# The Shortest Path Approach to Optimal Prefix-Free Codes for Unequal Letter Costs

Wojciech Rytter

**Abstract**

We show how the construction of certain optimal trees is reduced to the computation of a shortest paths in graphs with edge weights satisfying so called *Monge property*. The *Monge property* emerges, sometimes quite mysteriously, in many optimization problems and usually leads to running time improvements of at least a linear factor. A novel application of *memory reduced dynamic programming* combined with a Monge property and the SMAWK algorithm is given for a classical variation of the *Huffman coding problem*: optimal prefix-free codes for unequal letter costs. The problem consists of finding a minimal cost prefix-free code in which the encoding alphabet consists of unequal cost (length) letters, of lengths $\alpha$ and $\beta$. The problem is much more complicated than the Huffman coding and for general $\alpha$ and $\beta$ it is not even known if it is $\mathcal{NP}$-complete; also no polynomial time algorithm is known. There are quite efficient algorithms known, though, that are polynomial for *fixed* integral $\alpha$, $\beta$. Using a new approach the time complexity of the fastest known algorithms for these cases can be improved by a factor of $O(n^2)$. One linear factor is eliminated by limiting the space of subproblems (reducing the dynamic programming tables); an additional linear factor is eliminated by employment of the algorithm known as the SMAWK algorithm, for finding all row minima of a totally monotone 2-dimensional matrix in linear time. For example, for 1,2 trees, the best previously known algorithm works in $O(n^4)$ time, and our improvement gives an $O(n^2)$ time algorithm. The reduction of the space of subproblems is achieved by omitting some essential information related to partial subtrees. This is the most difficult and unorthodox part of the algorithm and it strongly affects the correctness proof which is unexpectedly complicated. The computed objects correspond to certain paths in a graph. However, due to the omission of essential information, some paths correspond to anomalous objects. The main point is that *shortest* paths correspond to good objects; correctness relies on deriving several new interesting combinatorial properties of lopsided trees.

**Joint work with Phil Bradford, Mordecai J. Golin and Lawrence L. Larmore.**