More about Matroids

Guest lecture in COMP150-Graph Theory
Anselm Blumer
21 November, 2019

Agenda

- intersections and unions of matroids
- a matroid that is vectorial but not graphical
- a digression on error-correcting codes
- spans, closed sets, and hyperplanes
- a matroid that is neither vectorial nor graphical
- planar graphs and duals of matroids
- a finite number of slides about infinite graphs
- some random slides about random graphs and other random objects
- infinite matroids

More about intersections of two matroids

- From last time: intersections of hereditary systems are hereditary systems, but the intersection of two matroids may not be a matroid
- This means that a problem defined in terms of the intersection of two matroids may not have a solution using a greedy algorithm, though Papadimitriou and Steiglitz give an algorithm for solving any 2-matroid intersection problem in time polynomial in the matroid oracles
- Details about Edmonds' Matroid Intersection Algorithm can also be found in Paul Wilhelm's lecture notes at

  https://www.mathematik.hu-berlin.de/~wilhelm/greedy-ausarbeitung.pdf

More about intersections of three matroids

- Problems defined in terms of the intersection of three matroids may be NP-complete

- For example, the problem of finding a Hamiltonian path in a directed graph can be defined in terms of a graphic matroid, a head partition matroid, and a tail partition matroid (Theorem 12.10 in Papdimitriou and Stieglitz)

Hamiltonian path as a 3-matroid intersection problem

- A Hamiltonian path in a directed graph has no two edges whose tails meet (independent in the tail partition matroid), no two edges whose heads meet (independent in the head partition matroid), and doesn't contain a cycle (independent in the graphical matroid for the corresponding undirected graph)

- Finding a set of edges that is simultaneously maximal in these three matroids solves the Hamiltonian path problem:
  The graph has a Hamiltonian path ⇔ the maximal set has |V|-1 edges

- Hamiltonian path is known to be NP-complete, so finding a polynomial-time algorithm for 3-matroid intersection is worth $1,000,000 (base 10) from the Clay Mathematics Institute (and probably a Turing award)

The Matroid Union Theorem

- (West 8.2.55) If M_1, M_2, … M_k are matroids with independent sets defined as subsets of I_1, I_2, … I_k and rank functions r_1, r_2, …r_k , then $M_1 \cup M_2 \cup \ldots \cup M_k$ is a matroid with independent sets being unions of sets that are independent in the component matroids and rank function
  $r(X) = \min\{ \sum r_i(Y) + |X - Y| \}$ where the min is over all $Y \subseteq X$

- Proved independently by Edmonds and Fulkerson (1965) and Nash-Williams (1966)

- West's Graph Theory textbook gives several applications of this theorem to covering and packing problems

Agenda

- intersections and unions of matroids
- a matroid that is vectorial but not graphical
-
-
-
-
-
-
-
-

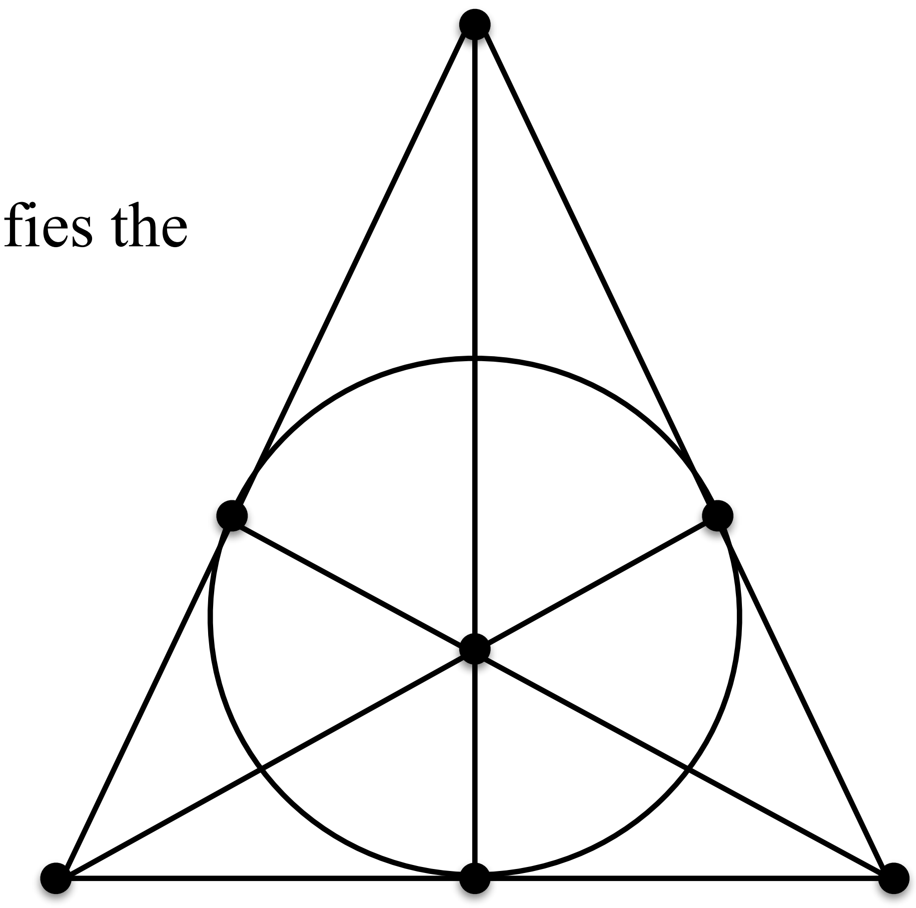
The Fano plane

- The Fano plane is a *projective plane* since it satisfies the following three properties:
(in fact it is the smallest projective plane)

- Any pair of points are on exactly one line
(the circle is really a line)

