15-451/651 Spring 18

Lecture 27: Review session

Caveats:

This slideshow has only some points of some topics, please read over lecture notes and materials which cover all topics. All topics (except voted-off lecture) are candidates for final, whether they are in these slides or not. Do not try to guess the “important” topics from the length of slides for each topic: the slide lengths depend on random factors like which one of us made them, and when. :)
Advanced core
Flows. LPs.

Complexity
NP-completeness. 🙁
Approx. 😊

Core
Upper/Lower bounds
Amortized Analysis
Hashing/fingerprinting.
DPs.

Uncertainty
Competitive analysis
Expert problems
Gradient Descent

Strategic Behavior
Basic game theory
Truthful mechanisms
Vickery auctions/VCG

Big Data
Compressing graphs
Streaming/Sketching
Quality/approx. tradeoff

Algebraic Algos
Polynomials and
error correction
Lecture 1 and 2: Upper & Lower Bounds

• Linear time selection
  • solving recurrences, linearity of expectation

• Tight lower bounds for sorting, max-finding, sorting-by-swaps
  • adversary method, comparison-tree method

Write a program which takes as input a list of $n$ distinct numbers $a_1, a_2, \ldots, a_n$, and outputs the elements of ranks $1, 2, 4 = 2^2, 8 = 2^3, 2^4, \ldots, 2^k$, where $2^k$ is the greatest power of 2 that is at most $n$.

Your algorithm should have $O(n)$ running time (deterministic worst-case or randomized expected). Please include a comment at the start of your program explaining your
Lecture 1 and 2: Upper & Lower Bounds

• Linear time selection
  • solving recurrences, linearity of expectation

• Tight lower bounds for sorting, max-finding, sorting-by-swaps
  • adversary method, comparison-tree method

(b) You are given $n$ unsorted distinct elements. Let $k = \lceil n^{0.9} \rceil$. Circle the tightest upper bound you can give on the number of comparisons to output the elements of ranks $1, 2, \ldots, k$ in that order. (Rank $1 =$ smallest.)

\[ \Theta(\log n) \quad \Theta(n^{0.9}) \quad \Theta(n^{0.9} \log n) \quad \Theta(n) \quad \Theta(n \log n) \quad \Theta(n^{1.9}) \]
• Theorem: Any deterministic comparison-based sorting algorithm must perform $\lg(n!)$ comparisons to sort $n$ elements in the worst case, i.e., for any sorting algorithm $A$ and $n \geq 2$, there is an input $I$ of size $n$ so that $A$ makes $\geq \lg(n!) = \Omega(n \log n)$ comparisons to sort $I$.

• Proof: Suppose there is a problem with $M$ possible outputs
  • For sorting $M = n!$ since for each output permutation $\pi$, there is an input with output $\pi$
  • Suppose for each possible output, there is an input for which that output is the only correct answer
    • For sorting there are inputs for which $\pi$ is the only correct answer

• Then there is a lower bound of $\lg M$
  • Consider a set of inputs in one-to-one correspondence with the $M$ possible outputs. Algorithm needs to find which of the $M$ outputs is correct for a given input, and each comparison can be answered in a way that removes at most half of the possible inputs
Lecture 3-4: Amortized Analysis & Union-Find

• Binary counter, maintaining array under inserts and deletes
  • amortized analysis
  • banker’s method and potential function (aka total money in bank)
  • “when is the next expensive operation coming, save enough to pay for it”

Consider a stack data structure that supports the following API:

- `top()`: Return the top of the stack. (cost = 1)
- `push(x)`: Put x on top of the stack. (cost = 1)
- `multipop(k)`: Pop and discard the top k elements of the stack. (cost = k)

With a potential of $\Phi(\text{stack}) = \boxed{}$, it is possible to prove the following bounds. All blanks must contain numerical constants such that if any one of them is made smaller the result will be false.

