7	3	8	2	4	9
1	4	7	5	9	3	8	2	6
3	6	9	4	8	2	1	5	7
8	5	2	1	6	7	4	9	3
2	9	3	6	1	5	7	8	4
4	8	1	3	7	9	5	6	2
6	7	5	8	2	4	9	3	1

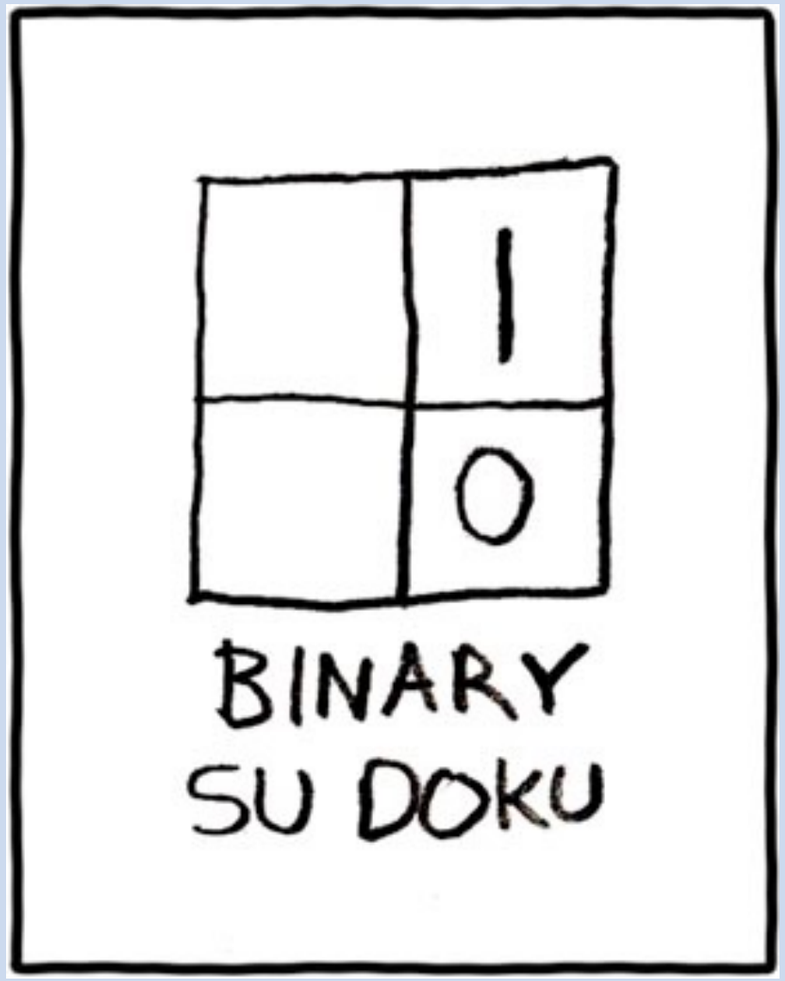
" Sudoku is a denial of service attack on human intellect. "

— Ben Laurie (founding director of Apache Software Foundation)

Sudoku is (probably) intractable

Remark. Natural generalization of Sudoku is **NP**-complete.