- Any pair of lines intersect in exactly one point

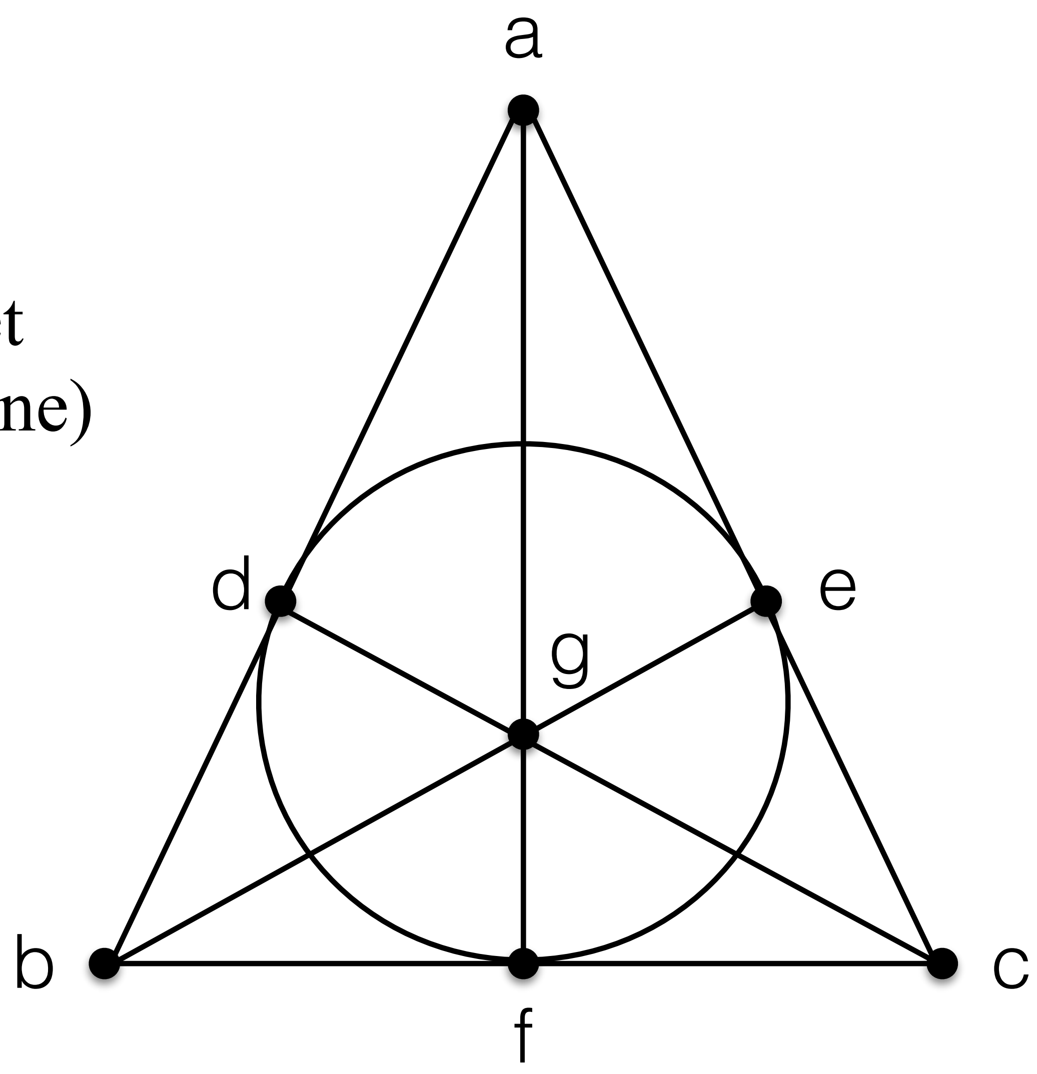
- There are four points, no three of which are collinear
(for example, the center point plus the three vertices of the large triangle)



The Fano plane as a matroid

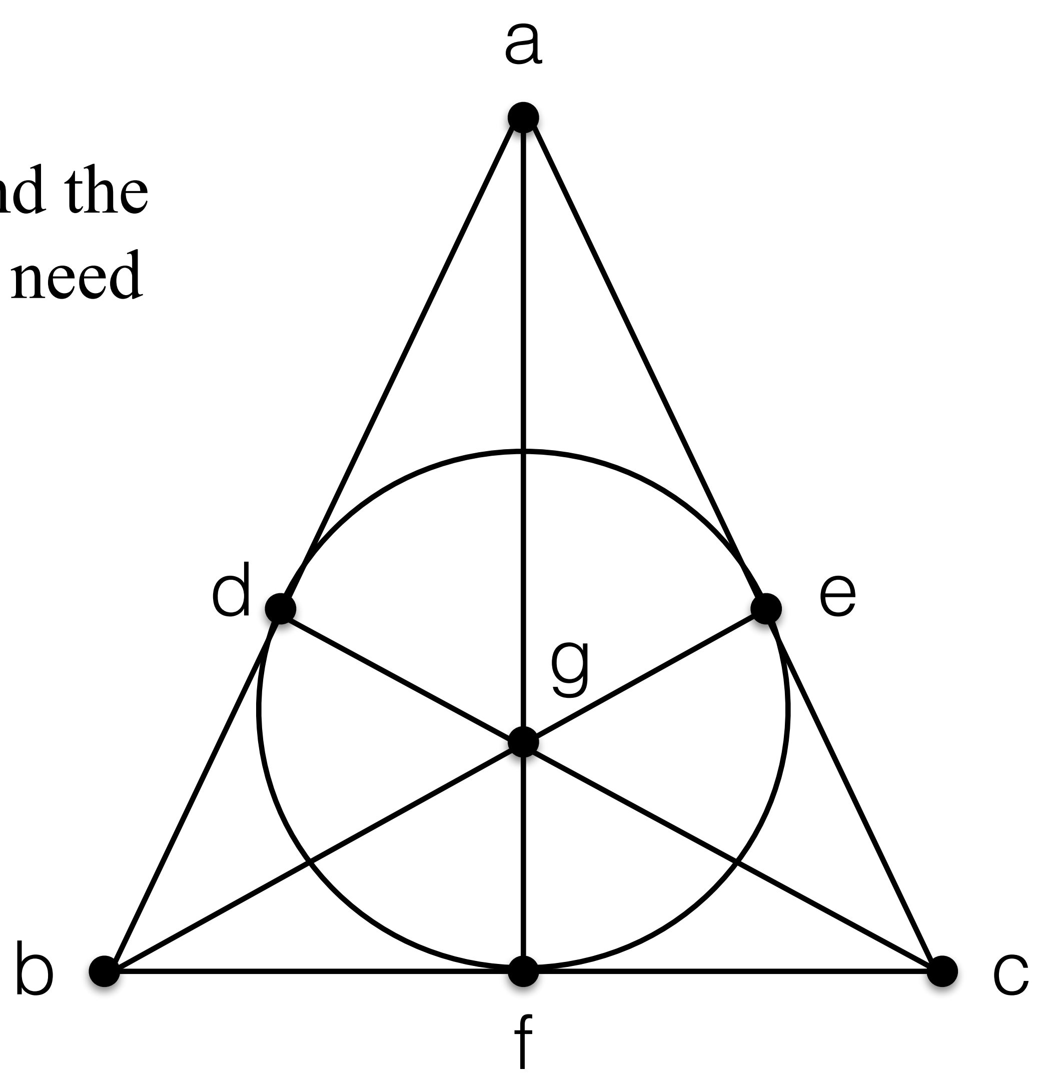
- The Fano plane is a matroid with ground set E = {a, b, c, d, e, f, g} (the points of the plane) and 3-element circuits defined as the lines (straight or circular) shown in this figure

- The Fano matroid is not graphical

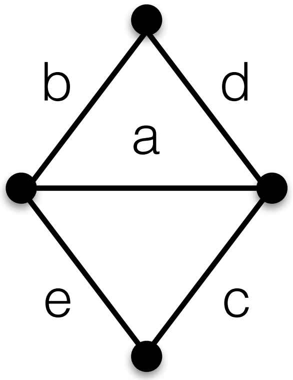


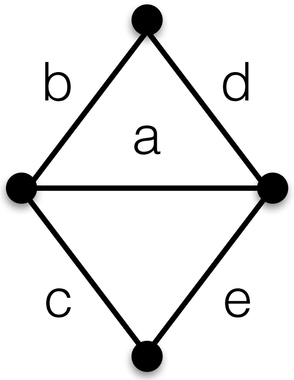
The Fano matroid is not graphical

- The outside lines {a,d,b}, {a,e,c}, {b,f,c}, and the vertical line {a,g,f} are circuits so they would need to be cycles in the graphical representation



- The first two share an edge, so there are two possible graphs:



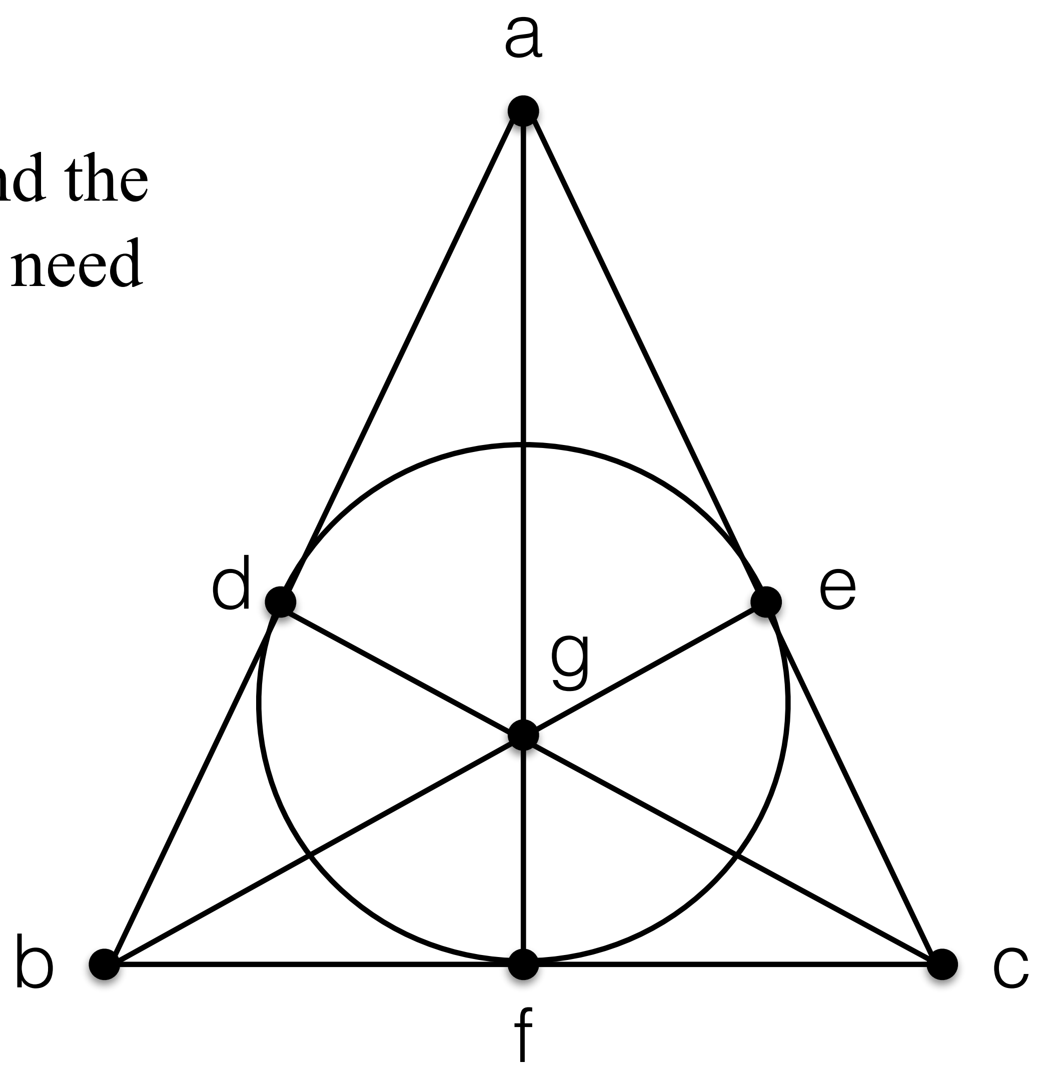


- In the left graph, it isn't possible to add edge to form the third cycle

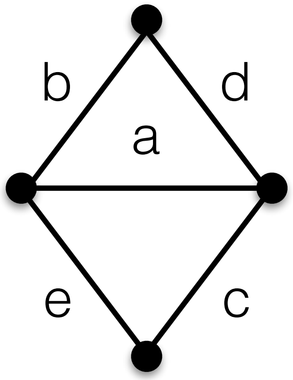
-

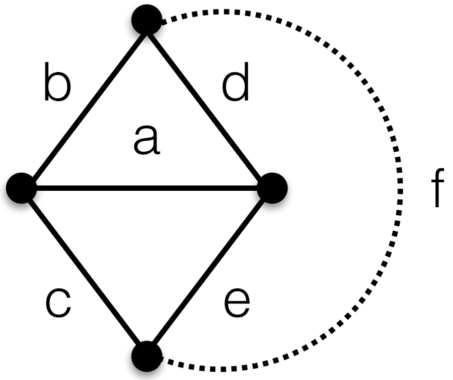
The Fano matroid is not graphical

- The outside lines {a,d,b}, {a,e,c}, {b,f,c}, and the vertical line {a,g,f} are circuits so they would need to be cycles in the graphical representation



- The first two share an edge, so there are two possible graphs:



