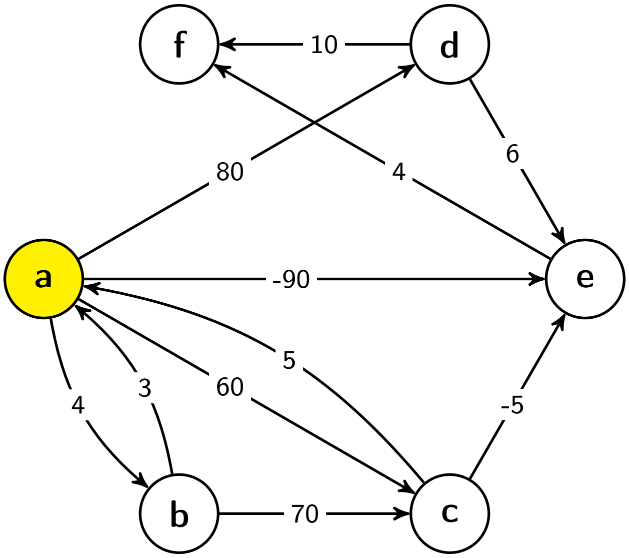
CSE 373 Final Review

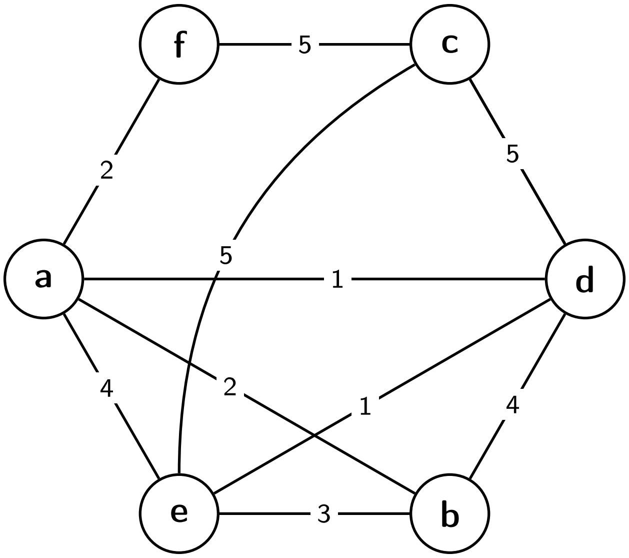
1. Dijkstra's



(a) Use Dijkstra's Algorithm to find the lengths of the shortest paths from a to each of the other vertices. For full credit, you must show the worklist at every step, but how you show it is up to you.

(b) Are any of the lengths you computed using Dijkstra's Algorithm in part (a) incorrect? For each length that is incorrect, explain what the correct answer is and why the answer from part (a) was incorrect.

2. MST's



(a) Use Kruskal's Algorithm to find two minimum spanning trees of the above graph.

(b) Imagine that the above graph had some negative edges in it. Would Prim's Algorithm necessarily return a correct result? Explain your answer in 1-2 sentences.

3. FootPen

A new social networking company called FootPen has arrived on the start-up scene. Because UW CSE students are known to be the best in industry, FootPen has asked you to help them implement several features. For each feature request,

- explain how to represent the problem as a graph,
- give a (high-level) idea for an algorithm to solve the problem, and
- give a runtime analysis of your algorithm.

(a) How can FootPen determine the number of people that have at least one friend in common with a particular user.

(b) FootPen would like to add a “make a new friend” feature. To facilitate this feature, FootPen has an algorithm that generates an “familiarity score” for every pair of friends. The familiarity score is a real number between 0.0 and 1.0 which represents how well two users know each other. The closer the score is to zero, the more friendly the two users are.

The “make a new friend” feature is intended to help users find a single user who they do not know but are likely to get along with. We can estimate a familiarity score between two users who are not friends by summing the scores of the shortest weight chain of friends that joins the two non-friends.

FootPen has run experiments and found that an estimated familiarity score of between 0.5 and 1 is likely to indicate that two users do not already know each other but would make good friends.

How can FootPen write the “make a friend feature”?

(c) Now that FootPen has become popular, the government has sent a request asking for the group of people who correspond with each other the most. FootPen has records the number of conversations each pair of users has in its database. How can FootPen identify the k users with the largest number of inter-communications?

4. P and NP

a) “NP” stands for ________________________________

b) What does it mean for a problem to be in NP?

5. Sorting

a) For each of the following situations, name the best sorting algorithm from among those we studied. There may be more than one answer deserving full credit, but you only need to give one answer for each

i) You need a very fast sort on average, and you can only use a constant amount of extra space.

ii) The array is in perfect sorted order.

iii) You have a large data set, but you know all the values are between 0 and 999.

iv) Copying your data is very fast, but comparisons are relatively slow.

b) Please provide an example of each of the following (or say NONE if no example exists) from among the sorting algorithms we studied. Reminder: A stable sort is one which preserves the original order of the input set, where the comparison algorithm does not distinguish between two or more items.

i) A stable comparison sort with a worst-case O(n^2) running time

ii) A stable comparison sort with a worst-case O(n) running time

iii) An efficient (i.e. O(n log n)) stable comparison sort

iv) A stable sort with a worst-case running time linear in n

6. Topics Prior to Midterm

a) Let $f : N \to N$ and $g : N \to N$ be increasing functions with $f(n)$ 6= $g(n)$ for any $n \in N$. Consider the statement:

$$f(n) \in \Omega(g(n)) \text{ and } g(n) \in \Omega(f(n))$$

Is this statement always, sometimes, or never true? Explain your answer in one sentence.

b) Insert 1, 2, 7, 6, 8, 3, 4, 5 into an AVL tree in that order. You do not have to show your intermediary steps but no work and a wrong answer will receive no credit.

c) For each of the following, decide which data structure is most appropriate to solve the problem:

AVL Tree | FIFO Queue | Hash Table | Linked List | Heap | Vanilla BST

i) An operating system needs to schedule when to run each of the current processes.

ii) You want to store 1MB of non-comparable data and you expect to run the find operation. very frequently

iii) A grocery store wants to make a database of their products
(they only have about 5000unique products) so that employees can look up
up the aisle that a particular product is in by name

d) You know in advance that you will never put more than six strings into a particular hash table.

i) If you had a choice between 6 and 7 as your table size, which would you choose? Why?

ii) Insert 1, 5, 9, 12, 37, 16 into your hash table using the hash function h(x) = x and separate chaining

iii) What is the load factor of your hash table from part (ii)?

iv) Insert the same numbers into an empty version of your hash table using the hash function h(x) = x and quadratic probing

v) Does your hash table from part (d) have primary clustering? What about secondary clustering?

vi) What is the load factor of your hash table from part (v)?