- amortized cost of `top()` $\leq \boxed{}$
- amortized cost of `push(x)` $\leq \boxed{}$
- amortized cost of `multipop(k)` $\leq \boxed{}$
Lecture 5: Hashing

• $h: U \rightarrow \{0, 1, 2, \ldots, M-1\}$ is universal if for all $x \neq y$,
  $$\Pr_{h \leftarrow H}[h(x) = h(y)] \leq \frac{1}{M}$$

• **Theorem:** If $H$ is universal, then for any set $S \subseteq U$ with $|S| = N$, for any $x \in S$, if we choose $h$ at random from $H$, the expected number of collisions between $x$ and other elements in $S$ is less than $N/M$.
  • $Ax \mod 2$ is universal

• **Definition:** A hash function family $H$ is $k$-universal if for every set of $k$ distinct keys $x_1, \ldots, x_k$ and every set of $k$ values $v_1, \ldots, v_k \in \{0, 1, \ldots, M-1\}$,
  $$\Pr[h(x_1) = v_1 \text{ AND } h(x_2) = v_2 \text{ AND } \ldots \text{ AND } h(x_k) = v_k] = \frac{1}{M^k}$$

• **Perfect Hashing:** can use $O(N)$ space to achieve constant worst-case lookup time to store a set of $N$ keys. Based on 2-level hashing scheme
Lecture 6: Data Stream Model

- Stream of elements $a_1, ..., a_t$

- $S_t$ is the multiset of items at time $t$, so $S_0 = \emptyset$, $S_1 = \{a_1\}$, ..., $S_i = \{a_1, ..., a_i\}$, $\text{count}_t(e) = \{i \in \{1, 2, ..., t\} \text{ such that } a_i = e\}$

- $e \in \Sigma$ is an $\epsilon$-heavy hitter at time $t$ if $\text{count}_t(e) > \epsilon \cdot t$

- Maintain $1/\epsilon$ identities and counters. Every time we discard an update, we decrement $1/\epsilon$ counts, so an $\epsilon$-heavy hitter will survive
Lecture 6: Data Stream Model

- Heavy hitters with stream deletions
- Query "What is count_t(e)?", should output est_t(e) with:
  \[ \Pr[|est_t(e) - count_t(e)| \leq \epsilon |S_t|] \geq 1 - \delta \]
- Want space close to our previous \(O(1/\epsilon)(\log(\Sigma) + \log t)\) bits
- Let \(h: \Sigma \rightarrow \{0,1,2, \ldots, 10/\epsilon\}\) be a 2-universal hash function
- Maintain an array \(A[0, 1, \ldots, k-1]\) to store non-negative integers

  when update \(a_t\) arrives:
  - if \(a_t = (\text{add}, e)\) then \(A[h(e)]++\)
  - else \(a_t = (\text{del}, e)\), and \(A[h(e)]--\)

- \(est_t(e) = A[h(e)]\)
In the string-matching problem, we have

- A text T of length m
- A pattern P of length n

**Goal:** output all occurrences of the pattern P inside the text T

Karp-Rabin: \( h_p(x) = x \mod p \) for \( x \in \{0,1\}^n \), think of \( x \) as an integer

Create \( x' \) by dropping the most significant bit of \( x \), and appending a bit to the right

- E.g., if \( x = 0011001 \), then \( x' \) could be 0110010 or 0110011

Given \( h_p(x) = z \), can we compute \( h_p(x') \) quickly?

Suppose \( x'_{lb} \) is the lowest-order bit of \( x' \), and \( x_{hb} \) is the highest order bit of \( x \)

\[ x' = 2(x - x_{hb} \cdot 2^{n-1}) + x_{lb} \]

Since \( h_p(a + b) = (h_p(a) + h_p(b)) \mod p \), and \( h_p(2a) = 2h_p(a) \mod p \),

\[ h_p(x') = (2h_p(x) - x_{hb} \cdot h_p(2^n) + x'_{lb}) \mod p \]

Given \( h_p(x) \) and \( h_p(2^n) \), this is just \( O(1) \) arithmetic operations mod p. \( O(m+n) \) total time.
Lecture 8: Dynamic Programming

• Store solutions to subproblems, avoid resolving repeatedly
• Top-down (via memorization), Bottom-up (via filling table)
  • standard table DP: longest common subsequence, knapsack
  • tree DP (for independent set), subset DP (for TSP)
Lecture 8: Dynamic Programming

(c) Let $P(n, d)$ be the number of properly parenthesized expressions of length $n$, where the maximum depth of nesting is at most $d$. You will write a dynamic program to compute this number.

