Combinatorial Mappings From Labelled Trees and Strings

Saverio Caminiti *

Labeled trees are of interest in practical and theoretical areas of computer science. Due to the growing interest in developing efficient tools for manipulating XML trees brings back labeled trees to the forefront of research.

An intriguing alternative to the usual representation of tree data structures in computer memories is based on coding labeled trees by means of node label strings.

This representation was first used by Prüfer in the proof of Cayley's theorem to show a one-to-one correspondence between unrooted labeled trees on n nodes and strings of length $n-2$. Later on, several other encodings have been introduced. We will consider two classes of encodings: the so called Prüfer-like codes and the Picciotto's codes.

Prüfer-like codes include those introduced by Neville, Moon, Deo, and Micikevičous. For codes belonging to this class we show how coding and decoding can be reduced to integer (radix) sorting. This approach easily yields to optimal linear time algorithms.

Picciotto's codes are based on works by Orlin, Knuth, Joyal, Eğecioğlu, and Remmel. Optimal algorithms for these codes are obtained through a general scheme to define bijective codes based on the transformation of a tree into a functional digraph. This general scheme also provides a theoretical instrument to understand former experimental results. Indeed, it has been observed that Picciotto's codes satisfy, much better than Prüfer-like codes, two properties (locality and heritability) fundamental to successfully use labeled tree codes in genetic algorithms.

These codes can also be exploited to efficiently generate random tree with several possible constrains.

*Computer Science Department, Sapienza University of Rome, Via Salaria 113, 00198 Roma, Italy. E-mail: caminiti@di.uniroma1.it.