EDAF05 Lab exercise 3

SPANNING USA

Thore Husfeldt
25 March 2009 9:26

DESCRIPTION

Find a minimum spanning tree of highway distances between 128 American cities

Files

The single input file for this exercise is in `/usr/local/cs/edaf05/lab3/`

Input

The input file is “USA-highway-miles.in”. It describes the intercity highway distance of 128 cities in the USA, based on data collected for the Stanford Graph Base.
It starts with listing the names of all cities, one per line. Names that contain whitespace are enclosed in quote marks. Then follows a line for every pair of cities with the distance given as an integer in square brackets.

Output

Print the total weight of the resulting tree. I get 16598.

Requirements

Minimal solution

Your algorithm has to work correctly (not only on this input – otherwise you could just write “print 16598”), and run in polynomial time. You can decide yourself if you want to use Prim’s or Kruskal’s algorithm. Note that I don’t require you to implement fancy tricks like a decrease-key priority queue and or a clever union–find data structures for connected components. The graph isn’t that big, and quadratic time will do just fine.

Good solution

Solve the problem in O(n log n) and optimise as much as you can. This involved either writing a good priority queue for Kruskal’s algorithm, or a union–find data structure for Prim’s. (It’s very good to know these things, but this is not primarily a data structures course.)

Tips

Be careful with reading the input file, names of cities don’t necessarily behave as you expect. Especially, if you get the answer 16394, your error is probably in the parsing stage, not in the algorithm.

About the input file

If you’re curious, the input file (like many of the files we use in this course) is based on the *Stanford GraphBase* by Donald Knuth. In this case, I modified Knuth’s file “miles.dat”, which you can find on the internet. It contains the correct intercity highway distances from 1949, based on Rand McNally’s *Standard Highway Mileage Guide* (1949). Knuth’s data includes slight modifications to those numbers (for example, the distances in miles.dat satisfy the triangle inequality) and some more information such as population and location.

```
Duckburg
"Gotham City"
Metropolis
Duckburg--"Gotham City" [2324]
Duckburg-Metropolis [231]
“Gotham City”-Metropolis [2298]
```

Format for the input file