COMP 182: Algorithmic Thinking
6 February 2014

A biologist comes back from a field trip with n mice, and she believes that each of these mice belongs to one of two different species, which we will call A and B. The biologist would like to divide the n specimens into two groups—those that belong to A and those that belong to B—but it's very hard for her to directly label any one specimen. So, she adopts the following approach.

For each pair of specimen i and j, she studies them carefully side by side and makes one of three decisions:

- *same*: she is certain that i and j belong to the same species;
- *different*: she is certain that i and j belong to two different species; or,
- *ambiguous*: she can't decide with certainty whether i and j belong to the same species or not.

So, now the biologist has a collection of n specimens and a collection of m judgments (*same* or *different*) for the pairs that were not declared ambiguous. The biologist would like to know if this data is consistent with the idea that each mouse is from one of species A or B. We declare the m judgments to be *consistent* if it is possible to label each specimen A or B in such a way that for each pair i and j, the following hold:

- if i and j were declared "same," then i and j have the same label, and
- if i and j were declared "different," then i are j have different labels.

Help the biologist by devising an efficient algorithm for the problem. First, formulate the problem. Then, give the pseudo-code of an $O(m+n)$ algorithm. Finally, analyze the algorithm's running time to show it indeed takes $O(m+n)$ time.