DT and AdaBoost

1. Here is some randomly generated data for two classes, A and B, with eight boolean attributes. Generate the first two levels of a decision tree for these classes. Specifically, you should identify which of the eight attributes to test first, based on information gain, and for either value of that attribute, which one to test next - then, for each of these four cases, which prediction to give (A or B). It is suggested (but not necessary) you write some code to help you, you can use this code later in your second programming project.

2. If you were to run Adaboost on the preceding problem, using decision stumps, assuming the use of information gain, consider the first iteration of the Adaboost algorithm. The first stump should be the same as the root of your previous tree. (R&N 751) Now:

   a. What would the error rate of the first stump be?
   b. What would be the hypothesis weight of the first stump?
   c. What would be the initial weights of each example?
   d. What would be the new weights of each example that was correctly classified, and each that was incorrectly classified?