Report on REGS on Extremal Problems in Combinatorics — D. B. West

Summer 2004 and Beyond

Overview

This summer I ran a large REGS group involving collaborative student research efforts on extremal problems in combinatorics. The group involved graduate students at various levels from both the Mathematics Department and the Computer Science Department, plus two undergraduates, for a total of about 20 students. The project was very successful, providing research experiences for many beginning students, mentoring relationships, and several publishable results.

Combinatorial problems lend themselves particularly well to such a project, because many of them can be described in the terminology of a single introductory course and understood by beginning students. Also, the questions can be studied on examples to gain understanding, and proving a conjecture on a special subclass of instances can be of great interest.

The group met three times per week, nominally from 2–4pm, but usually some of the students stayed till about 5pm, so I was meeting with students about nine hours per week once the research was underway. During the first two or three meetings, including one before the summer term started, I briefly described more than a dozen problem areas. Students selected from these, and in the second phase they gave presentations describing some of what was known about the problem they chose. They prepared these talks on their own, mostly on the basis of library sources. Most of us went to a conference during the first week of summer term, so others used that time to prepare their talks.

For a couple of weeks, we had two such presentations per meeting. Gradually I coaxed the students into working on examples of the problems presented, and several working groups formed. Some of the original problems were discarded as efforts coalesced around problems that students found appealing and accessible.

By mid-July, there were no more presentations. Working sessions involved students collaborating in groups of one to four students. I moved from group to group, listening to what had been discovered and making suggestions. The results are being presented in our regular combinatorics research seminar during the fall. Subsequently there will be talks given at regional meetings, and papers will be published.

The principle behind this project is the research analogue of the collaborative study sessions used in the undergraduate graph theory course (Math 412) and by some instructors in the undergraduate "transition to proofs" course (Math 347). Small groups in the same room facilitate efficient contact for the instructor/advisor, many students can be talking and thinking independently at the same time, students engaged in discussion together feel more comfortable speaking up and contributing their ideas, etc.

I expect that three or four papers will result from this summer's activities. The project was so successful that we are continuing it this fall, with the participation of Prof. Kostochka, our new post-doc Stephen Hartke, and 15–20 students, including some new students. Since everyone has more duties now than during the summer, we meet only once each week, 2–4pm on Thursdays, separate from our long-running research seminar at 3pm on Tuesdays.

Personnel

In this REGS group, several of the younger graduate students from Mathematics were supported with VIGRE funds. This included Qi Liu, Lale Ozkahya, Jennifer Vandenbussche, Joseph Wright, and Gexin Yu. Two undergraduates, Noah Prince and Derrick Cheng, received REU funding to participate. Some more senior graduate students participated with support as graders for summer courses, such as Hailong Hu, Jeong-Ok Choi, and Weiting Cao. More senior students who participated occasionally included Naeem Sheikh, Kyung-Won Hwang, Hemanshu Kaul, Jeong-Hyun Kang, Kittikorn Nakprasit, and Seog-Jin Kim (the last three had defended their theses). The regular participants from Computer Science were David Bunde, Kevin Milans, Erin Wolf, and Dan Cranston.

Some of these investigations may lead to thesis research by the early-career graduate students who were involved. Short reports on their experiences were submitted by Choi, Hu, Liu, Ozhahya, Yu, and Vandenbussche.

Summary of Results

Progress was made in several areas.

Pagenumber of graphs. When laying out circuits on a chip, crossing wires must go on separate layers; we wish to minimize the number of layers. Given a graph, we order the vertices linearly, as on the spine of a book, and embed the edges on "pages". Edges whose endpoints alternate along the spine must go on different pages, since on a single page thee would cross. The minimum number of pages used, over all vertex orderings and assignments of edges to pages, is the *pagenumber* or *book thickness* of the graph, written $\mathrm{bt}(G)$.

