# Artificial Intelligence Assignment 4
# Medical Diagnosis

**Aakash Kumar (2012CS10201)**
**Aashish Arora (2012MT50581)**

We have implemented the EM algorithm in this assignment.
The implementation is straight forward. First we needed to either complete the data set given to us or fill the CPT with random values maintaining consistency. We tried both and the results were same. To complete the data set, we use the follwing method:

Take a data point (a line in .dat file) and check if it has a question mark in it. If it has no question mark then insert the data into a vector and assign a weight of 1 to it. Otherwise for every possible value that the question mark can take, insert a data point and assign a weight of 1/(number of possible values) to each of the points inserted.

Now we have a complete data set. Now we fill the CPT by simple counting.
After completing the CPT, we use this CPT to find the probability that the question mark is equal to one of the possible values in a dataset.

We repeat this procedure till time remaining is less than five seconds. Then we rewrite the file with the current CPT.

We have used a smoothing factor 0.001 for calculating CPT using the data set.

Collaboration: Ayush Sakhuja