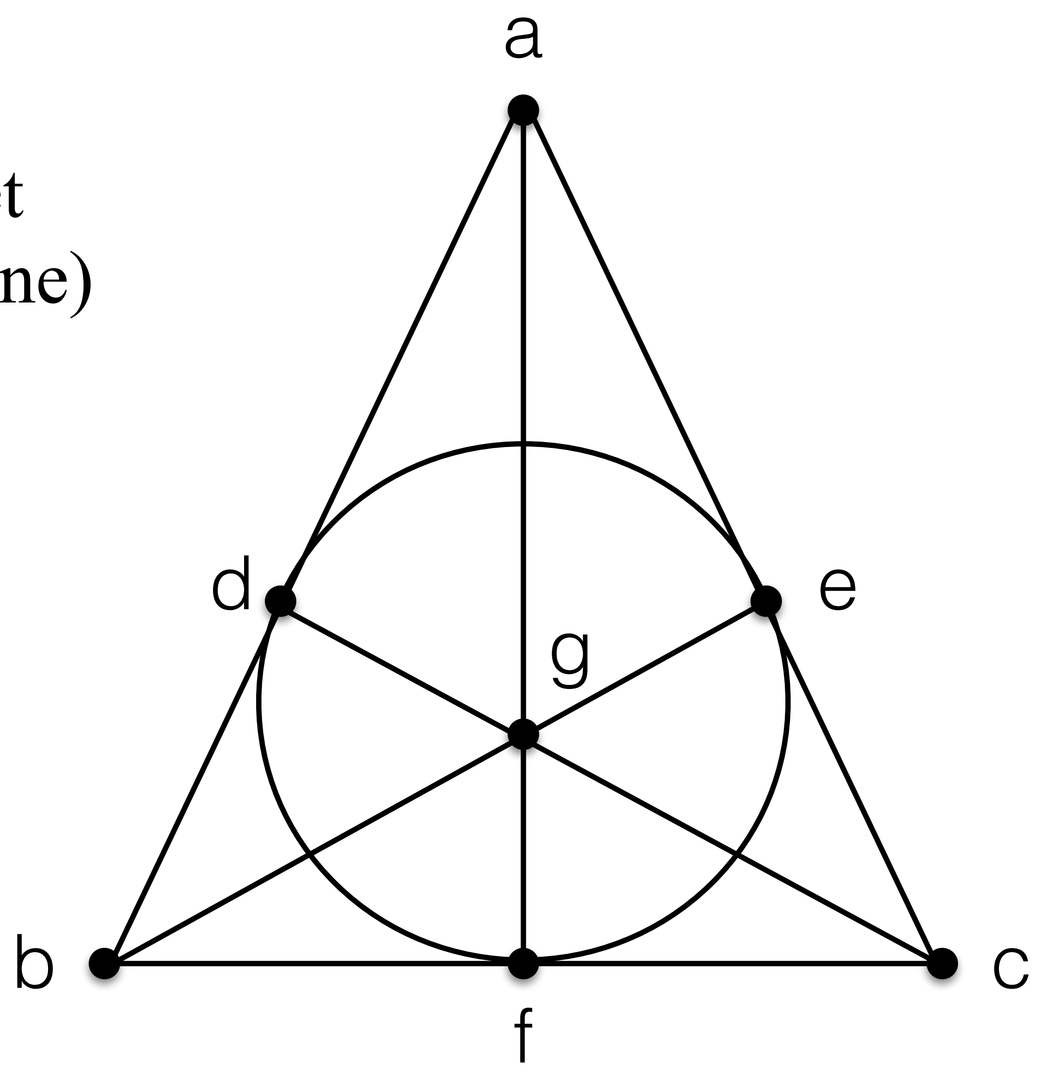


- In the left graph, it isn't possible to add edge to form the third cycle

- In the right graph, it is possible to add f so it forms a cycle {b,f,c} but then it's impossible to form a cycle for {a,g,f}

The Fano plane as a matroid

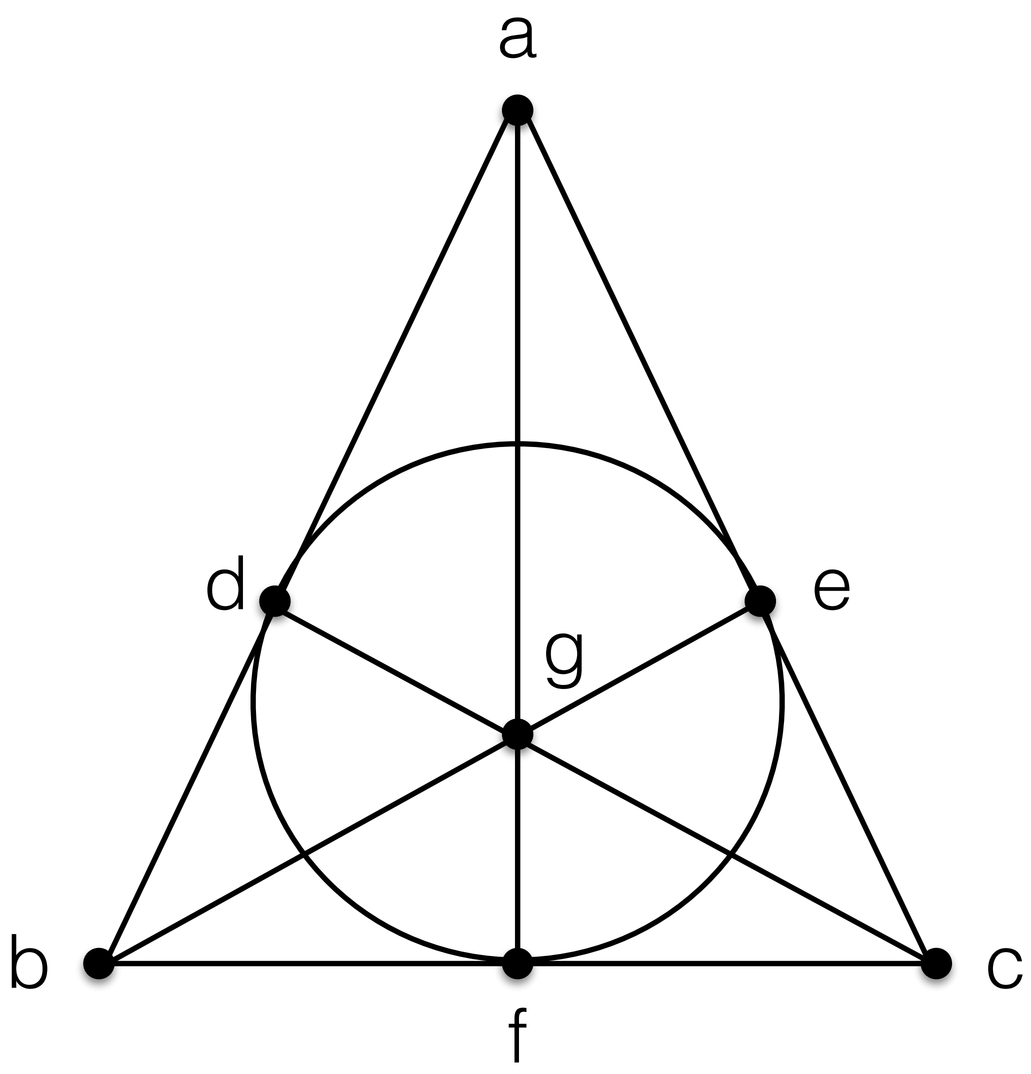
- The Fano plane is a matroid with ground set E = {a, b, c, d, e, f, g} (the points of the plane) and 3-element circuits defined as the lines (straight or circular) shown in this figure
- The Fano matroid is not graphical
- The Fano matroid is vectorial since the points can be labelled with binary vectors satisfying the appropriate independence properties



The Fano matroid *is* vectorial

- Columns a…g of this matrix, viewed as three-element binary vectors, represent the Fano plane as a vectorial matroid

a	b	c	d	e	f	g
1	0	0	1	1	0	1
0	1	0	1	0	1	1
0	0	1	0	1	1	1



Agenda

- intersections and unions of matroids
- a matroid that is vectorial but not graphical
- a digression on error-correcting codes

Communication and error detection

- Suppose we want to transmit the characters ‘acb’ using binary, but the communication channel is unreliable

- We can encode ‘acb’ into binary using ASCII:
  acb ➔ 1100001 1100011 1100010 ➔ acb
  (note that the actual order of transmission is 100001111000110100011)

- If a bit error occurs on the fourth bit, this will be received as ‘icb’

- Inserting ‘parity check bits’ allows the receiver to detect that there was an error:

  No error: acb ➔ $\underline{1}1100001\underline{0}1100011\underline{1}1100010$ ➔ acb
  (the underlined bits give each byte even parity)

  A single bit error on any byte can be detected by the receiver:
  acb ➔ $\underline{1}1100001\underline{0}1100011\underline{1}1100010$ ➔ $\underline{1}1101001\underline{0}1100011\underline{1}1100010$ ➔ $\underline{?}$cb

Communication, error detection, and error correction

- If the receiver detects an error it needs to send a request for retransmission, but two-way communication isn't always available

- An alternative is to add more parity check bits to allow the receiver to correct errors without communicating back to the transmitter

- A simple example:
  Send two 'parity check bits' with each bit 0 ➔ 000 ➔ 0 and 1 ➔ 111 ➔ 1

- Now a bit error can be corrected: 0 ➔ 000 ➔ 100 ➔ 000 ➔ 0

- (But double errors won't be corrected: 0 ➔ 000 ➔ 101 ➔ 111 ➔ 1)

- Communication is much more reliable, at the expense of reducing the communication rate by a factor of 3

What does this have to do with matroids?

