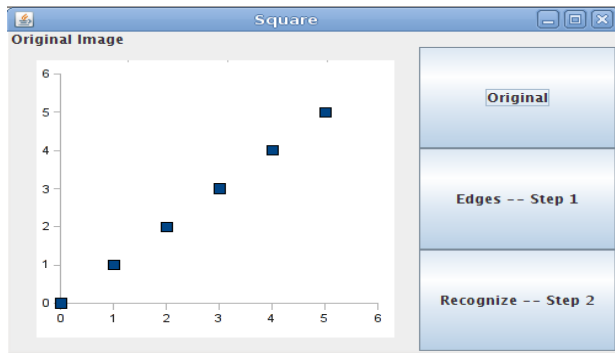


Reverse Engineering Graphs: Obtaining Data Points from Scatterplots

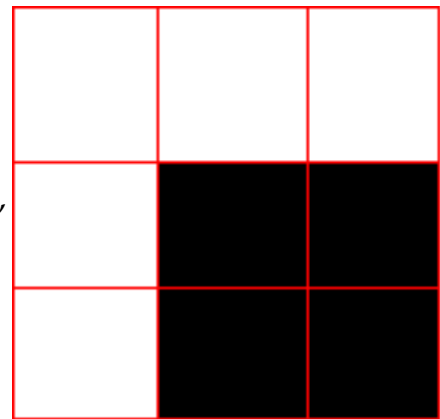
By: Maya Wei

Abstract: Various programs exist to take data points and use them to render a graph. However, once the data are put into visual form, there is a loss of numerical information if the original data cannot be obtained. This project seeks to take data from a graph; in essence, the purpose is to reverse engineer a given graph. This will provide for a set of data points which can be used for various other numerical purposes, not simply the graph form in which they are presented.



Background: The project deals in basic methods of image analysis and shape differentiation. It relies much on edge detection, which can be from very basic techniques to very advanced algorithms. The graph used was created in Open Office Calc; there are no gridlines and no colored background. The image is saved as .png.

Methodology: Currently, the edge detection method in this program is highly primitive. It takes the color of the point at $(0, 0)$ and makes the assumption that the point's color is also the background color. It then looks at each individual point (x, y) and the points' neighbors $(x+1, y)$, $(x-1, y)$, $(x, y+1)$, $(x, y-1)$. If any of the points' neighbors are the same color as $(0, 0)$, then that means that (x, y) is touching an edge. For example, the black squares at $(1, 1)$, $(2, 1)$, and $(1, 2)$ would be considered edges. The bottom right-image shows the graph with edges highlighted.



Shape recognition has yet to be implemented.

Expected Results: With the edges known, it will now be possible to group the points into axis, number, and point categories. Once the center of each "point" is found, it will be possible to determine the "point's" relation to the axis.

