Automated Proof and Discovery in Three Combinatorial Problems

Ph.D. Thesis Defense

Paul Raff

Department of Mathematics
Rutgers University

August 14, 2009
For Partial Fulfillment of Ph.D. Requirements

Automated Proof and Discovery in Three Combinatorial Problems

Ph.D. Thesis Defense

Paul Raff

Department of Mathematics
Rutgers University

August 14, 2009
For Partial Fulfillment of Ph.D. Requirements

Outline

Avoiding Differences

Starting and ending with The Triangle Conjecture

- We will investigate the quantity $f_\Delta(n)$, defined by

$$f_\Delta(n) = \max\{|X| \mid X \subseteq [n] \text{ and } X \text{ avoids differences in } \Delta\}.$$

Avoiding Differences

Starting and ending with The Triangle Conjecture

- We will investigate the quantity $f_\Delta(n)$, defined by

$$f_\Delta(n) = \max\{|X| \mid X \subseteq [n] \text{ and } X \text{ avoids differences in } \Delta\}.$$

- Motivated by the Triangle Conjecture of Schützenberger and Perrin.

Avoiding Differences

Starting and ending with The Triangle Conjecture

- We will investigate the quantity $f_\Delta(n)$, defined by

$$f_\Delta(n) = \max\{|X| \mid X \subseteq [n] \text{ and } X \text{ avoids differences in } \Delta\}.$$

- Motivated by the Triangle Conjecture of Schützenberger and Perrin.

Result Enumeration schemes for computing $\{f_\Delta(n)\}|_{n=1}^{\infty}$ and *proving* its behavior.

Result An asymptotic version of the Triangle Conjecture.

Spanning Trees in Grid Graphs

Graphs of the form $G \times P_n$ or $G \times C_n$

- We will extend the methods used by Desjarlais and Molina to compute the sequence $\{\tau_G(n)\}_{n=1}^{\infty}$, where

$$\tau_G(n) = \text{number of spanning trees in } G \times P_n.$$

- Enumeration schemes that calculate and *prove* full information:
  - Recurrence
  - Generating function
  - Closed-form formula.

Result Spanning tree sequences are *divisibility sequences*.

The Firefighter Problem

On $\mathbb{Z} \times \mathbb{Z}$

- We will introduce the problem.

Introduction

The Triangle Conjecture

- The Triangle Conjecture deals with *codes*, which are subsets C of the set

$$\mathcal{A}_m = \{x^i y x^j \mid i + j \leq m\}$$

where $C^\star$ exhibits unique factorization.

Example

$\{y, xy, yx\} \subseteq \mathcal{A}_2$ is *not* a code, for

$$yxy = y \cdot xy$$
$$yxy = yx \cdot y$$

Conjecture (Schützenberger-Perrin 1980)

If $C \subseteq \mathcal{A}_m$ is a code, then $|C| \leq m$.

- Shortly afterward, Shor exhibited a code $C \subseteq \mathcal{A}_{15}$ with 16 elements.

Important Point

Shor's counterexample relied on finding large sets avoiding prescribed differences.

Definition

Generally, $f_\Delta(I;S)$ is the *size* of the largest subset of I that avoids differences in Δ and elements in S. Additionally, $f_\Delta(I) = f_\Delta(I;\emptyset)$ and $f_\Delta(n;S) = f_\Delta([n];S)$.

Lemma (R.)

If $1 \in S$ *then*

$$f_\Delta(n;S) = f_\Delta(n-1;S-1)$$

otherwise,

$$f_\Delta(n;S) = \max\{f_\Delta(n-1;S-1), 1 + f_\Delta(n-1;\Delta \cup (S-1))\}.$$

- Given Δ, S, the number of different parameters needed in the enumeration scheme to computer $f_\Delta(n; S)$ is finite.

Theorem (R.)

Given any finite Δ, S, the sequence $(f_\Delta(n; S))$ is eventually pseudoperiodic.

- A theorem-prover proving the structure of $(f_\Delta(n; S))$ has been implemented.

Definition

$$\mu(\Delta) = \lim_{n \to \infty} \frac{f_\Delta(n)}{n}.$$

$\mu(\Delta)$ is rational.

Theorem (R. - Asymptotic Version of Triangle Conjecture)

$$\mu(X - X) \leq \frac{1}{|X|}.$$

Proof.

Let $X = \{x_1, x_2, \ldots, x_k\}$ and consider

$$\{x_1 + 0, x_2 + 0, \ldots, x_k + 0\}$$
$$\{x_1 + 1, x_2 + 1, \ldots, x_k + 1\}$$
$$\{x_1 + 2, x_2 + 2, \ldots, x_k + 2\}$$
$$\vdots$$

To avoid differences, we can only have one element from each set. Each $n \in \mathbb{N}$ is represented at most $k (= |X|)$ times in this family of sets. □

History

- The Matrix Tree Theorem will compute the number of spanning trees of any graph G.
- The Matrix Tree Theorem does *not* provide any information about the number of spanning trees of an infinite family of graphs.