- It turns out that the matrix representation of the Fano matroid is the parity check matrix of a code that will correct one error among each group of seven bits, at the expense of slowing communication by a factor of 1.75

- If (a, b, c, d) are four bits to be transmitted,
  add three parity-check bits calculated by:
  $e = b \oplus c \oplus d$, $f = a \oplus c \oplus d$, $g = a \oplus b \oplus c$ ($\oplus$ is XOR)

- Two examples:
  0101 ➔ 0101011
  1110 ➔ 1110001

- Suppose the fourth bit is corrupted:
  0101 ➔ 0101011 ➔ 0100011
  1110 ➔ 1110001 ➔ 1111001

a	b	c	d	e	f	g
1	0	0	1	1	0	1
0	1	0	1	0	1	1
0	0	1	0	1	1	1

- To correct the error, multiply the parity check matrix by the received bit vector

What does this have to do with matroids?

- To correct the error, multiply the parity check matrix by the received bit vector (in other words, XOR the columns corresponding to 1-bits)

- For 0101 ➔ 0101011 ➔ 0100011
  this gives $(1, 1, 0)^T$

- For 1110 ➔ 1110001 ➔ 1111001
  this gives $(1, 1, 0)^T$

- It's not a coincidence that in both cases this gives a column identical to the fourth column of the matrix, signifying that the fourth bit needs to be flipped back

a	b	c	d	e	f	g
1	0	0	1	1	0	1
0	1	0	1	0	1	1
0	0	1	0	1	1	1

- This coding scheme is called the (7,4) Hamming code

The (7, 4) Hamming code

- If (a, b, c, d) are four bits to be transmitted, add three parity-check bits calculated by $e = b \oplus c \oplus d$, $f = a \oplus c \oplus d$, $g = a \oplus b \oplus c$

- To detect and correct an error, multiply the parity check matrix by the received bit vector (in other words, XOR together the columns corresponding to 1-bits)

- Since the columns of the parity check matrix consist of all seven nonzero three-bit vectors, every possible received word (a,b,c,d,e,f,g) results in either a column of all zeroes or one of the columns of the matrix

- In the first case (a,b,c,d,e,f,g) was a codeword and there was no error (or possibly 3 or more errors)

- In the second case the location of the column indicates that correcting a single bit would change (a,b,c,d,e,f,g) to a codeword

a	b	c	d	e	f	g
1	0	0	1	1	0	1
0	1	0	1	0	1	1
0	0	1	0	1	1	1

- In other words, every possible seven-bit vector is at Hamming distance 0 or 1 from a codeword (the Hamming distance between two bit vectors is the minimum number of bits that need to be flipped to convert one vector into the other)

Agenda

- intersections and unions of matroids
- a matroid that is vectorial but not graphical
- a digression on error-correcting codes
- spans, closed sets, and hyperplanes
-
-
-
-
-

Span functions for hereditary systems

- Given a hereditary system M on ground set E, the span function σ_M is a function from subsets of E to subsets of E given by $\sigma_M(X) = X \cup \{e \in E \mid Y + e \text{ forms a circuit for some } Y \subseteq X\}$

- Span functions have the following properties (8.2.25)
  (s1) expansive: $X \subseteq \sigma(X)$
  (s2) order-preserving: $Y \subseteq X$ implies $\sigma(Y) \subseteq \sigma(X)$
  (s3) Steinitz exchange: If e is not in $\sigma(X)$ but e is in $\sigma(X+f)$ then f is in $\sigma(X+e)$

- (8.2.26) If adding e to a set X doesn't increase its rank then e was in the span of X

Matroids can be defined in terms of span functions

- (incorporation) $r(\sigma(X)) = r(X)$ for all $X \subseteq E$
- (idempotence) $\sigma(\sigma(X)) = \sigma(X)$ for all $X \subseteq E$
- (transitivity of dependence) If *e* is in the span of X and X is a subset of the span of Y then *e* is in the span of Y
  [$e \in \sigma(X)$ and $X \subseteq \sigma(Y) \Rightarrow e \in \sigma(Y)$]
- (West 8.2.27)

Terminology related to span functions

- *Spanning sets*: $X \subseteq E$ with $\sigma(X) = E$
- *Closed sets* (or *flats*, or *subspaces*): $X \subseteq E$ with $\sigma(X) = X$
- *Hyperplanes*: maximal proper closed subsets of E
- Span functions are sometimes called *closure* functions

Agenda

- intersections and unions of matroids
- a matroid that is vectorial but not graphical
- a digression on error-correcting codes
- spans, closed sets, and hyperplanes
- a matroid that is neither vectorial nor graphical
-
-
-
-

The Vámos matroid

- Described by Peter Vámos in 1968 in an unpublished manuscript

- The ground set consists of the eight red dots in the figure

- Every set of four elements is a base except for the five planes colored blue

- The Vámos matroid is neither graphical nor vectorial (see James G. Oxley, *Matroid Theory*, 2006)

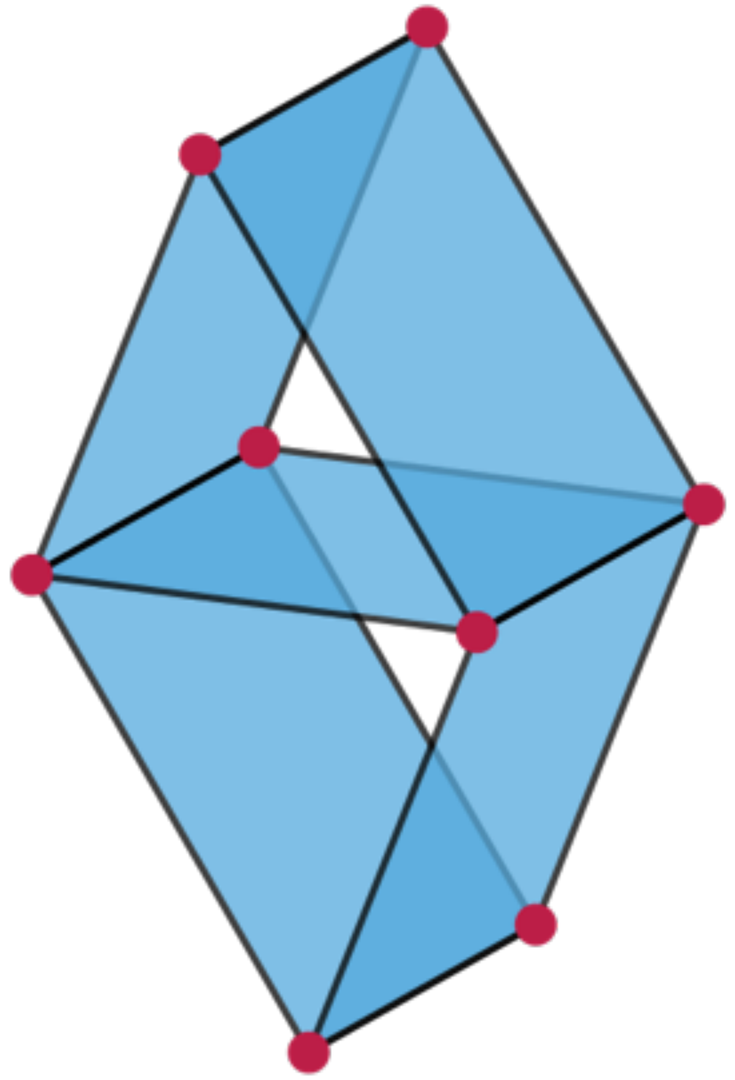


Figure by David Eppstein, Creative Commons CC0 1.0 Universal Public Domain Dedication

