Mathematics 357 – Combinatorics
Final Project Information
March 7, 2017

General Information

As announced in the syllabus given out at the beginning of the semester, it will be possible to do a final project for the Combinatorics course that will replace a final examination and count for the Project Course requirement for the Mathematics major. According to the preferences people submitted at the beginning of the semester, some people wanted to do a final exam instead of a project. I expect that is still true, so there will still be a final exam given at the scheduled time for our class.

For those doing a project, the final product will be paper of about 15 pages and an oral presentation about 15-20 minutes in length. Because a rather large majority of the class has chosen to do projects, I have decided to alter the course schedule and devote the final class meetings (May 1,3,5,8) to the project presentations. Everyone is expected to attend, and I may ask a (small portion of a) question on the final related to something in one of the presentations.

Several suggested topics are given below. All of them lead to other interesting results and applications of combinatorial mathematics. Many of them deal with aspects of the subject outside of the strictly *enumerative* combinatorics focus of Beeler's textbook. I will also be happy entertain any ideas you might have designing a project topic of your own. If there is some combinatorial subject you are interested in, or that makes contact with another course you have taken, please do not hesitate to discuss it with me and get my approval before you start to work.

You will work in groups of two on these projects, or individually if you would prefer. If you need help putting a group together, be sure to talk to me *well before March 20.*

Schedule, Deadlines, Groundrules

- **On or before Monday, March 20** – Inform me *by email* which general topic you want to work on and with whom you will be working. Ideally, each group will work on a different topic, although as noted below the areas are large enough that if more than one group wants to try that, there will be ways to "split up" the topic into several parts. We can discuss the possibilities before this is due if you would like.
- **Friday, April 7** – Each group will submit, *by email*, an *annotated bibliography* of the sources to be used for your project. You should identify *at least five books, articles, or web sites, including no more than two web sites*, that will be relevant. *Important Note: I have no problem if you want to consult Wikipedia pages to get a general overview of a topic or to learn the definition of an unfamiliar term. BUT, Wikipedia articles are NOT suitable items for your bibliography – they do not go into the details in as much depth as you will need.* There are a few general references on reserve for the class in the Science Library – I have included pointers to those in the project topic descriptions below. Those are fine as items in your bibliography. For each of your sources, write up a short paragraph giving a rough description of how that source relates to your main

topic, what kind of information you will take from it, and how you will be using it (including a preliminary estimate of how reliable you think the information there is). The project descriptions below contain some first places to look, but you should plan to spend some time searching for additional sources. This annotated bibliography will count for 10% of the final project grade.

- **Between April 10 and April 21** – Each team will meet with me during office hours (or at another time if that is not convenient) for a progress report and a chance to discuss any questions that have come up as you have started to work on the project. I will be happy to meet to talk over any aspect of the project at other times too, of course.
- The oral presentation will receive a separate grade, and will count for 30% of the total on the final project.
- **Monday, May 8** – All final project papers will be due (via email) by 5:00pm. The grade for the paper component will account for 60% of the total on the final project. **This is an** ***absolutely firm*** **deadline–no extensions will be possible, so please don't ask. If you don't leave things until the last minute, there should be no difficulty in meeting this schedule.**

In working on this paper, you should follow the same procedures you would follow in preparing a research paper for any other course, and of course the College Policy on Academic Honesty applies here, as it does to all of your work. Your grade will depend on the thoroughness of your research, the degree of independent thought about the subject revealed through your work, the organization of the paper, and the quality of your writing. Do not assume that because this is a mathematics paper writing does not matter! A word to the wise: I tend to be rather severe when I read papers with large numbers of misspellings, or many grammatical and other technical errors. The same thing happens with papers consisting largely of masses of undigested quotations from your sources, particularly when those quotations involve technical discussions of mathematics you do not understand fully. To avoid having this happen to you, take me up on the following offer: I will be happy to read a preliminary draft of your paper and give you comments. (Of course, this means that the writing *absolutely cannot* be left until the afternoon of May 8!)

