Concordia's CSE 2010 Summer Camp
6 – 13 June, 2010

Syllabus

We will move at a comfortable leisurely pace through the following topics in the order given here. It is not our ambition to cover all of them.

1 The stable marriage algorithm

In a society composed of a number of single women and the same number of single men, a compulsive matchmaker gets everybody married. The resulting arrangement may or may not be stable; we call it unstable if, and only if, there are a man and a woman — call them Jean-Louis and Anouk — such that Jean-Louis prefers Anouk to his wife and Anouk prefers Jean-Louis to her husband: these two have every incentive to elope and destroy the fabric of their society.

For illustration, consider five women and five men, where

Alice	prefers	Irwin to Garry to Henry to Jesse to Fidel,
Betty	prefers	Garry to Fidel to Jesse to Irwin to Henry,
Carol	prefers	Fidel to Henry to Garry to Jesse to Irwin,
Diane	prefers	Garry to Irwin to Jesse to Henry to Fidel,
Ellen	prefers	Jesse to Fidel to Irwin to Henry to Garry,
Fidel	prefers	Diane to Alice to Carol to Betty to Ellen,
Garry	prefers	Ellen to Alice to Diane to Betty to Carol,
Henry	prefers	Carol to Diane to Betty to Alice to Ellen,
Irwin	prefers	Ellen to Carol to Diane to Betty to Alice,
Jesse	prefers	Betty to Ellen to Diane to Carol to Alice.

One way to achieve stability here is to marry

Alice & Garry, Betty & Jesse, Carol & Fidel, Diane & Henry, Ellen & Irwin;

another way is to marry

Alice & Garry, Betty & Jesse, Carol & Henry, Diane & Fidel, Ellen & Irwin.

It is a fact that stability can be achieved no matter what the preference lists are: we will discuss an efficient algorithm — designed in 1962 by David Gale and Lloyd Shapley — that, given these lists, finds a stable pairing of the men and the women.

2 Modular arithmetic

For every choice of an integer n and a positive integer m, there are integers q and r such that $n = qm + r$ and $0 \leq r < m$. The r is called *the residue of n modulo m* or just simply $n \bmod m$. For a fixed m, classification of integers by their residues modulo m generalizes the classification of integers as even and odd: saying that $n \bmod 2 = 0$ is just a fancy way of saying that n is even and saying that $n \bmod 2 = 1$ is just a fancy way of saying that n is odd. Every child knows that

$$\begin{aligned} \text{even} + \text{even} &= \text{even} \\ \text{even} + \text{odd} &= \text{odd} \\ \text{odd} + \text{even} &= \text{odd} \\ \text{odd} + \text{odd} &= \text{even} \end{aligned} \qquad \begin{aligned} \text{even} \cdot \text{even} &= \text{even} \\ \text{even} \cdot \text{odd} &= \text{even} \\ \text{odd} \cdot \text{even} &= \text{even} \\ \text{odd} \cdot \text{odd} &= \text{odd} \end{aligned}$$

This generalizes to *arithmetic modulo m:* for instance, if $x \bmod 6 = 2$ and $y \bmod 6 = 5$, then $(x + y) \bmod 6 = 1$ and $xy \bmod 6 = 4$.

	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	3	4	5	0	1	2
4	4	5	0	1	2	3
5	5	0	1	2	3	4

addition modulo 6

	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3	0	3	0	3
4	0	4	2	0	4	2
5	0	5	4	3	2	1

multiplication modulo 6

We will discuss an efficient algorithm (the *extended Euclidean algorithm*) for solving the equation $cx \bmod m = 1$ when c and m are relatively prime (meaning that their greatest common divisor is 1) and an efficient algorithm for computing $a^b \bmod m$.

3 Primes and primality testing

Some twenty-three centuries ago, Euclid of Alexandria proved that there are infinitely many prime numbers. We will review the proof to refresh the students' memory and then we will discuss the following four conjectures, known as *Landau's problems*:

1. The conjecture that there are infinitely many primes of the form $n^2 + 1$.
2. *The Goldbach conjecture:* Every even integer greater than 2 is the sum of two primes.
3. *Twin prime conjecture:* There are infinitely many primes p such that $p + 2$ is prime.
4. Legendre's conjecture that for every integer n there is a prime between n^2 and $(n + 1)^2$.