Agenda

- intersections and unions of matroids
- a matroid that is vectorial but not graphical
- a digression on error-correcting codes
- spans, closed sets, and hyperplanes
- a matroid that is neither vectorial nor graphical
- planar graphs and duals of matroids
-
-
-

Duals of matroids

- A dot above the name of a set will denote the complement with respect to the ground set E: $\dot{X} = E - X$

- The *dual* of a matroid can be defined in terms of complements of bases: If M is a matroid on ground set E with bases $\{B_i\}$, then M* is a matroid on E with bases $\{\dot{B}_i\}$

- The fact that this defines a matroid was proved by Whitney in 1935 using the base exchange property
  (for every e in $B_1 - B_2$ there is an f in $B_2 - B_1$
  so that $B_1 - \{e\} + \{f\}$ is a base)

- (West 8.2.34) The rank function of the dual of a matroid satisfies
  $r^*(X) = |X| - (r(E) - r(\dot{X}))$

Duals of connected planar graphs

- A connected planar graph G = (V, E) has a natural dual G* = (V*, E*) formed by associating each face of G (including the unbounded face) with a vertex in V* and each edge in E with with an edge e* connecting the faces on opposite sides of e.

- A set of edges X ⊆ E is a spanning tree for G if and only if { e* ∈ E* | e ∈ Ẋ } is a spanning tree for G*
(This is Exercise 6.1.21 in West)

Graph duals and matroid duals

- The *cycle matroid* of a graph, G, is the matroid with ground set E(G) and circuits (minimal dependent sets) given by the cycles of G

- The bases of a cycle matroid M(G) of a connected graph G are the spanning trees of G

- Thus the bases of M(G*) are the complements of the bases of M(G)

- The greedy algorithm finds minimum-weight bases in M(G) and simultaneously finds maximum-weight bases in M(G*)

Bonds in graphs and matroids

- An *edge cut* of a graph G, denoted [S, V(G)-S], is the set of edges having one endpoint in S and the other endpoint in V(G)-S
  (S should neither be empty nor all of V(G))
- A *bond* is a minimal edge cut
- The *bond matroid* of G is the matroid whose circuits are the bonds of G
- The bond matroid is the dual matroid of the cycle matroid
- G is planar if and only if its bond matroid is graphic
  (West 8.2.44, proved by Whitney in 1933)

Agenda

- intersections and unions of matroids
- a matroid that is vectorial but not graphical
- a digression on error-correcting codes
- spans, closed sets, and hyperplanes
- a matroid that is neither vectorial nor graphical
- planar graphs and duals of matroids
- a finite number of slides about infinite graphs
-
-

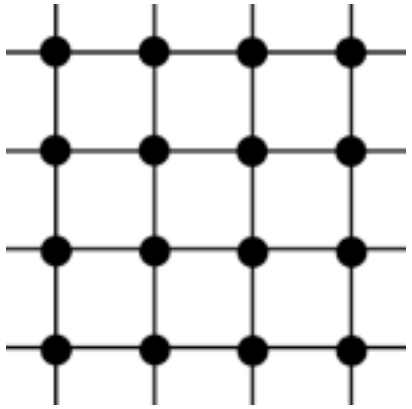
Infinite graphs

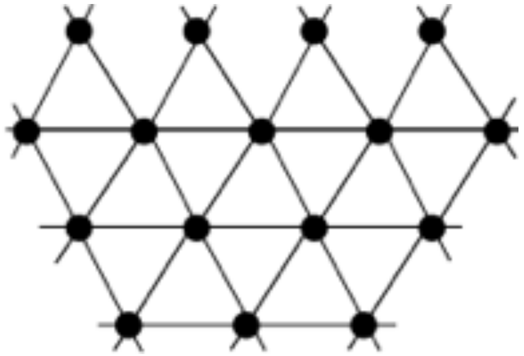
- What is an infinite graph?
  According to Wolfram MathWorld:
  A graph that is not finite is called infinite.
- If every vertex has finite degree, the graph is *locally finite*
- If the set of vertices is countably infinite and there is at most one edge (or two directed edges) between any pair of vertices then the set of edges is countable (since the union of a countable number of countable sets is countable)
- It’s also possible to define graphs on sets of vertices with higher cardinality

Examples of infinite graphs

- An infinite sequence of distinct vertices $v_1, v_2, v_3, \ldots$ together with edges $(v_1, v_2), (v_2, v_3), (v_3, v_4), \ldots$ is called a *ray*

- The square grid graph and the triangular grid graph

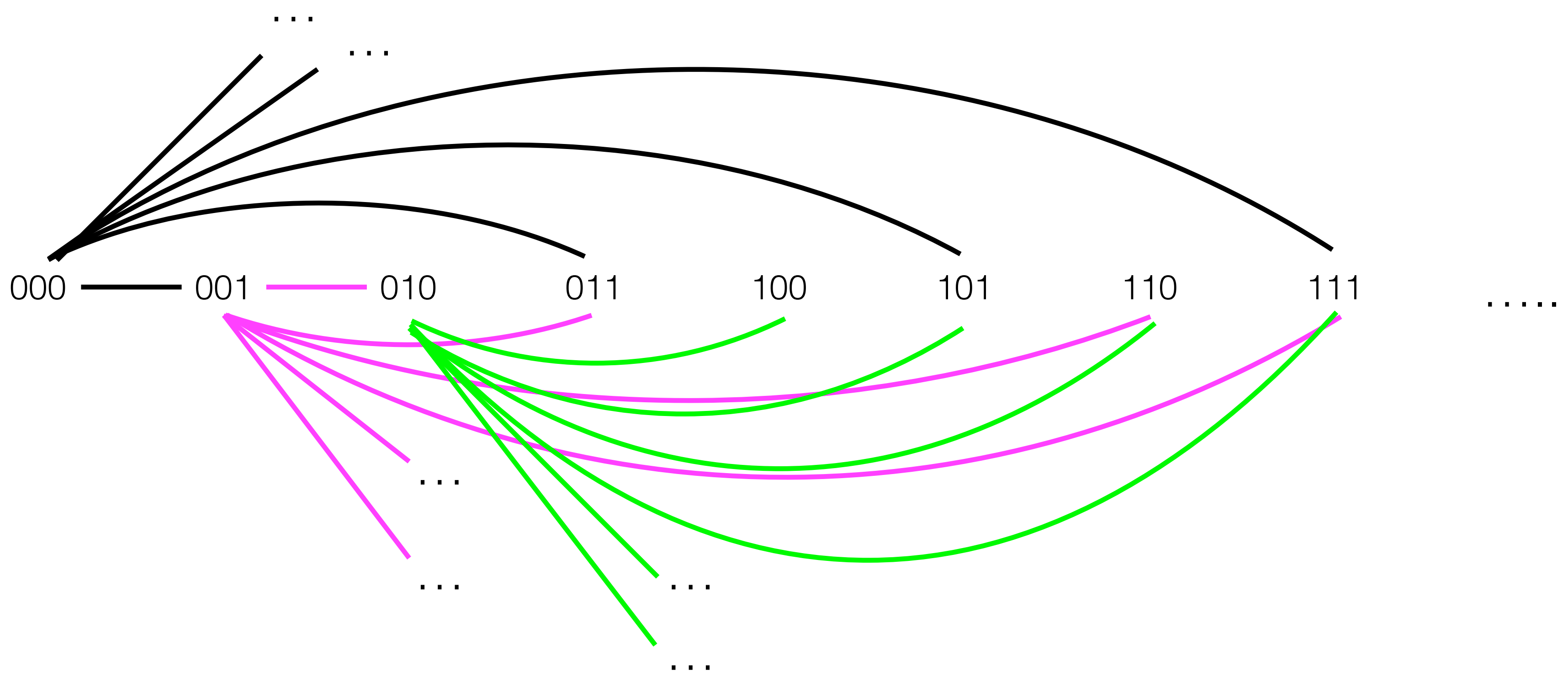




- One could define variants of these that are not locally finite, for example each vertex in the square grid graph could have an edge to every other vertex in the same row or column