Your papers should be word-processed, using one side of the paper only, double-spaced. Equations can be entered by hand if necessary. You can also use the MS Word Equation Editor if you have it. Papers may also be typeset using TeX or LATeX (the software I use to produce course handouts). This has a "learning curve" but it produces the best-looking typeset mathematical formulas. Your paper should include a bibliography listing all the sources you consulted. Direct quotations should be identified with foot- or end-notes. If you don't understand all the technical details of an argument, ask me about it, and we can work through it.

Suggested Topics

I. Hall's "Marriage Theorem" and Matrix Applications

The first group of topics here deals with a basic combinatorial result that guarantees that arrangements exists, rather than counting how many of them there are. This theorem is often explained by the following sort of (dated and sexist, I know!) analogy. In a group of seven boys and six girls of marriageble age,

- girl 1 knows boys a,b,c
- girl 2 knows boys b,c
- girl 3 knows boys c,e,g
- girl 4 knows boys a,b
- girl 5 knows boys a,b,c
- girl 6 knows boys d,e,f

Question: Is it possible to find each of the girls a husband from among the boys she knows? (Note: the husbands have to be distinct boys(!) – take a moment to think about this before reading on!) It is not hard to see that the answer is *no* here because the four girls 1,2,4,5 know only three different boys (a,b,c) between them. The general problem is: Given a collection of finite sets $A_1, \ldots, A_n$, under what conditions is it possible to find a *transversal* of the collection, namely a collection of elements $x_i \in A_i$, $i = 1, \ldots, n$ with the x_i all distinct? The result is that this is true if and only if for each $S \subseteq [n]$,

$$\left| \bigcup_{k \in S} A_k \right| \geq |S|.$$

There is also a "matrix form" of Hall's theorem that deals with $0,1$-matrices. The translation is pretty direct and it leads to a way to split up a $0,1$-matrix as a sum of the *permutation matrices* we encounted earlier in the semester. For instance, if every row and column sum in a $0,1$-matrix A is d, then the A is the sum of d permutation matrices. A project here could consist of learning a proof of this result, then looking at any of the following variants or applications dealing with matrices.

A. A *doubly stochastic* matrix is one where the entries are all non-negative and each row and column sums to 1. (These appear in the context of Markov Chains in probability theory – see the next topic below.) Among the interesting properties of these matrices is that they are all *convex linear combinations of permutation matrices.* In other words, if A is doubly-stochastic, then

$$A = \lambda_1 P_1 + \cdots + \lambda_k P_k$$

where the λ_i are non-negative reals with $\lambda_1 + \cdots + \lambda_k = 1$ and the P_i are permutation matrices. This result is one that you should try to *prove yourself* before looking up a proof(!)

B. A more applied project here would be look up the *Markov chains* mentioned in the previous description, especially the *discrete-time case*, and understand how they are used to model stochastic processes where the transition probabilities are constant. This is less combinatorics and more linear algebra, but it's an interesting and fun topic: recommended especially for those with interests on the applied side of mathematics.

There are many good probability textbooks with chapters on on Markov Chains, including a free online one generated by the Chance project at Dartmouth, available at `http://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/Chapter11.pdf`

C. The *permanent* of a square matrix M is the scalar obtained by expanding along any row or column as in the computation of the determinant, but using *all plus signs* rather than alternating $+$ and $-$. The permanent of a 0,1-matrix M gives a way to count the number of different "transversals" (collections of 1's, one in each row and one in each column) that M contains. In 1926, van der Waerden conjectured that

$$\mathrm{perm}(M) \geq \frac{n!}{n^n}$$

for any $n \times n$ *doubly stochastic* matrix; and that the minimum was achieved if and only if M was the $n \times n$ matrix all of whose entries are $1/n$. The corresponding result for a 0,1-matrix M with constant row and column sum s would be that

$$\mathrm{perm}(M) \geq \frac{s^n n!}{n^n}$$

This conjecture remained unresolved until 1980(!) when it was settled (affirmatively) by Egorychev. For this project, you would learn this proof (which contains a lot of good linear algebra – eigenvectors and symmetric bilinear forms!!) and give an expository account of it. The section 5.3 of *Combinatorial Theory* by Hall (on Reserve in the Science Library for our class) is a good reference for this.