**Example:** The string $()()()$ has length 8 and depth 1, $(())()$ has length 10 depth 2, and $(()()())$ has length 16 and depth 4.

(a) Fill in the blanks in the following recurrence for $P(n, d)$ (where $n \geq 0$ and $d \geq 0$):

$$P(n, d) = \begin{cases} 
0 & \text{if } n \text{ is odd} \\
0 & \text{if } n > 0 \text{ and } d = 0 \\
\sum & \text{otherwise}
\end{cases}$$

(b) The runtime to compute $P(n, d)$ using dynamic programming is $O(\underline{\text{_________}})$. 
Lecture 9: Shortest Paths

- SSSP: Dijkstra (for non-negative weights)
  - Bellman-Ford (for negative weights, detects negative cycle)

- APSP: Floyd-Warshall
  - Johnson’s algorithm (BF, make edge-lengths non-neg, n Dijkstras)
  - Matrix Multiplication
Lecture 10-11: Network Flows

• s-t Flow Network: capacities on edges
• Ford-Fulkerson. Residual graph.

(a) What is the value of the maximum S-T flow in the graph below? __________

(b) Draw the residual graph that results after running the Ford-Fulkerson algorithm to completion on the above graph. Label the edges with their residual capacities. (Vertices shown below for your convenience).
Lecture 10-11: Network Flows

- s-t Flow Network: capacities on edges
- Ford-Fulkerson. Residual graph.

- Max-flow Min-Cut theorem. Integer Max-flow if capacities are integers.
- Solves Bipartite matching.

- Poly-time rules: Edmonds-Karp fattest path, E-K shortest augmenting path.
- Min-cost max-flow
In lecture we saw how to compute a maximum matching in a bipartite graph using a maximum flow algorithm. Let’s adapt this idea to find a 2-factor in a bipartite graph (or determine that no 2-factor exists).

(a) Think of a construction to convert any bipartite graph $G$ into a flow network $G'$ such that $G$ has a 2-factor if and only if $G'$ has a maximum flow that is “large enough”. Illustrate your construction on the graph below by adding two additional vertices ($s$ and $t$), and adding some edges. For each edge give it a direction. Label each edge with a capacity.
Lecture 12: Game Theory

• Assume players have independent randomness
• $V_R(p, q) = \sum_{i,j} \Pr[\text{row player plays } i, \text{ column player plays } j] \cdot R_{i,j} = \sum_{i,j} p_i q_j R_{i,j}$
• $V_C(p, q) = \sum_{i,j} \Pr[\text{row player plays } i, \text{ column player plays } j] \cdot C_{i,j} = \sum_{i,j} p_i q_j C_{i,j}$

• What is $V_R(p, q) + V_C(p, q)$?
  • 0, since zero-sum game

If $p = (.5, .5)$ and $q = (.5, .5)$ what is $V_R$?

$V_R = .25 \cdot (-1) + .25 \cdot 1 + .25 \cdot 1 + .25 \cdot (-1)$

If $p = (.75, .25)$ and $q = (.6, .4)$ what is $V_R$? $V_R = -0.1$

Von Neumann: Given a finite 2-player zero-sum game,

$$lb = \max_p \min_q V_R(p, q) = \min_q \max_p V_R(p, q) = ub$$
Lecture 13: Linear Programming

Variables: S, P, E
Objective: Maximize 2P + E subject to
Constraints: 

S + P + E = 168
E ≥ 56
S ≥ 60
2S + E − 3P ≥ 150
P + E ≥ 70
P ≥ 0
Start at vertex of the feasible region (polyhedron in high dimensions)
Look at cost of objective function at each neighbor
Move to neighbor of minimum cost
Always make progress, but could take exponential time (in high dimensions)
Lecture 14: Seidel’s Algorithm