History
(continued)

- Desjarlais and Molina created an enumeration scheme to compute $\tau_{P_2}(n)$, the number of spanning trees of $P_2 \times P_n$.

- To compute $\tau_{P_2}(n)$, they also computed $\tau'_{P_2}(n)$, defined as the number of spanning *forests* of $P_2 \times P_n$ with the special property that the two vertices on the right end are in different components.

This yields the enumeration scheme

$$\tau_{P_2}(n) = 3\tau_{P_2}(n-1) + \tau'_{P_2}(n-1)$$
$$\tau'_{P_2}(n) = 2\tau_{P_2}(n-1) + \tau'_{P_2}(n-1)$$

History
(continued)

- From this, they deduced that $\tau_{P_2}(n)$ satisfies the recurrence

$$\tau_{P_2}(n) = 4\tau_{P_2}(n-1) - \tau_{P_2}(n-2)$$

with the initial conditions $\tau_{P_2}(1) = 1$, $\tau_{P_2}(2) = 4$. $\tau'_{P_2}(n)$ also satisfies the same recurrence but $\tau'_{P_2}(2) = 3$.

- We generalize and formalize this framework.

Contributions

- This can be extended to a formal framework if we consider the enumeration scheme that counts, for a graph G on n vertices, *all* values $a_G(n; P)$ for all partitions P of $[n]$.
- We must compute, for all P and P', the number of different ways we can append edges to *transition* from a spanning tree represented in $\tau_G(n; P)$ to one represented in $\tau_G(n; P')$. This naturally yields a matrix, A_G.

Example

If $G = K_3$ our transition matrix is

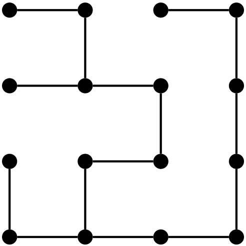
$$\begin{bmatrix} 16 & 8 & 8 & 8 & 3 \\ 4 & 3 & 2 & 2 & 1 \\ 4 & 2 & 3 & 2 & 1 \\ 4 & 2 & 2 & 3 & 1 \\ 3 & 2 & 2 & 2 & 1 \end{bmatrix}$$

Results

- Full sequence information for all graphs up to 5 vertices, with plans to find all information for all graphs on 6 vertices (about 25% complete).
- Interesting conjectures:
  - Characteristic polynomial always factors into polynomials of degrees a power of 2.
  - Coefficients of characteristic polynomial alternate in sign – suggests potential reformulation in terms of Inclusion-Exclusion.
  - Recurrence of minimum order for $P_k \times P_n$ has order 2^{k-1}.
  - Recurrence of minimum order for $K_k \times P_n$ has order k.

Spanning Tree Sequences are Divisibility Sequences

A spanning tree of $G \times P_{2n}$ can be split into three parts: a left tree, a right tree, and the middle edges.



COMP(P,MID) is the set of partitions that is compatible with the left-hand partition and the middle edges.

Lemma (Split-Merge Lemma)

$$\sum_{\text{MID} \in \binom{[\nu]}{k}} \sum_{P \in \mathcal{P}_\nu(p)} \tau_G(n;P) \sum_{P' \in \text{COMP}(P,\text{MID})} \tau_G(n;P')$$

$$= \binom{k-1}{p-1} \tau_G(n) \sum_{P \in \mathcal{P}(e)} \left(\prod P\right) \tau_G(n;P).$$

Given a graph G, a fire is placed at a specified vertex and at discrete time intervals $t \geq 0$, $f(t)$ firefighters are placed on unoccupied vertices, and then the fire spreads to adjacent vertices that are not protected nor already on fire.

Question

If G is finite, can the fire be contained with vertices that are neither on fire nor protected? If G is infinite, can the fire be contained?

Theorem (Wang-Moeller)

If $f(t) = 1$, then no fire can be contained in the two-dimensional grid. If $f(t) = 2$, then any finite fire can be contained in the two-dimensional grid.

Theorem (Ng-R.)

If f is periodic and the average number of firefighters per turn is $1.5 + \varepsilon$, then any finite fire can be contained in the two-dimensional grid.

This theorem can be extended further to deal with non-periodic functions.

Theorem (R.)

The function f defined by

$$f(t) = \begin{cases} 3 & \text{if } t \text{ is odd} \\ 0 & \text{if } t \text{ is even} \end{cases}$$

can not give a convex solution to a point fire in the two-dimensional grid.