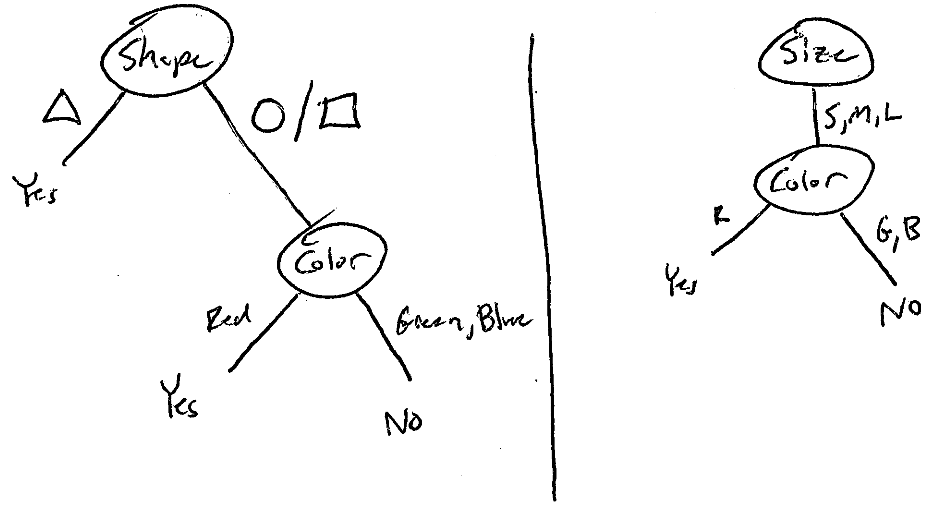
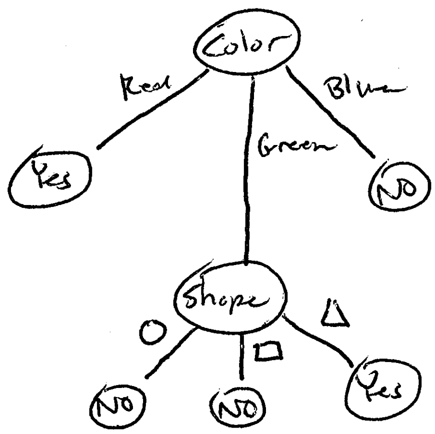
1.



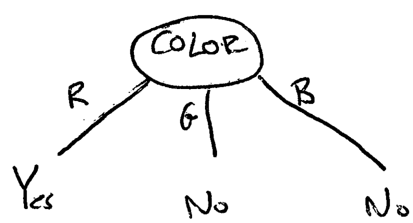
2. The easiest way to do this would be to work backwards. By taking all of the correct results from the leaves and working your way up the tree you will narrow down the possible attributes one by one until the possible attributes is narrowed down to only one, which would be the solution.

3.



Out of 15 tests, none failed.

4.



Out of 15 tests, only one failed. (Was classified false)

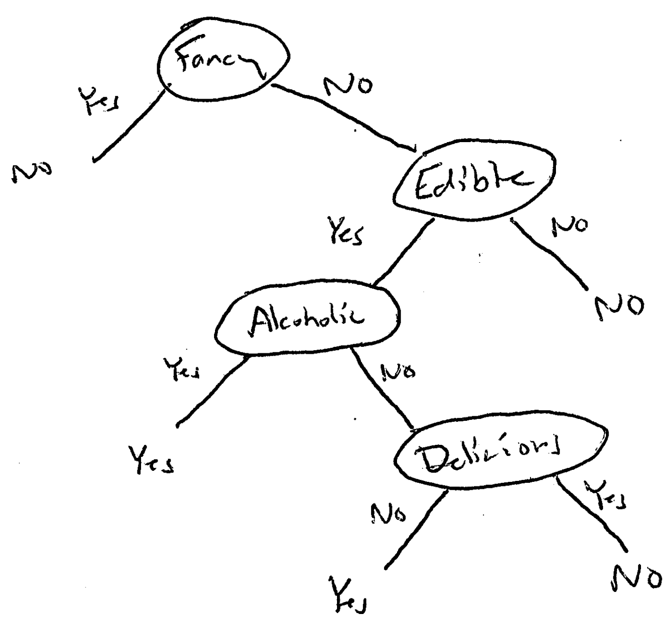
It was the case of "any-size, green, triangle."

5. I found that I was getting no errors at about 65 cases.

6. After altering the input to have 4.6% false data it returned an incorrect data set containing 7.6923% misclassified instances. This was surprising as I didn't expect the result to get worse. I expected the data to only have the ones incorrect that I had changed. My reason being that there were no errors before I made the alterations.