We will prove *Fermat's little theorem* (if p is a prime and a is an integer, then $a^p - a$ is a multiple of p) and we will describe the *Miller-Rabin randomized primality test.*

Several programming assignments related to primes are on page 25 of *1001 problems in classical number theory* by J. M. de Koninck and Armel Mercier.

4 RSA public-key cryptosystem

Encryption is a way of rendering a communication private. The sender enciphers each message M into a ciphertext $E(M)$ before transmitting it to the receiver. The receiver (but no unauthorized person) knows the deciphering function D that recovers the original message from the received ciphertext: $D(E(M)) = M$. An eavesdropper who hears the transmitted message hears only the ciphertext, which makes no sense to him since he does not know how to decrypt it.

In a *public-key cryptosystem*, each user places his encryption procedure E in a publicly available directory, but keeps the corresponding decryption procedure D secret. To make this system workable, both E and D must be easy to compute; to make it secure, the publicly known E must give away no hints about the secretly kept D.

In 1977, Ron Rivest, Adi Shamir, and Leonard Adleman met these specifications in a most elegant way. Their *RSA algorithm* was one of the first great advances in public-key cryptography; it is the most commonly used encryption algorithm in electronic commerce protocols and it is included in the most popular web browsers. It uses the two algorithms in #2 and an algorithm that tests its integer input for being a prime.

We will describe the RSA algorithm and we will explain how it can be used not only for encryption, but also for appending to the message the sender's signature that can be authenticated.

5 The Collatz conjecture

Every positive integer s_0 defines a sequence $s_0, s_1, s_2, \ldots$ by

$$s_{n+1} = \begin{cases} 3s_n + 1 & \text{if } s_n \text{ is odd,} \\ s_n/2 & \text{otherwise.} \end{cases}$$

For instance, 7 defines the sequence

$$7,\ 22,\ 11,\ 34,\ 17,\ 52,\ 26,\ 13,\ 40,\ 20,\ 10,\ 5,\ 16,\ 8,\ 4,\ 2,\ 1,\ 4,\ 2,\ 1,\ \ldots\ .$$

The Collatz conjecture, proposed by Lothar Collatz in 1937, states that, for every s_0, the sequence eventually reaches the cycle $4, 2, 1, 4, 2, 1, \ldots$. This has been verified by computer for all s_0 up to $20 \cdot 2^{58}$ ($\approx 5.764 \cdot 10^{18}$), but the conjecture remain open and seems to be exceedingly hard: Shizuo Kakutani reported that

> *. . . For about a month everybody at Yale worked on it, with no result. A similar phenomenon happened when I mentioned it at the University of Chicago. A joke was made that this problem was part of a conspiracy to slow down mathematical research in the U.S.*

Its discussion leads to binary representation of integers, to tag systems, and to the the Halting Problem.

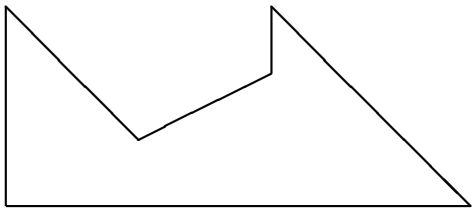
Other dynamical systems may also illustrate unexpected behaviour from simple code. For example, plot a subsequence (e.g., 500 points at 2 pixel intervals across a window) of

$$x_{n+1} = \mu x_n(1 - x_n).$$

and vary μ in [2,4]. A similar project involves the tent map.

6 The Art Gallery Theorem

The room with six walls whose floor plan is shown here cannot be guarded by a single stationary guard: no matter where in the room you stand, some part of the room is hidden from your view.



However, two stationary guards suffice to survey this room. More generally, the *Art Gallery Theorem* asserts that very polygonal room with n walls can be guarded by $\lfloor n/3 \rfloor$ stationary guards. (Here, as usual, $\lfloor x \rfloor$ denotes x rounded down to the nearest integer.)

We will explain the pretty proof of this theorem. To make it self-contained, one has to prove that every polygon can be triangulated. A relatively simple proof of this auxiliary theorem is reproduced, for instance, in Subhash Suri's lecture notes (page 13).

An attractive programming challenge is to design an algorithm that, given the sequence of corners of the n-gon in a their cyclic order (such as $[4,3]$, $[7,2]$, $[0,0]$, $[0,3]$, $[2,1]$, $[4,2]$, $[4,3]$ in our example), finds the set of at most $\lfloor n/3 \rfloor$ guards.