The Rado graph

- First constructed by Ackermann in 1937
- Rediscovered by Erdős and Rényi in 1963 and Richard Rado in 1964



- Vertices are numbered 0,1,2,… and vertex *j* is connected to all vertices with bit *j* set to 1 (bit 0 is the low-order bit)

Properties of the Rado graph

- It is not locally finite, in fact every vertex has infinite degree
- Every finite graph and every countably infinite graph is isomorphic to an induced subgraph of the Rado graph
- If one constructs a graph at random starting with a countable set of vertices, and connects each pair of vertices independently with probability 0.5, the resulting graph is (with probability 1) isomorphic to the Rado graph (so the Rado graph is self-complementary)
- The above statement is still true if 0.5 is replaced by any fixed probability $0 < p < 1$
- The automorphism group of the Rado graph is a simple group with size equal to the cardinality of the continuum

Agenda

- intersections and unions of matroids
- a matroid that is vectorial but not graphical
- a digression on error-correcting codes
- spans, closed sets, and hyperplanes
- a matroid that is neither vectorial nor graphical
- planar graphs and duals of matroids
- a finite number of slides about infinite graphs
- some random slides about random graphs and other random objects
-

Random graphs

- What should we get if we ask for a random graph?

Random graphs

- What should we get if we ask for a random graph?
- There are two popular and closely related models of random simple graphs with *n* vertices:

  - the Erdős-Rényi model (1959), G(*n*, *m*) or $G_{n,m}$ or sometimes $G_{n,N}$ where *m* (or *N*) is the number of edges

  - the Gilbert model (1959), G(*n*, *p*) or $G_{n,p}$

Random graphs

- What should we get if we ask for a random graph?

- There are two popular and closely related models of random simple graphs with *n* vertices:

  - the Erdős-Rényi model (1959), G(*n*, *m*) or $G_{n,m}$ or sometimes $G_{n,N}$ where *m* (or *N*) is the number of edges

  one can generate a graph in the $G_{n,m}$ model by starting with a graph with no edges and selecting one of the C(*n*,2) possible edges at random, then selecting one of the remaining C(*n*,2)-1 edges at random, … until the graph has *m* edges

Random graphs

- What should we get if we ask for a random graph?
- There are two popular and closely related models of random simple graphs with n vertices:

  - the Erdős-Rényi model (1959), G(n, m) or $G_{n,m}$ or sometimes $G_{n,N}$ where m (or N) is the number of edges

  - the Gilbert model (1959), G(n, p) or $G_{n,p}$

  one can generate a graph in the $G_{n,p}$ model by flipping a biased coin with P(heads) = p for each of the C(n,2) edges and adding that edge if the coin is heads

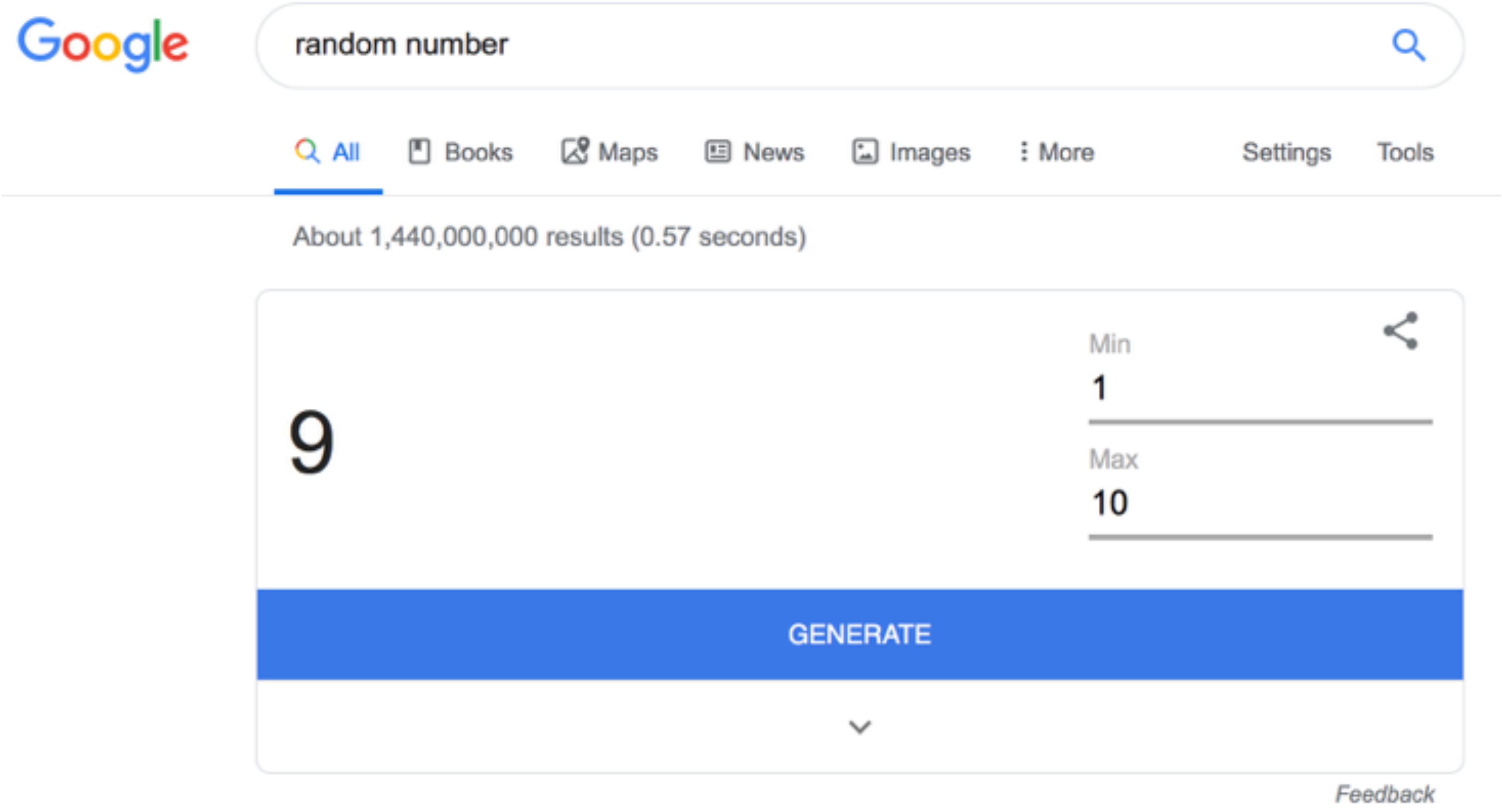
Random numbers

- These models reduce the problem of generating random graphs to the problem of generating random numbers or random bits

- So what is a random number?