- Order constraints randomly: \( a_{i_1} \cdot x \leq b_{i_1}, \ldots, a_{i_{m-2}} \cdot x \leq b_{i_{m-2}}, c_1, c_2 \)
  - Leave \( c_1, c_2 \) at the end
- Recursively find the optimum \( x^* \) of \( a_{i_2} \cdot x \leq b_{i_2}, \ldots, a_{i_{m-2}} \cdot x \leq b_{i_{m-2}}, c_1, c_2 \)
- Case 1: If \( a_{i_1} \cdot x^* \leq b_{i_1} \), then \( x^* \) is overall optimum
  - \( O(1) \) time
- Case 2: If \( a_{i_1} \cdot x^* > b_{i_1} \), then we need to intersect the line \( a_{i_1} \cdot x = b_{i_1} \) with each other line \( a_{i_j} \cdot x = b_{i_j} \) and solve a 1-dimensional problem in \( O(m) \) time
Lecture 15: Duality

\[ P = \max(2x_1 + 3x_2) \]
\[ \text{s.t. } 4x_1 + 8x_2 \leq 12 \]
\[ 2x_1 + x_2 \leq 3 \]
\[ x_1, x_2 \geq 0 \]

\[ 4y_1 + 2y_2 + 3y_3 \geq 2 \]
\[ 8y_1 + y_2 + 2y_3 \geq 3 \]
\[ y_1, y_2, y_3 \geq 0 \]

and we seek \[ \min(12y_1 + 3y_2 + 4y_3) \]

• Dual of the dual is the primal!

• Can we get better upper/lower bounds by looking at more complicated combinations of the inequalities, not just linear combinations?
### Lecture 15: Duality

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>O</th>
<th>U</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>P</strong></td>
<td><strong>D</strong></td>
<td><strong>I</strong></td>
<td><strong>O</strong></td>
</tr>
<tr>
<td><strong>I</strong></td>
<td>✓</td>
<td>X</td>
<td>✓</td>
</tr>
<tr>
<td><strong>O</strong></td>
<td>X</td>
<td>✓</td>
<td>X</td>
</tr>
<tr>
<td><strong>U</strong></td>
<td>✓</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

I means infeasible
O means feasible and bounded
U means unbounded

Check means possible
X means impossible
Lecture 16: NP-completeness

• **P**: (decision) problems solvable in poly-time

• **NP**: (dec.) probs with short “proofs” that can be verified in poly-time

• **NP-complete**: prob in NP, and every problem in NP reduces to it
  • circuit-SAT, 3SAT, CLIQUE, IND-SET, Vertex Cover, Set Cover, 3-Coloring, Hamilton Path, Partition

• to show problem Q is NP-complete:
  • show that Q in NP
  • reduce from known NP-complete problem to unknown problem
    “so the new problem is at least as hard as some NP-complete problem”
The students of 15-451/651 like forming clubs. Each club consists of exactly 3 distinct members. Students may belong to multiple clubs (or may belong to none at all.) We want to pick a subset of students to designate as “officers”, such that each club contains at least one officer. A set $S$ of students is good if every club intersects $S$.

E.g., if there is a student who belongs to every club, we can just designate her as officer, and all clubs will contain an officer. If all clubs are disjoint, we need to pick one officer per club.

The decision problem Officers is: given the roster $R$ of the class (with $n$ students), a list of clubs $C_1, \ldots, C_m$ (each a subset of $R$ of size 3), and an integer $K > 0$, does there exist a good set of size $K$?

Show this problem is $\text{NP}$ complete. (Hint: One way is to use vertex cover.)

Officers is in $\text{NP}$ because a verifier can efficiently check a proposed solution as follows:

To prove that Officers is $\text{NP}$-hard, we can reduce from

$Q = \text{__________________________} \to Q' = \text{__________________________}$

Specifically, given an instance $I$ of $Q$ we create an instance $f(I)$ of $Q'$ as follows:

Now we show that

$I$ is a YES-instance of $Q \iff f(I)$ is a YES-instance of $Q'$

as follows:

Finally, time taken to write down $f(I)$ is _________.
Lec 17: Approx Algos

• Derive some “surrogate” for OPT value.
• Show we pay not too much more.

