Assignment: A6

Due: 20 November 2012

You are to explore the use of the Genetic Algorithm to find a decision tree that can classify scanned images of the 26 lower-case characters (i.e., a-z). Several aspects of this approach deserve careful attention:

- **Organism**: choose an appropriate organism
    1. Pick attributes that allow GA computation
    2. Compare binary GA to subtree GA
- **Parent Selection Methods (study option 1 at least)**:
    1. Study performance as a function of straight percentage or probabilistic choice
    2. Compare with respect to percentage selected (make enough offspring)
- **Crossover Methods**: Study 2 of 3 at least
    1. Single point
    2. Two-point
    3. Random each bit
- **Mutation**:
    1. Study mutation percentage
- **Data Management**:
    1. Describe how you select training and testing data

In addition, the results need to be presented in a strong statistical framework; this means computing statistics (e.g., mean, variance) over several trials (how many?), and showing confidence intervals.

Finally, the analysis and interpretation are the essential parts of the report; use these to present your findings, understanding and remaining problems.

In this assignment, the major goal is to explore the use of genetic algorithms.

There is a set of sample images on the class data sub-directory.