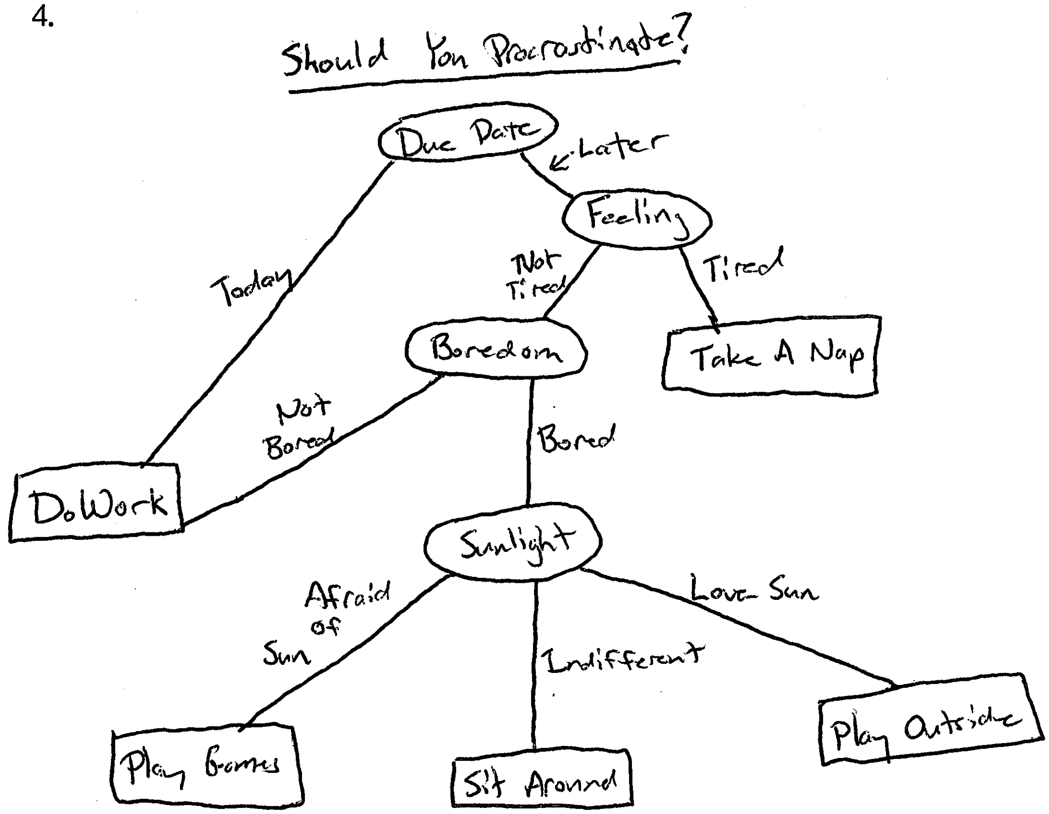
7. With the totally random data both ID3 and J48 came out to have correct and incorrect classifications of almost exactly 50/50 +-5%. When I upped the training data to 7000 the incorrect and correct results were even closer to exactly 50/50.

8. **Simple example:**
    1. Approximately 16 distinct samples.
    2. In this sample I considered it simple because the final outcome is either a yes or a no response. In this example we are looking to see if an item is affordable by college students. The only two true cases are when the item isn't fancy, is edible, and is alcoholic or disgusting.
    3. When I tested my simple example it came out with 100% of the results as being correctly classified. When I scaled it down to 5 different samples it still got every one of them correct. I believe this is because my cases are so specific it is very difficult not to categorize it correctly. As for upscaling this, it works just as well for the same reasons.
    4.



    5. Provided at the end of the lab report.

9. **Complex Example:**
    1. Approximately 32 distinct samples.
    2. In the complex sample I considered it such because the final outcome had many varying responses, making it more specific. This sample would help a college student decide whether or not to procrastinate on work or not.
    3. For this complex example I was getting incorrect classifications up until I reached 50 samples. After that I quit receiving misclassifications. By scaling down it got worse and by scaling up it fixed the problem by giving more samples to learn from.
    4.



5. When given a case similar to the needle in the haystack example, you would need a large set of data if it is randomly generated. In the case were the the randomly generated data doesn't have a positive example there is no way to know what the positive outcome would be. You could use a small data set on the other hand if you provide all of the possible positive results. On the other hand if the attributes are mostly irrelevant this is also not a great case. By providing unimportant information there is nothing to be learned from the dataset, and if there is, it is mostly minuscule.