Introduction to Algorithms: 6.006
Massachusetts Institute of Technology October 21, 2008
Professors Ronald L. Rivest and Sivan Toledo Handout 3

Problem Set 4

This problem set is divided into two parts: Part A problems are programming tasks, and Part B problems are theory questions.

Part A questions are due **Tuesday, November 4** at **11:59PM**.

Part B questions are due **Thursday, November 6** at **11:59PM**.

Solutions should be turned in through the course website in PDF form using LaTeX or scanned handwritten solutions.

A template for writing up solutions in LaTeX is available on the course website.

Remember, your goal is to communicate. Full credit will be given only to the correct solution which is described clearly. Convoluted and obtuse descriptions might receive low marks, even when they are correct. Also, aim for concise solutions, as it will save you time spent on write-ups, and also help you conceptualize the key idea of the problem.

Exercises are for extra practice and should not be turned in.

Exercises:

- CLRS 22.2-3 (page 539)
- CLRS 22.2-8 (page 539)
- CLRS 22.3-9 (page 548)
- CLRS 22.3-10 (page 549)

Part A: Due Tuesday, November 4

1. **(50 points)** $2 \times 2 \times 2$ Rubik's Cube

   We say that a configuration of the cube is k levels from the solved position if you can reach the configuration in exactly k twists, but cannot reach the it in any fewer twists.

   Download `ps4_rubik.zip` from the class website.

   (a) **(20 points)** For this problem, we will use breadth-first search to recreate the column labeled f in the chart seen at `http://en.wikipedia.org/wiki/Pocket_Cube`. Write a function `positions_at_level` in `level.py` that takes a nonnegative integer argument `level`, and returns the number of configurations that are `level` levels from the solved configuration (`rubik.I`), using both quarter twists and half twists (twisting the cube by 90 or 180 degrees).

The code in `rubik.py` only defines the `rubik.quarter_twists` move set, so you should start by defining a new move set that includes half twists as well. Do not modify `rubik.quarter_twists` because you will need it for the next part of this problem.

Test your code using `test_level.py`, and submit it to the class website. Testcases above level 8 are commented out, since they may require more memory than many computers have.

(b) **(30 points)** Now we will solve the Rubik's cube puzzle by finding the shortest path between two configurations (the start and goal). For this part of the problem, we will limit the move set to only allow quarter twists (half twists are not allowed).

Your code from part (a) could easily be modified to find shortest paths, but a BFS that goes as deep as 14 levels will take a few minutes (not to mention the memory needed). A few minutes might be fine for creating a Wikipedia page, but we want to solve the cube fast!

Instead, we will take advantage of a property of the graph that we can see in the chart. In particular, the number of nodes at level 7 (half the diameter) is much smaller than half the total number of nodes.

With this in mind, we can instead do a two-way BFS, starting from each end at the same time, and meeting in the middle. At each step, expand one level from the start position, and one level from the end position, and then check to see whether any of the new nodes have been discovered in both searches. If there is such a node, we can just read off parent pointers (in the correct order) to return the shortest path.

Write a function `shortest_path` in `solver.py` that takes two positions, and returns a list of moves that is a shortest path between the two positions.

Test your code using `test_solver.py`. Check that your code runs at close to the same speed as level 7 from part(a) in the worst case, after modifying it to use just the quater twist move set.

(c) **(Optional)** Go out and impress your friends with new 2x2x2 Rubik's Cube solver you just created! You can test your code using `test_human_solver.py`, which will ask you to input the current configuration of a your Rubik's cube, and then give you the shortest solution path in human-readable symbols (you may need to read `rubik.py` to understand these symbols though).

Part B: Due Thursday, November 6

1. **(15 points)** Connected Components

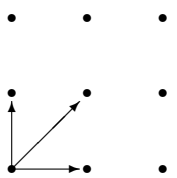
Given an undirected graph $G = (V, E)$, a *connected component* of G is a set of verticies $C \subseteq V$ for which the following two properties hold.

- every two vertices in C are connected by a path.
  $v_1, v_2 \in C \implies \exists$ a path in G $v_1 = x_0 \rightarrow x_1 \rightarrow \ ... \ \rightarrow x_k = v_2$
- no edge connects a vertex inside the set to a vertex outside the set.
  $v \in C$ and $w \in V/C \implies \{v, w\} \notin E$
  (V/C is the set V minus the set C)

Give an $O(V + E)$-time algorithm for partitioning an undirected graph into connected components. That is, given a graph $G = (V, E)$ return a set $S = \{C_i\}$ where each C_i is a connected component of G and $\bigcup C_i = V$

2. **(20 points)** Eliminating Cycles by Removing One Edge

   For each of the following statements, prove the statement or give a counter example to show that it is false. If you give a counter example, give one with the fewest possible vertices. Use LaTeX to draw counter-example graphs if necessary (the solution template contains a drawing of the following graph to get you started).



   (a) **(10 points)** If DFS on a graph G produces exactly one back edge, then it is possible to remove an edge from G to make the graph acyclic.

   (b) **(10 points)** If G is cyclic but can be made acyclic by removing one edge, then DFS will encounter exactly one back edge.

3. **(15 points)** Graphs and Matrices

   The *incidence matrix* of an undirected graph $G = (V, E)$ is an $|V| \times |E|$ matrix U, in which every column corresponds to an edge $e = \{i, j\} \in E$. The entries of column $e = \{i, j\}$ are all zero except for the entries in rows i and j which are both 1.

   Consider the matrix $A = UU^T$. (that is, A is the matrix product of U and its transpose; the i, j entry of a product $X = Y * Z^T$ is $X_{i,j} = \sum_k Y_{i,k} Z j, k$ where the summation is over a column of Y and a column of Z.) Describe two ways in which the entries of A relate to properties of the graph G.