2-by-2 Sudoku



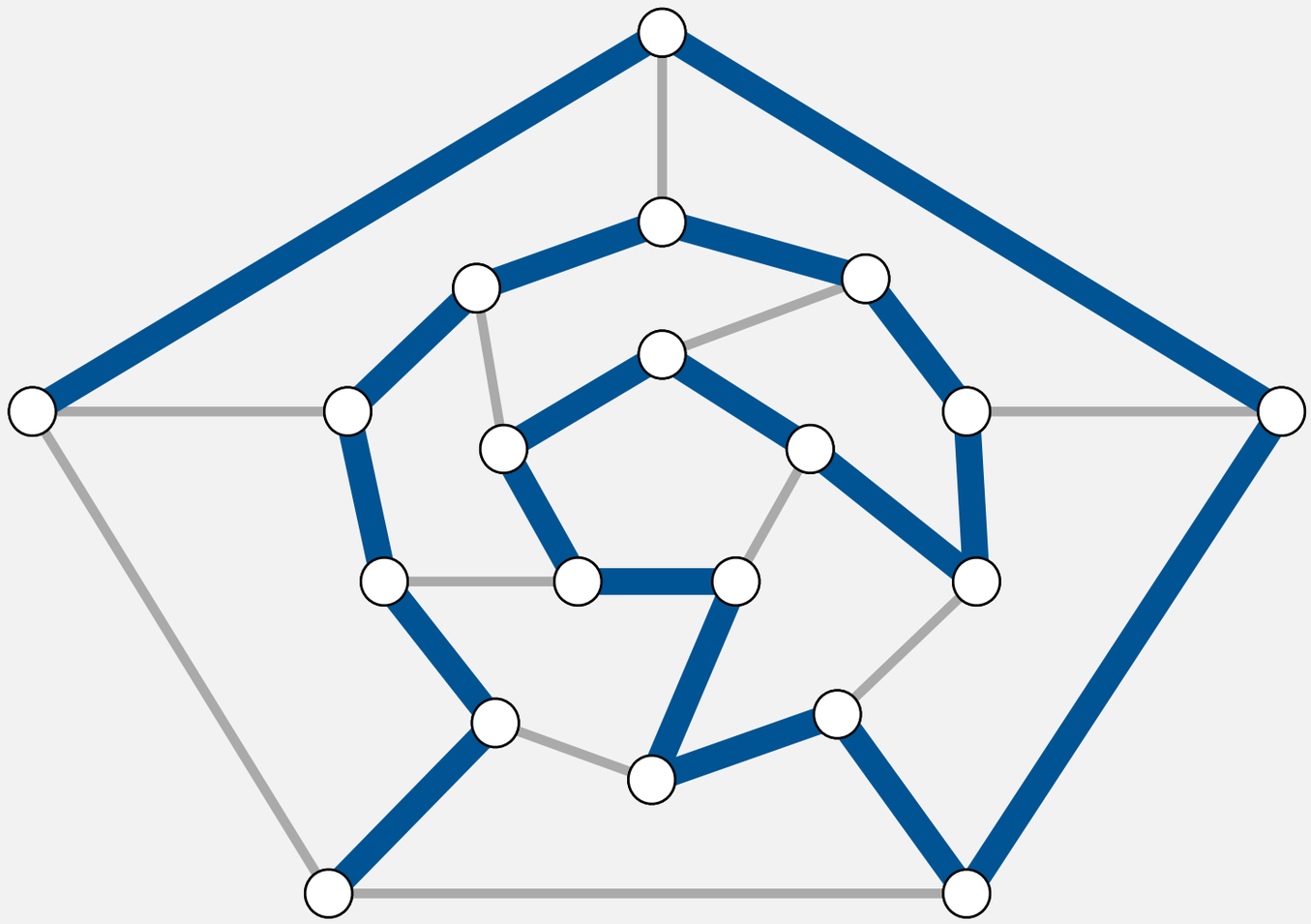
http://xkcd.com/74

25-by-25 Sudoku

J	4	N									C		B	2	M	P				E		H		O
H	D		O		6						8			1	A	B	G	C	E	5	L			F
	8		I		A	K	O	3	B	M		L	F	5	1		H	7		C			6	J
B			A				G	L		N	J		H	6	8					D	M	1	2	7
	L	1	5		M		4	2	N			P				D	J		6	9	B	8	A	
F	H		N	O	4	5				D				M	J		I			6		9	C	8
5				M		6	F							K	9	A	C				1		L	
	1				I	2		J	K		7		A	B					N		H	O		
6	A		E	G	9			C		L			O		2	5	7	1	8	F		J	K	M
I	J			K	D	L					1				E	G		3	H				B	5
M	5	3	L	7	N	A	C	I			F	B		G			K	E			O	2	J	H
	F				B	G		O		1	9			E		7		L	5	K	D	6		
K							1			5	O	H			6			9		N				
D	G					J	5	H	3			K	P		B			N		1	C	E	8	
1		C		B	7	F	6	K	D	2		M		N			4		J				5	9
L	I			5			A	E		B		1	7		F		N	J				C		D
8	6	A	H						C	O						I					F	5	7	
3	C	B	1				L		F	9				A	4				7	8	2	N		6
		E	G			7		1	5	C			L			2				H				K
		F			O						H	J		4	C				D	3	E	I	1	L
	N	6	F	H					M	E	K	3				9	P					G	O	2
G	O	5	3	C	P		E	8		F		6							4	B	J	7		I
	9	I	D	8	L	B		6		G			4	H	5	J			C	A		F		1
		J		1	G			F	7				5	9	N	L		2	A		6			C
	B				C			9				A			G		8					K	D	E

Hamilton path

Goal. Does there exist a path that visits every vertex exactly once?



visit every edge exactly once

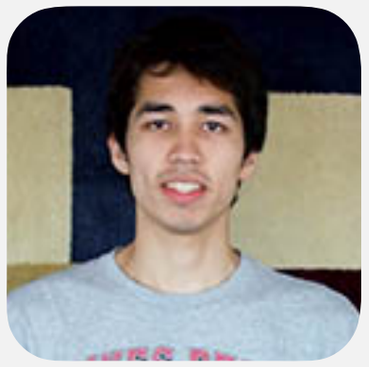
Remark. Linear-time algorithm for EULER-PATH; HAMILTON-PATH is **NP**-complete.

Credits

Faculty lead preceptors and graduate student AIs.







Undergraduate graders and lab TAs. Apply to be one next semester!

Ed tech. Most developed by undergrads!

iClicker

piazza

Princeton Salon

gradescope

codePost

Quizzera

“ Algorithms and data structures are love.

Algorithms and data structures are life. ”

— anonymous COS 226 student