The Goulden–Jackson cluster method for monoid networks and an application to lattice path enumeration

Yan Zhuang
Department of Mathematics
Brandeis University
zhuangy@brandeis.edu

Given a finite or countably infinite set A, let A^* be the set of all finite sequences of elements of A, including the empty sequence. We call A an *alphabet*, the elements of A *letters*, and the elements of A^* *words*. By defining an associative binary operation on two words by concatenating them, we see that A^* is a monoid under the operation of concatenation, and we call A^* the *free monoid* on A.

The combinatorial framework of free monoids can be generalized using what are called "monoid networks". A monoid network consists of a digraph G with each arc assigned a set of letters from an alphabet A. Monoid networks were first introduced by Gessel [1] in an equivalent form called "G-systems", and they were recently reintroduced in their current form by Zhuang [3] to solve problems in permutation enumeration. These constructions are closely related to finite-state automata and the transfer-matrix method.

The Goulden–Jackson cluster method allows one to determine the generating function for words in a free monoid A^* by occurrences of words in a set $B \subseteq A^*$ as subwords in terms of the generating function for what are called "clusters" formed by words in B, which is easier to compute. As its name suggests, this celebrated result was first given by Goulden and Jackson [2], but has seen a number of extensions and generalizations over the years. We give a generalization of the cluster method for monoid networks, which gives a way of counting words in A^* corresponding to walks between two specified vertices in G by occurrences of subwords in a set B.

The monoid network cluster method can be used to count lattice paths by occurrences of subwords, and here we focus on Motzkin paths. We consider both regular Motzkin paths and Motzkin paths bounded by height, and our results include bivariate and multivariate generating functions for these paths by ascents, plateaus, peaks, and valleys—all of which are statistics that are determined by occurrences of various subwords in the underlying word of the Motzkin path—as well as generating functions for Motzkin paths with restrictions on the heights at which these subwords can occur, yielding both closed-form and continued fraction formulae.

References

[1] Ira Martin Gessel. *Generating Functions and Enumeration of Sequences*. PhD thesis, Massachusetts Institute of Technology, 1977.

[2] I. P. Goulden and D. M. Jackson. An inversion theorem for cluster decompositions of sequences with distinguished subsequences. *J. London Math. Soc. (2)*, 20(3):567–576, 1979.

[3] Yan Zhuang. Monoid networks and counting permutations by runs. Preprint, 2015.