Additional web sources for the Art Gallery Theorem include AMS Feature Column by Joseph Malkevitch and an interactive applet by Alexander Bogomolny (The theorem is also alluded to in *Obsession,* Episode 16, Season 2 of the TV series "Numb3rs"; we plan to screen this episode for the students.)

7 Ramsey's theorem

In every group of six people, there are always three people such that every two of them are friends or every two of them are strangers. This is easy to check: Take any person and call her A. If A has at least three friends among the remaining five people, then call them B, C, D and consider two subcases: if two of B, C, D are also friends of each other, then these two people and A are the group of three we are looking for; if no two of B, C, D are friends of each other, then B, C, D are the group of three we are looking for. If A has at most

two friends among the five people other than A, then some three of these five people are strangers to A; call them B, C, D and repeat the argument with the roles of friends and strangers interchanged.

More generally, Frank Plumpton Ramsey proved in 1928 that for every positive integer k there is a positive integer n such that

> *in every group of n people, there are always k people*
> *such that every two of them are friends or every two of them are strangers.*

When $R(k)$ denotes the smallest n with this property, we have $R(3) = 6$: $R(3) \leq 6$ is shown by the case analysis given above and $R(3) > 5$ is shown by five people seated at a round table so that everybody is a friend of the two people sitting next to them and a stranger to the other two.

We will survey the known bounds on $R(k)$ for general k and go into some of the proofs.

8 The Friendship Theorem

An excellent introduction to graph-theoretic concepts is provided by a proof of the following theorem:

> *If, in a finite group of people, every two people have precisely one common friend, then there is a politician (meaning everybody's friend) in this group.*

9 Hamiltonian cycles

In 1857, Sir William Rowan Hamilton invented a puzzle, which he called *The Icosian Game*. In one version of this puzzle, called "The Travellers Dodecahedron or A Voyage Around the World", each of the twenty vertices of the dodecahedron (the polyhedron with 12 pentagonal faces, 20 vertices, and 30 edges) is labeled by the name of an important city (the list goes from Brussels to Zanzibar) and the objective is to travel through the edges of the dodecahedron in such a way that each of the twenty cities gets visited precisely once.

The solution, shown here, is easy to find. When the graph of the dodecahedron is replaced by other graphs, the problem of finding the *Hamiltonian cycle* — a cycle passing through the edges in such a way that each vertex gets visited precisely once — may get more difficult: try the Petersen graph or this graph or this graph.

We will dicuss two efficient algorithms that, given a graph G (represented by a list of vertices and a list of edges), attempt to find a Hamiltonian cycle in G.

The *Bondy-Chvátal algorithm* fails to find a Hamiltonian cycle only if G has relatively few edges compared to the number n of its vertices or if its edges are distributed in a relatively lopsided way. More precisely, the *degree* of a vertex v means the number of vertices that can be reached from v along a single edge. The algorithm fails to find a Hamiltonian cycle only if, for some k in the range $0 < k < n/2$, there are at least k vertices of degree at most k and at least $n - k$ vertices of degree at most $n - k$.

The *Chvátal-Erdős algorithm* fails to find a Hamiltonian cycle only if G has a structural fault of a different kind: a set S of vertices no two of which are joined by a single edge along with a set of fewer than $|S|$ vertices, whose removal breaks the graph into two or more pieces.

Design of efficient computer implementations of these two algorithms presents a number of interesting challenges.

10 Binomial coefficients

Combinatorial proofs of binomial identities: some subset of identities 125–137 and 151–162 in Chapter 5 of *Proofs that really count: the art of combinatorial proof* by Arthur T. Benjamin and Jennifer J. Quinn. (On a far more advanced level, there is the book $A{=}B$ by Marko Petkovšek, Herbert Wilf and Doron Zeilberger, which can be downloaded free from the web.) Bounds on binomial coefficients such as

$$\binom{2n+1}{n+1} \le 4^n, \quad \text{and} \quad \binom{2n}{n} \ge \frac{4^n}{2n} \text{ whenever } n \ge 1,$$

which are a prerequisite for #11

11 Erdős's proof of Bertrand's postulate

In 1845, Joseph Bertrand conjectured a far stronger proposition:

> *There is a prime between every positive integer and its double.*

This was proved in 1859 by Pafnuty Chebyshev; his proof was later streamlined first by Edmund Landau and then by Srinivasa Ramanujan. In 1931, Paul Erdős — eighteen years old at that time — found an elegant new proof. We will present its details.

