Discovery of Huffman Codes

Notes to the instructor

The project is designed for a junior level Data Structures and Algorithms course.

Story of the invention of Huffman codes is a great story that demonstrates that students can do better than professors. David Huffman was a student in an electrical engineering course in 1951. His professor, Robert Fano, offered students a choice of taking a final exam or writing a term paper. Huffman did not want to take the final so he started working on the term paper. The topic of the paper was to find the most efficient (optimal) code. What professor Fano did not tell his students was the fact that it was an open problem and that he was working on the problem himself. Huffman spent a lot of time on the problem and was ready to give up when the solution suddenly came to him. The code he discovered was optimal, that is, it had the lowest possible average message length. The method that Fano had developed for this problem did not always produce optimal code. Therefore, Huffman did better than his professor. Later Huffman said that likely he would not have even attempted the problem if he had known that his professor was struggling with it.

The project uses quotes from Fano's and Huffman's papers to look into greedy algorithms, prefix-free codes, Huffman encoding, binary tree representations of codes, unit and amount of information. Both Fano and Huffman encodings are produced by greedy algorithms. However, in Fano's approach the greedy algorithm does not find an optimal solution while in Huffman's version it does.

The project is divided into several parts. Each part contains a reading assignment (read selected quotes from the original sources and some additional explanations) and a list of tasks on the material from the reading.