One special family proposed for study long ago by Leighton (due to the motivation from VLSI) is the cartesian product of complete graphs, $K_n \,\square\, K_n$. There seems to have been little or no progress on this. Q. Liu, N. Prince, and J. Vandenbussche began preliminary study of the behavior of pagenumber under cartesian products. The modest objective so far is to understand $\mathrm{bt}(G \,\square\, K_2)$. An upper bound of $\mathrm{bt}(G) + 1$ arises using two copies of an optimal embedding of G on the same pages (with opposite vertex orders), plus one more page for the copies of K_2. Also $\mathrm{bt}(G \,\square\, K_2) \geq \mathrm{bt}(G)$, since $G \subseteq G \,\square\, K_2$. The students obtained some technical sufficient conditions for $\mathrm{bt}(G \,\square\, K_2)$ to equal either $\mathrm{bt}(G)$ or $\mathrm{bt}(G) + 1$.

The kth power of the path P_n is the graph P_n^k with vertices $v_1, \ldots, v_n$ such that $v_i v_j$ is an edge if and only if $|j - i| \leq k$. This is the unique largest n-vertex graph with bandwidth k, where the *bandwidth* of a graph is the minimum, over all labelings of the vertices by distinct integers, of the maximum difference between adjacent labels. Assigning to page i all the edges whose lower endpoint has index congruent to i modulo k yields an embedding on k pages, but in fact $\mathrm{bt}(P_n^k) = k - 1$.

The graph P_n^k is an example of a *k-tree*, a graph that can be obtained from K_k by iteratively adding one vertex whose neighborhood in the current graph is a clique. *Treewidth* is an important structural parameter in modern graph theory; one characterization is that the graphs with treewidth k are precisely the subgraphs of k-trees. Ganley and Heath proved that k-trees, and hence graphs with treewidth k, have pagenumber at most $k + 1$, and they conjectured that the best possible bound is k. Togasaki and Yamazaki proved that k is the maximum value for graphs with "pathwidth" k, a class intermediate between bandwidth k and treewidth k (the complete bipartite graph $K_{mk,k}$ for large m achieves the bound).

Students J. Vandenbussche and G. Yu proved that the Ganley–Heath conjecture is false, at least for $k = 3$. Iterated pigeonholing techniques yield a construction of large 3-trees that cannot be embedded in 3 pages. It appears that the idea can be extended to construct k-trees that do not embed on k pages; the students are continuing to work on this now. They also developed an algorithm that produces embeddings on 3 pages for a class of 3-trees that properly contains those treated by Togasaki and Yamazake.

Pebbling of graphs. Graph pebbling models the movement of resources in a network, where moving the resources consumes some of the resources, such as fuel. In the classical model, we begin with a distribution of pebbles at the vertices of a graph. A *pebbling move* removes two pebbles from one vertex and moves one of them to an adjacent vertex; the other pebble is consumed. The *pebbling number* of a graph G is the minimum t such that *every* distribution of t pebbles on the vertices of G permits a pebble to be moved to any target vertex. The *optimal pebbling number* is the minimum t such that *some* distribution of t pebbles permits this. Graph pebbling is currently a “hot topic”. Glenn Hurlbert maintains a website devoted to results on pebbling.

This summer, students David Bunde, Bryan Clark, Dan Cranston, Kevin Milans, and Erin Wolf proved a number of complexity results about pebbling and extremal results about optimal pebbling.

Among the former is a characterization of when a given set of pebbling moves can be ordered into a valid sequence of pebbling moves. As a consequence, when designing sequences of pebbling moves, one may focus on choosing which pebbling moves to make as opposed to the order in which to make them. These leads to several NP-completeness results.

Milans obtained a short probabilistic proof of a lower bound on the optimal pebbling number of the k-dimensional hypercube. It turns out that the bound was already known, but this is a much simpler proof.

With contributions from many participants, we proved this summer that the maximum value of the optimal pebbling number of an n-vertex graph is $\lceil 2n/3 \rceil$, achieved by paths and cycles. The values for paths and cycles were known before, but the students simplified the proofs. The upper bound is new.