II. "Ramsey Theory."

Another large and very active branch of contemporary combinatorics is the area called *Ramsey Theory.* The basic idea of many of the results in this subject is that in large enough collections of objects, some sort of *order* (= regularity, or symmetry) must appear. For instance, here are two iconic statements of Ramsey-type results:

1. Ramsey's Theorem: For any given integer c and integers $n_1, \ldots, n_c$, there is another integer R (depending on c and the n_i) such that if the edges of the complete graph on R vertices are colored using c colors, then there exist complete subgraphs of sizes $n_1, \ldots, n_c$ whose vertices all have the same color i (!) (For instance, it's not too hard to see that with $c = 2$, and $n_1 = n_2 = 3$, then $R = 6$: In any 2-coloring of the vertices of the complete graph on 6 vertices there are two monochromatic triangles, one of each color.)
2. van der Waerden's Theorem: For any given integers, n, c, there is a number V (depending on n and c) such that if the integers in $[V]$ (or any other V consecutive integers) are colored with c colors, then there is a monochromatic *arithmetic progression.*

The main goal here would simply to understand and then prove to the rest of the class how such an unlikely-sounding thing can be true(!)

III. The next group of topics rely mostly on the definitions of graphs, directed graphs, and so on. The basic material on graphs from Chapter 11 of Beeler is a prerequisite, and you should probably start out by reading that.

A. Matchings and The König-Egerváry theorem.

There is a also a graph form of Hall's Theorem from the topic group I above, which gives conditions under which a matching from V_1 to V_2 exists in a *bipartite* graph $G = (V_1, E, V_2)$. For this project, you would consider a result which can be seen as a generalization of that theorem. The König-Egerváry theorem in effect identifies the size of the biggest possible "partial" matching in such a graph, even if a full matching does not exist. There is also an algorithm (called the "Hungarian algorithm") for solving the *optimal assignment problem* based on this result. For this topic, you would present a discussion of the K-E theorem and its proof, include solutions for Exercises 1,2,3 from Chapter 9 of Bryant's text *Aspects of Combinatorics* (on reserve for the class in the Science Library), and discuss the application to the optimal assignment problem. A good source to get started on this topic is the first part of Chapter 9 of Bryant and Chapter 11 in Cameron, *Combinatorics.*

B. Directed Graphs and The Max Flow-Min Cut Theorem.

Directed graphs are frequently used to model flows through networks (for example traffic flows over road or rail systems, fluid flows through pipeline systems, and so forth). In this area, there is a well known theorem, first developed by Ford and Fulkerson, called the Max Flow-Min Cut Theorem. This theorem relates the maximum possible flow through a network to the capacities of its individual components. For this topic, you would present the precise statement of the theorem, give a proof, and then illustrate the result with one or more worked-out examples, the more realistic the better! Good sources to get started on this topic include Chapter 12 in Brualdi, *Introductory Combinatorics*, the second part of Chapter 9 of Bryant *Aspects of Combinatorics.*

C. Edge- and/or Vertex-Coloring of Graphs.

Ever wonder how the Registrar's Office manages to schedule all the different classes at Holy Cross so that no two classes share the same room assignment at the same time? Many problems like this one can be phrased in terms of *edge-colorings* of graphs – that is assignment of some color to each edge in a graph so that no two edges meeting at any vertex have the same color. Similarly one could ask – what is the smallest number of colors needed to color the the *vertices* of a graph in such a way that no vertices connected by an edge share the same color. Although these problems seem similar, they quickly go in very different directions, and I think it would be possible to do either one for an interesting project. For the edge-colorings, you would research a basic result called Vizing's Theorem, which states that any graph whose maximum vertex degree is d can edge-colored with either d or $d + 1$ colors. Surprisingly, no more than $d + 1$ colors are ever needed! Also include solutions to Exercises 3,4,5,6 of Chapter 7 in Bryant *Aspects of Combinatorics.* Then you would look at the application of edge-coloring of graphs to the *time-tabling problem*, a generalization of the classroom scheduling problem mentioned above. The vertex-coloring problem is quite a bit more subtle and famous results here include the Four-Color Map Theorem, which was only proved through extensive computer calculations in 1976. Bryant, *Aspects of Combinatorics* has chapters on both types of problems. These are also discussed in several of the other books on reserve.