12 Van der Waerden's theorem

No matter how $1, 2, 3, 4, 5, 6, 7, 8, 9$ are coloured in two colours, there is always a monochromatic arithmetic progression of length three, meaning $a, a + d, a + 2d$ with $d > 0$. This is easy to check: First, switching colours (red↔white) if necessary, we may assume that 5 is red. Next, to avoid red arithmetic progression $1, 5, 9$, we must make at least one of 1 and 9 white; flipping the board ($1 \leftrightarrow 9$, $2 \leftrightarrow 8$, $3 \leftrightarrow 7$, $4 \leftrightarrow 6$) if necessary, we may assume that 1 is white. Finally, 7 is either red or white; in either of these two cases, the integers coloured so far trigger off a chain reaction that ends up, one way or another, in a monochromatic arithmetic progression of length three.

More generally, Bartel Leendert van der Waerden proved in 1926 that for every positive integer k there is a positive integer n such that

> *no matter how* $1, 2, \ldots, n$ *are coloured in two colours, there is always a monochromatic arithmetic progression of length* k*, meaning* $a, a+d, a+2d, \ldots, a+(k-1)d$ *with* $d > 0$.

When $W(k)$ denotes the smallest n with this property, we have $W(3) = 9$: the case analysis sketched above shows that $W(3) \leq 9$ and the colouring

1	2	3	4	5	6	7	8
red	red	white	white	red	red	white	white

shows that $W(3) > 8$. Finding $W(4)$ is an interesting programming assignment.

We will survey the known bounds on $W(k)$ for general k and, depending on the students' preferences, we may go into some of the (relatively difficult) proofs.

13 Towers of Hanoi

This puzzle provides a simple example of a problem for which a recursive solution is both shorter and more intuitive (if explained carefully!) than the iterative solution (which requires insight to find). In addition, its discussion leads to binary representation of integers and to Gray codes.

14 Points and lines

James Joseph Sylvester asked in 1893 for a proof of the following proposition:

> *if finitely many points in the plane do not lie on a line, then some line passes through precisely two of them.*

Forty years later, Tibor Gallai proved it. We will explain the short and simple proof, often cited as a paragon of beauty, found by by Leroy M. Kelly in 1943.

A corollary of the Sylvester-Gallai theorem, esily proved by induction on n, atates that

> *If n points in the plane do not lie on a line, then at least n distinct lines pass through pairs of them.*

The "at least n" here cannot be strengthened to "at least $n+1$": when $n-1$ points lie on a line and an additional point lies off this line, then precisely n distinct lines pass through pairs of these n points.

15 The de Bruijn-Erdős theorem and projective planes

The corollary in #14 generalizes as DE BRUIJN-ERDŐS THEOREM:

> *If E is a family of proper subsets of an n-point set V such that every two distinct points of V belong to precisely one member of E, then $|E| \geq n$.*

The extremal families E here (those consisting of precisely n sets) come in two flavours, typified respectively by

$$\{1,2,3,4,5,6\},\{1,7\},\{2,7\},\{3,7\},\{4,7\},\{5,7\},\{6,7\}$$

and by

$$\{1,2,4\},\{2,3,5\},\{3,4,6\},\{4,5,7\},\{5,6,1\},\{6,7,2\},\{7,1,3\}.$$

Families of the first kind, so-called pencils, consist of one set of size $n-1$ and $n-1$ sets of size 2. In families of the second kind, n equals k^2-k+1 for some positive integer k, every member of E consists of precisely k points of V, and every point of V belongs to precisely k members of E (in addition to every two distinct points of V belonging to precisely one member of E); these families are called *projective planes of order $k-1$*.

We will prove the de Bruijn-Erdős theorem (including its characterization of the extremal families) and we will discuss, without proofs, the question for which values of m there are projective planes of order m. If m is a prime or a power of a prime ($m = 2,3,4,5,7,8,9,11,\ldots$), then there exists a projective plane of order m; it has been known since 1938 that there exists no projective plane of order 6; one of the most celebrated problems in combinatorics was the question whether there is a projective plane of order 10.

In 1988, this question was answered in the negative by a team of researchers in our department led by Clement Lam. (This achievement was reported in the New York Times and Dr. Lam's own account of the adventure is here).