Among the questions still being explored is the maximum of the optimal pebbling number for n-vertex graphs with minimum degree at least k (conjectured to be about $2n/(k+1)$ when $k \geq 2$), algorithmic results for trees, other open problems relating pebbling parameters to connectivity, etc.

Graph decomposition. A *decomposition* of a graph is an expression of it as a union of pairwise edge-disjoint subgraphs. Given a family $\mathbf{F}$ of graphs, the $\mathbf{F}$-*thickness* is the minimum number of subgraphs in a decomposition of G into subgraphs belonging to $\mathbf{F}$.

Fan Chung showed that every connected n-vertex graph decomposes into $\lceil n/2 \rceil$ trees. This is sharp, because each tree uses at most $n - 1$ edges. Decomposition results can be studied in more detail by seeking tighter upper bounds over smaller families or by narrowing the set of trees used for decompositions, which may require more subtrees.

In fact, Chung’s inductive decomposition construction uses very restricted trees: caterpillars of diameter at most four. Thus restricting to caterpillars does not change the extremal result. Indeed, the famous conjecture of Gallai is that every connected n-vertex graph can be decomposed into $\lceil n/2 \rceil$ paths, so Chung’s result obtaining this bound using caterpillars can be viewed as a partial result in that direction. This suggests studying tree thickness and caterpillar thickness on classes of sparse graphs.

Students Q. Liu and D. Cheng studied extremal problems for these parameters under embeddability and/or minimum girth conditions. Only connected n-vertex graphs are considered.

When G is an n-vertex tree, $\theta_{\mathbf{C}}(G) \le \lceil n/4 \rceil$, with equality for the tree having $\lfloor n/2 \rfloor$ three-vertex paths with a common endpoint. Beyond trees are cacti: connected graphs in which every edge appears in at most one cycle. When G is a cactus, $\theta_{\mathbf{T}}(G) \le \lceil n/3 \rceil$. However, $\theta_{bC}(G)$ may be as high as $\lceil n/2 \rceil$ which is already the maximum over all n-vertex connected graphs. When $n \equiv 1 \bmod 3$, the cactus $H_{n,3}$ achieving both bounds consists of $(n-1)/3$ pairwise disjoint triangles plus a single vertex having an edge to each triangle.

Classes where the restriction to caterpillars does not force the answer up to $n/2$ arise from a girth restriction. For triangle-free outerplanar graphs, $\theta_{\mathbf{C}}(G) \le \lceil 3n/8 \rceil$, achieved by the cactus $H_{n,4}$ consisting of disjoint 4-cycles plus one vertex having a neighbor in each 4-cycle.

A girth restriction also yields an interesting result for tree thickness: If G has girth g, where $g \ge 5$, then $\theta_{\mathbf{T}}(G) \le \lfloor n/g \rfloor + 1$. When the girth is 4, the arguments break down, and the results are weaker. When G is outerplanar, $\theta_{\mathbf{T}}(G) \le \lfloor n/4 \rfloor + 1$, as expected, with equality for $H_{n,4}$. In this and the results above, the upper bounds arise from inductive constructive algorithms.

When all planar triangle-free graphs are allowed, the best available upper bound for tree thickness is $\lceil n/3 \rceil$. In fact, this upper bound holds more generally, for all triangle-free graphs that do not contain the subgraph $K_{3,3}$. However, the best available lower bound is $\lceil n/4 \rceil$, from the cactus $H_{n,4}$. Thus the extremal problems for tree thickness of planar graphs and tree thickness of planar triangle-free graphs remain open.

Another open problem arises for caterpillar thickness of triangle-free planar graphs. It seems that the answer should be again $\lceil 3n/8 \rceil$, achieved by $H_{n,4}$, but the upper bound has not yet been proved.

Since the extremal examples for these problems tend to be cacti with cut-edges, another set of problems arises by requiring the input graphs (families of planar or outerplanar graphs under various girth restrictions) to be 2-edge-connected or 2-connected. The students are studying these additional problems this fall.