D. Planar Graphs.

As you have no doubt noticed, there are some graphs which can be drawn in the plane so that edges meet only at vertices, and others, such as the complete bipartite graph $K_{3,3}$, where there is no "room" to draw the edges without having them intersect at other points. Graphs of the first type are called *planar* graphs; the others are called non-planar. One interesting question is whether it is possible to *characterize* which graphs are planar and which are not. There is a famous theorem due to Kuratowski which gives a complete answer to this question, and which would be your focus for this project. The background needed for the proof is covered in Chapter 13 of Brualdi does not give a complete proof of the Kuratowski theorem. There are relatively readable versions of the proof available online as lecture notes from courses elsewhere. If you want to work through one, come to me to discuss which ones would be appropriate as sources.

IV. The third group of topics below have more prerequisites in linear and abstract algebra. Give these a look. They indicate some interesting applications of algebraic ideas to combinatorial questions!

A. Error-Control Codes.

For this project, you would start in by learning some of the basics of the theory of linear codes over finite fields – one of the most interesting applications of abstract algebra and combinatorics to a real world problem (in my opinion at least!). Learn the basics of the Hamming distance and how it measures error-correcting capacity of a code. Then go on to linear codes, generator and parity-check matrices. Finally look at the definition of a *perfect code* and see how this reduces to a great combinatorial problem. Find the key examples of perfect codes that are known (Hamming and Golay codes) and describe them. To get started on this topic, you might begin by consulting the coding theory textbook *Fundamentals of Error-Correcting Codes* by Huffman and Pless (on reserve in the Science Library for our class), or any one of the other coding theory texts in our library. There is also a decent set of lecture notes for a minicourse on coding theory that I cowrote with a friend at Loyola Marymount in Los Angeles available at `http://mathcs.holycross.edu/~little/SACNAS2004.pdf` that you may find useful. There are a number of different ways one might go here, some more combinatorial, some more algebraic.

B. Latin Squares, Algebra, and Finite Projective Planes.

A *Latin square* is an $n \times n$ array in which the entries $\{1, 2, 3, \ldots, n\}$ appear exactly once each in each row and each column. For example,

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \\ 3 & 1 & 4 & 2 \\ 4 & 3 & 2 & 1 \end{pmatrix}$$

is a 4×4 Latin square. (Note: this is also the group table for multiplication modulo 5(!) if we leave out the class of 0.) Latin squares occur in many applications of combinatorics such as the design of statistical experiments.

There is an interesting connection between Latin squares and algebra, starting with the observation that any finite group table gives an example! For a deeper connection, pairs of orthogonal Latin squares have many connections with finite affine and projective geometries and finite fields. For this topic, you should choose some particular aspect of this story and present it, along with examples, etc. There's much more than can be done in 15 pages and a lot of current research here, so you'll need to be *selective*. If you want to try this topics, see me after you've had a chance to look over some of this so we can talk about the choices. Chapter 10 of Brualdi, Chapter 5 of Bryant *Aspects of Combinatorics*, and Chapter 13 of *Combinatorial Theory* by Hall (on Reserve in the Science Library for our class) are good references for this.

C. Möbius Inversion.

The Inclusion-Exclusion Principle that we have studied is a first instance of a general pattern concerning partial order relations on finite sets. The general version is called the *Möbius inversion formula*. This has many applications in number theory and other algebraic subjects that touch on combinatorics. For this project, you would learn about the basics and show how the general Möbius Inversion formula applies in several different settings. Section 6.6 in Brualdi contains everything you will need here, but be aware that that section builds on two other sections (4.5 and 5.7) that we have not covered in class. So you will probably want to read those too.