

## HW 6: due Thursday, November 12 in class

For this assignment, submit your answers as problems as a single pdf file named hwk6.pdf, alongside any code you may have written and required additional text files, using provide with the command:

```
homework% provide comp150ns hwk6 hwk6.pdf p3.txt code-you-wrote
```

For this problem, you will be using a two data sets:

- The Karate Club network, a classic example of community structure in networks.
- PPI, a graph of physical protein-protein interactions in yeast.

1. Implement the Louvain method for community detection.
2. Run your implementation of Louvain on the Karate Club network. List the clusters you obtain. How do these clusters compare to the eventual split of the club? What, if anything, does this say about the validity of your clusters?
3. Run Louvain on the PPI network. How might you evaluate these clusters using the MIPS annotation for individual proteins? Describe your scoring metric, and provide the clusters you obtain and their corresponding scores as p3.txt.

Note: data available at <http://www.cs.tufts.edu/comp/150NS/private/datasets.html>