Code Jam 2016 - Round 1C

Analysis: Slides!

Slides!: Analysis!

Small dataset

The Small dataset has bounds that suggest we can construct all possible sets of slides, but this turns out to be overly optimistic. We represent a set of slides as a directed graph *G*, with each node representing a building, and a directed edge from node *i* to node *j* representing a slide leading from building *i* to building *j*. The most straightforward construction tries either including or not including each of the **B** * (**B**-1) possible slides, for a total of $2^{\mathbf{B} * (\mathbf{B}-1)}$ possible sets of slides. Unfortunately, even for **B** = 6, this is too many: there are approximately 2^{30} sets to check, or about a billion.

However, one observation allows us to dramatically cut down the number of sets we have to examine. Notice that there can never be a cycle as part of any valid path from building 1 to building **B**. If there were a cycle, then we could generate new, valid paths by traversing that cycle arbitrarily many times before continuing to our destination, meaning that the number of valid paths would be infinite. (The *G* that we use can still contain a cycle that is not on any valid path; however, removing that cycle would not affect the number of valid paths, and thus we only need to consider graphs *G* with no cycles at all.)

This means that any valid path from building 1 to building **B** cannot visit the same building twice, so each path can have length at most **B**. As a result, running a depth-first search on *G* starting from node 1 will take O(**B**) time for each path found. If we find more than **M** paths, then we can terminate our search immediately, since this set of slides cannot be valid. This means that our worst-case running time to test any given set of slides is O(**M***B**). We can also calculate a smaller upper bound on the number of sets we have to examine: for each pair of slides *i* and *j*, exactly one of three possibilities must be true:

- There is a slide from *i* to *j*.
- There is a slide from *j* to *i*.
- There is no slide from *i* to *j* and no slide from *j* to *i*.

Since there are **B** * (**B**-1) / 2 different pairs of slides, this gives us an upper bound of $3^{\mathbf{B} * (\mathbf{B}-1) / 2}$ possible sets of slides. For **B** = 6, this number is around fourteen million, which is a manageable number of sets to check.

Another helpful observation that makes the small even more tractable is that since our graph has no cycles, it is a directed acyclic graph, and so it has a topological sorting. So, for any correct solution, we could renumber the buildings (other than 1 and **B**) such that every slide's end building has a larger number than its start building. Since this is true, we only need to consider slides that go from lower to higher building numbers.

Large dataset

The Large dataset requires a more efficient approach. A natural first question to ask is: what is the maximum number of paths from building 1 to building **B** that we can possibly construct? One straightforward construction that seems to yield a large number of paths is to construct a slide from building *i* to building *j* for every pair of positive integers *i*, *j* with 1 ≤ *i* < *j* ≤ **B**. To compute the number of paths for this set of slides, notice that every path from building 1 to building **B** corresponds uniquely to a set of distinct integers from the set {2, ..., **B**-1} representing the buildings visited along that path. For example, if **B** = 5, then the set {2, 4} would correspond to the path 1 -> 2 -> 4 -> 5, and the empty set would correspond to the path 1 -> 5. Since there are **B**-2 integers strictly between 1 and **B**, each of which can be either absent or present in a set, there are $2^{\mathbf{B}-2}$ unique sets that can be constructed, and thus $2^{\mathbf{B}-2}$ possible paths from 1 to **B**.

But is this the largest possible number of paths we can construct? It turns out that it is. We can show this by the pigeonhole principle. We assume that there exists some set of slides that yields some number **M** > $2^{\mathbf{B}-2}$ paths, and derive a contradiction. Each path corresponds to some set of distinct integers {2, ..., **B**-1} representing the buildings visited along that path, and since there are $2^{\mathbf{B}-2}$ distinct such sets, it follows that two paths of slides must visit the exact same buildings. This means there is some pair of buildings *i* and *j* such that *i* is visited before *j* on one of these paths, but is visited after *j* on the other path. (If there were no such pair, then these two paths would be exactly the same, since no building can ever be visited twice.) Since we can reach building *i* from building *j* and vice versa, it follows that there is a cycle between the two buildings. This contradicts what we showed earlier, meaning that we cannot construct a set of slides with exactly **M** paths for **M** > $2^{\mathbf{B}-2}$.

Now we show how to extend the ideas above to handle the case where **M** < $2^{\mathbf{B}-2}$. We start by constructing all possible slides from building *i* to building *j* for every pair of positive integers *i*, *j* with 2 ≤ *i* < *j* ≤ **B**. Notice that there are exactly $2^{\mathbf{B}-1-i}$ ways to get from building *i* to building **B**. Each path from building *i* to building **B** maps uniquely to a subset of distinct integers from the set {i + 1, ..., **B**-1}. This set contains **B**-1-i integers, so there are $2^{\mathbf{B}-1-i}$ possible subsets that we can choose. If we build a slide from building 1 to building *i* for *i* strictly between 1 and **B**, then this increases the number of ways to get from 1 to **B** by exactly $2^{\mathbf{B}-1-i}$, since there are that many ways to get from building *i* to building **B**. This suggests a method for generating a network with exactly **M** slides. We start by writing **M** in binary. If the *i*-th digit of **M** (counting from the right, starting from 1) is a 1, then we add a path between building 1 and building **B**-i. This will add 2^{i-1} new paths to our slide network. If we repeat this process for each value of *i*, then we will end up with a network with a number of paths from 1 to **B** exactly equal to **M**. This process will work if **M** has at most **B**-2 digits, meaning **M** ≤ $2^{\mathbf{B}-2}$-1. Since **M** = $2^{\mathbf{B}-2}$ is the largest value we are able to construct, this gives us a method for constructing all values of **M** between 1 and $2^{\mathbf{B}-2}$. We have previously shown a construction for **M** = $2^{\mathbf{B}-2}$, which is equivalent to the solution for **M** = $2^{\mathbf{B}-2}$ - 1 with an additional path from building 1 to building **B**.

This means that a sequence of slides is therefore possible to construct if, and only if, **M** ≤ $2^{\mathbf{B}-2}$, and the above construction works for any such **M**. The sequence itself is computed by the construction above.

To illustrate the above method, here are valid answers for all cases with **B** = 5, and **M** = 1 through 8:

```
M = 1 M = 2 M = 3 M = 4 M = 5 M = 6 M = 7 M = 8

00010 00100 00110 01000 01010 01100 01110 01111
00111 00111 00111 00111 00111 00111 00111 00111
00011 00011 00011 00011 00011 00011 00011 00011
00001 00001 00001 00001 00001 00001 00001 00001
00000 00000 00000 00000 00000 00000 00000 00000
```

Observe that the solutions only differ in their first lines. The first lines of the solutions for **M** = 1 through 7 are 1, 2, ..., 7 in binary plus an extra 0 at the end. The first line of the solution for **M** = 8 is 7 in binary, plus an extra 1 at the end: the direct connection from building 1 to building 5 that brings the total to 8. For **M** ≥ 9, the answer is `IMPOSSIBLE`.