• E.g., $p_{\text{max}}$, $\frac{\sum_j p_j}{m}$ are lower bounds for OPT makespan
  • or size of maximal matching for vertex cover

• Can write integer LP, relax to get LP, “round” LP to get integer sol.
Lec 18-19: Online Algos and Experts

• Algorithms that work without knowing future
• competitive ratio = max_{all inputs I} ALG(I)/OPT(I).

• Ski-rental, List Update (MoveToFront)
  • potential function capturing difference between us and other algo “B”.

• Multiplicative Weights:
  • reduce the weight of every wrong expert by (1/2) or (1-epsilon)
  • #mistakes <= (1+eps)(#mistakes of best expert) + O(\log N/eps).
Lec 20: Gradient Descent

- Minimizing convex function:
  - Move small step in direction of negative gradient
    \[ x_{t+1} \leftarrow x_t - \eta_t \nabla f(x_t) \]
  - If done carefully, get within eps of optimum in time \( O(1/\text{eps}^2) \).
  - Proof via potential = \( \| x - x^* \|^2 \)

- In HW, approximately solve \( Ax = b \) quickly without Gaussian elim. (when \( A \) is “well-conditioned”).
Lec 21-22: Auctions and VCG

- We want to allocate items to maximize social welfare = sum of values.
- But players want to maximize their own utility = value – price.
- Use prices to incentivize players to bid truthfully.
- Vickery (second-price auction), VCG.
  - get bids, find the allocation that maximizes the social welfare (treating bids as truth).
  - charge players = externality: how much they caused others value to decrease.

Consider the following instance of a matching market, with items A, B, C and buyers a, b, c. The valuations of a for the items are $v_{Aa} = 12, v_{Ba} = 10, v_{Ca} = 7$ (or in short, the valuations of a for the items are 12,10,7), the valuations of b for the items are 10, 3, 5, and valuations of c are 8, 6, 5.
Lecture 23: Graph Compression

• Multiplicative spanners
  • Greedy algorithm to find them
  • Use high girth to prove they have a small number of edges

• Additive spanners

  • $Q$ is a shortest path from $e$ to $v$ in $G$!
  • $d_Q(e, v) \leq 1 + d_P(c, v)$
  • $d_{P'}(u, v) = d_P(u, c) + 1 + d_Q(e, v) \leq d_P(u, c) + d_P(c, v) + 2 = d_P(u, v) + 2$
Lecture 24: Sketching Vectors

- Sketches are linear so can be updated in the presence of insertions and deletions to coordinates.

- CountSketch estimates the Euclidean norm of a vector.
  - Good to know Chebyshev’s inequality.

- 1-Sparse Finder: Let \( p = \text{poly}(n) \) be a prime, and \( z \) a random integer mod \( p \).
  
  Maintain \( A = \sum_i x_i \), \( B = \sum_i x_i \cdot i \), and \( C = \sum_i x_i \cdot z^i \text{mod } p \).
  
  Can output a non-zero item of a vector with \( O(\log^2 n) \) bits of space.
Lecture 25: Sketching Graphs

• (Modified) Boruvka’s spanning tree algorithm for finding a spanning tree by arbitrarily choosing an outgoing edge of each group of vertices in each iteration

\[
a_1 = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}
\]

• Run Modified Boruvka on the sketches

\[
\sum_{j \in S} c_i a_j = c_i \left( \sum_{j \in S} a_j \right) \rightarrow e \in \text{support} \left( \sum_{j \in S} a_j \right) = E(S, V \setminus S)
\]
Lec 26: Polynomials

• Non-zero degree-d polynomial has at most d roots.
• Given d+1 points \((a_i, b_i)\) there is unique poly \(P(x)\) passing thru them.
  • Lagrange interpolation finds this polynomial
  • Chinese remainder theorem uses exactly the same principle!

• Can use to error correct!
  • View message of length n+1 as poly of degree n.
  • Evaluate at n+2k+1 locations (2k more than necessary)
  • Interpolate correctly if at most k errors.