Degree Splitting. An n-vertex graph G is *self-complementary* if it is isomorphic to its complement, meaning that the complete graph K_n decomposes into two copies of G. More generally, one can ask whether a graph H with an even number of edges decomposes into two isomorphic subgraphs. Jamison and Stevens provided some sufficient conditions for such isomorphic decomposition of various trees.

Motivated by questions about 3-dimensional polytopes, Malkevitch suggested a weaker notion of “isomorphic-degree” factorization. Say that a *splitting* of G is a decomposition of G into two spanning factors that have the same degree list.

Various necessary conditions come quickly to mind. If G is splittable, then no odd degree exceeds twice the maximum of the other vertex degrees. Also, the number of odd degree vertices plus twice the number of even degrees that are split into two odd numbers in the subgraphs must be a multiple of 4. When splitting a tree, various conditions arise from the requirement that the two subgraphs must have the same number of components.

Even when the problem is restricted to caterpillars, characterizing those that are splittable turns out to be rather difficult. Students W. Cao, J. Choi, and L. Ozkahya have studied this problem. They have broken the question into a numerical step involving solutions of linear equations to obtain the vertex degrees of the subgraph and a structural step in which the edges are distributed to achieve those degrees.

Cut-and-paste sorting. Given a permutation of $[n]$ in word form, we are allowed the operation of cutting out a string and reinserting it somewhere else, possibly reversed. In the worst case, how many such operations are needed to produce the identity permutation?

Student Dan Cranston (a thesis student of West although he is now in the CS department) found an algorithm that sorts every permutation of n numbers in at most $\lceil 2n/3 \rceil$ such steps. He also gave an elegant parity argument to show the existence of a wide class of permutations needing at least $n/2$ such steps.

A later literature search turned up a claim that these results appear in unpublished work of Sudborough. Efforts are ongoing to contact him to determine whether he does have the same results.

Forced cliques in non-majorizable graphs. The *Turán graph* $T_{n,r}$ is the n-vertex graph whose vertices are partitioned into r sets of sizes differing by at most 1 so that vertices are adjacent if and only if they belong to distinct sets in the partition. Turán's Theorem, the archetypal result in extremal graph theory, states that $T_{n,r}$ is the unique n-vertex graph with most edges among those not containing the complete graph K_{r+1}.

Among the many proofs of Turán's Theorem, the proof by Erdős shows that an n-vertex graph G not containing K_{r+1} is *degree-majorized* by an r-partite graph H, meaning that $d_i \le d'_i$ for $1 \le i \le n$, where $d_1, \ldots, d_n$ and $d'_1, \ldots, d'_n$ are the vertex degrees of G and H, respectively, in nonincreasing order.

Various authors have studied the number of copies of K_{r+1} that are forced to appear when the number of edges has a particular value larger than the number of edges in $T_{n,r}$. Many years ago, West proposed the analogous question for degree-majorizability: When G is an n-vertex graph with maximum degree k that is not degree-majorizable by an r-partite graph, how many copies of K_{r+1} must appear in G? Student Joe Wright worked on extending West's results of the mid-80s on this question.

$L(2,1)$-labeling. A more refined notion of graph coloring that has applications in radio-frequency assignment is to require vertices that are close together to receive colors that are far apart. An *$L(d_1, d_2, \ldots)$-labeling* is a vertex coloring using natural numbers, such that the colors on vertices at distance i apart receive colors that differ by at least i. In particular, in an $L(2,1)$-labeling, colors on adjacent vertices differ by at least 2, and colors on vertices with a common neighbor must be distinct. The objective is to minimize the global difference between the largest and smallest colors used.

The big conjecture in this subject is by Griggs and Yeh, stating that a graph with maximum degree k has an $L(2,1)$-labeling with span at most k^2. In her recent thesis, Jeong-Hyun Kang proved this for 3-regular graphs that have spanning cycles.

For trees with maximum degree k, the minimum span is always k or $k+1$. Students W. Cao, J. Choi, L. Ozkahya, and S.-J. Kim tried to prove a characterization of which trees have which value. They made progress toward this, but at some point we received an email from South Africa from an individual who claimed to have done this. However, we have not been able to obtain the paper.