Programming Assignment Five: Graph Algorithms

Out: Nov. 30, 2018; Due: Dec. 13, 2018.

I. Motivation

1. To give you experience in implementing a graph data structure using the adjacency list representation.
2. To give you experience in implementing a few graph algorithms.

II. Programming Assignment

You will read from the **standard input** a description of a directed graph and then report two things on that graph:

1. Whether the graph is a directed acyclic graph (DAG).

2. Determine the total weight of a minimum spanning tree (MST) of the graph when it is treated as an undirected graph.

1. Input Format

The first line in the input specifies the number of nodes in the graph, N. The nodes in the graph are indexed from 0 to $N-1$. Each subsequent line represents a **directed** edge by 3 numbers in the form:

```
<start_node> <end_node> <weight>
```

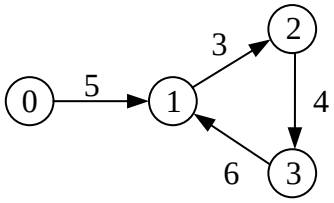
where both `<start_node>` and `<end_node>` are integers in the range $[0, N-1]$, representing the start node and the end node of the edge, respectively, and `<weight>` is an **integer** representing the edge weight.

An example of input is

```
4
0 1 5
```

```
1 2 3
2 3 4
3 1 6
```

It represents the following directed graph:



Typically, we will describe the graph in a file. However, since your program takes input from the standard input, you need to use the Linux input redirection "<" on the command line to read the graph from the file.

2. Output Specification

Your program writes to the **standard output**. It will first tell whether the graph is a DAG or not. Then, it will calculate the total weight of a MST of the original graph when treated as an undirected graph.

a) Telling whether the graph is a DAG or not

If the graph is a DAG, your program should print:

```
The graph is a DAG
```

Otherwise, print:

```
The graph is not a DAG
```

b) Calculating the total weight of an MST

You should treat all the directed edges in the original graph as undirected edges, and obtain the total edge weight of a minimum spanning tree for the undirected graph. If there exists a spanning

tree, your program should print the total edge weight of an MST. Specifically, it should print the following line:

```
The total weight of MST is <total_weight>
```

where `<total_weight>` specifies the total edge weight of an MST.

If there exists no spanning tree, your program should print:

```
No MST exists!
```

For the example graph shown above, the valid output should be:

```
The graph is not a DAG
The total weight of MST is 12
```

III. Program Arguments and Error Checking

Your program takes no arguments. You do not need to do any error checking. You can assume that all the inputs are syntactically correct.

IV. Implementation Requirements and Restrictions

- You must implement the graph data structure using the **adjacency list representation**.
- You must make sure that your code compiles successfully on a Linux operating system. You are required to write your own `Makefile` and submit it together with your source code files. **Your compiled program should be named as `main` exactly.**
- We highly encourage the use of the STL, such as `vector`, `deque`, `map`, etc., for this project, with the exception of two prohibited features: The C++11 regular expressions library (whose implementation in gcc 4.7 is unreliable) and the thread/atomics libraries (which spoil runtime measurements). Do not use other advanced libraries (e.g., boost, pthreads, etc).
- Output should only be done where it is specified.

V. Hints

- You may want to define the following three abstract data types: node, edge, and graph.
- To improve the performance of your program, we recommend you to add `-O2` option when compiling your programs.
- For large test cases, I/O could be the runtime bottleneck. To reduce the I/O time, we recommend you to put the following two statements at the beginning of your main function:

```
std::ios::sync_with_stdio(false);
std::cin.tie(0);
```

VI. Testing

We have supplied three input files `g1.in`, `g2.in`, and `g3.in` for you to test your program. The outputs of our program for these three inputs are also provided. They are `g1.out`, `g2.out`, and `g3.out`. All these files are located in the Programming-Assignment-Five-Related-Files.zip. To do the test, type a similar command to the following into the Linux terminal once your program has been compiled:

```
./main < g1.in > test.out
diff test.out g1.out
```

If the `diff` program reports any differences at all, you have a bug.

These are the minimal amount of tests you should run to check your program. Those programs that do not pass these tests are not likely to receive much credit. You should also write other different test cases yourself to test your program extensively.

VII. Submitting and Due Date

You should submit all the related `.h` files, `.cpp` files, and the `Makefile` via the online judgement system. The `Makefile` compiles a program named `main`. Please also include a pdf file containing all of your source code. See announcement from the TAs for the details about how to submit these files. The submission deadline is 11:59 pm on Dec. 13, 2018.

VIII. Grading

Your program will be graded along four criteria:

1. Functional Correctness
2. Implementation Constraints
3. General Style
4. Performance

Functional Correctness is determined by running a variety of test cases against your program, checking your solution using our automatic testing program. We will grade Implementation Constraints to see if you have met all of the implementation requirements and restrictions. For example, we will check whether you implement the graph data structure using the adjacency list representation. General Style refers to the ease with which TAs can read and understand your program, and the cleanliness and elegance of your code. For example, if it is hard for the TAs to check whether your graph data structure is implemented using adjacency list, it will lead to General Style deductions. Part of your grade will also be determined by the performance of your algorithm. We will test your program with some large test cases. If your program is not able to finish within a reasonable amount of time, you will lose the performance score for those test cases.