Random numbers

- We could ask Google to give us a random number:



random number

About 1,440,000,000 results (0.57 seconds)

9

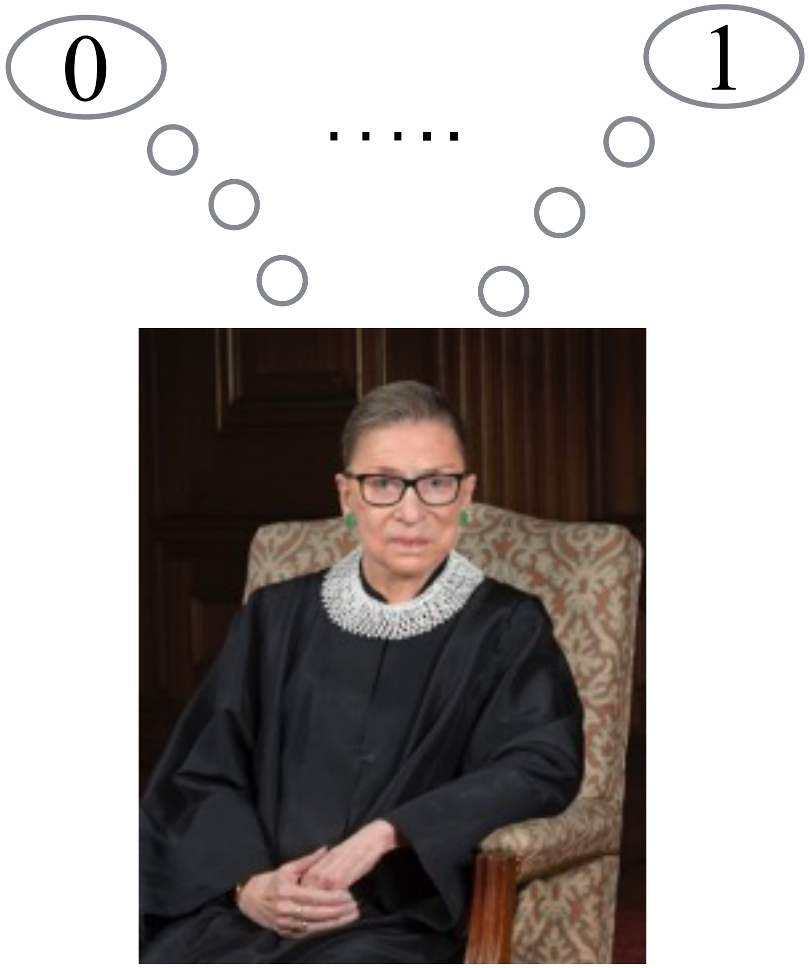
Min
1

Max
10

- Apparently 9 is a random number (and possibly 1,440,000,000 as well)

Random bits

- We could ask a Random Bit Generator (RBG)



Random Bit Generators according to NIST
(the National Institute of Standards and Technology)

- Go to https://csrc.nist.gov/glossary/term/Random-Bit-Generator

- Random Bit Generator (RBG)

  Abbreviation(s) and Synonym(s):

  RBG

  Definition(s):
  A device or algorithm that outputs a sequence of binary bits that appears to be statistically independent and unbiased. An RBG is either a DRBG or an NRBG.

  Source(s):
  NIST SP 800-90A Rev. 1

- This is a 110-page (base ten) document) available at
  https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-90Ar1.pdf

- caveat on page 2:
  The precise structure, design and development of a random bit generator is outside the scope of this document.

- So let’s try to look up DRBG and NRBG

Random Bit Generators according to NIST
(the National Institute of Standards and Technology)

- DRBG

  Abbreviation(s) and Synonym(s):
  Deterministic Random Bit Generator

  Definition(s):
  None

- NRBG

  Abbreviation(s) and Synonym(s):
  Non-deterministic Random Bit Generator

  Definition(s):
  None

Statistically independent and unbiased bits

- Suppose you have a sequence of statistically independent but biased bits b_1, b_2, b_3, b_4, … where $P(b_i = 1) = p$
- You can generate a sequence of statistically independent and unbiased bits by the following algorithm:

```
for each pair ($b_{2k-1}$, $b_{2k}$)
    if ($b_{2k-1} \neq b_{2k}$) output $b_{2k-1}$
```

- For example, the sequence 0,0,1,0,1,1,0,0,0,1, … would result in 1,0,…
- The probability that any iteration of the loop outputs a bit is $2p(1-p)$, so the expected time between outputs is $1/p(1-p)$
- This can be improved by using more clever algorithms

Random matroids

- Donald E Knuth of Stanford University wrote an early paper on random matroids:

  Random Matroids, by Donald E Knuth, *Discrete Mathematics* 12 (1975) pp. 341-358

Agenda

- intersections and unions of matroids
- a matroid that is vectorial but not graphical
- a digression on error-correcting codes
- spans, closed sets, and hyperplanes
- a matroid that is neither vectorial nor graphical
- planar graphs and duals of matroids
- a finite number of slides about infinite graphs
- some random slides about random graphs and other random objects
- infinite matroids

Infinite matroids

- One would like to define infinite matroids using an extra axiom:
  An infinite set is independent when all of its
  finite subsets are independent

- This turns out to cause problems with duality, leading Rado (1966) to challenge matroid theorists to come up with a better definition

- In 1969, Higgs proposed B-matroids as a solution, but it took until 2013 to prove that this definition is fully satisfactory
  (see Bruhn, Diestel, Kriesell, Pendavingh, and Wollan: *Axioms for infinite matroids*, available on arXiv)

Independence axioms for B-matroids

- The axiom sets for B-matroids are defined in terms of a maximality property of a set *I* of subsets of E:

  (Max) If I $\subseteq$ X $\subseteq$ E and I $\in$ *I* then the set
  {I'$\in$ *I* | I $\subseteq$ I' $\subseteq$ X } has a maximal element

- The independence axioms can then be written:

  (I1) The empty set is independent
  (I2) Any subset of an independent set is independent
  (I3) If I and J are independent and J is maximal but I is not,
  then there is an *x* in J - I so that (I + *x*) is independent
  (IMax) The set of all independent sets satisfies (Max)

Base axioms for B-matroids

- B-matroids can be defined in terms of bases, *B* :

(B1) *B* is nonempty

(B2) Base exchange - if B_1 and B_2 are bases and x is in B_1-B_2
then there is a y in B_2-B_1 so that $(B_1-x)+y$ is a base

(BMax) The set of all subsets of members of B satisfies (Max)

Circuit axioms for B-matroids

- B-matroids can be defined in terms of circuits, $\mathcal{C}$:

(C1) The empty set is not a circuit

(C2) No circuit is a proper subset of a circuit

(C3) If X is a subset of a circuit C and $\{C_x \mid x \in X\}$ has the property that

$$x \in C_y \Leftrightarrow x = y \quad \text{for all } x, y \in X$$

then for any z that is in C but not in any C_x, there is a circuit C' containing z that is contained in $(C \cup \bigcup_{x \in X} C_x) - X$

(CMax) The set of all $\mathcal{C}$-independent sets satisfies (Max)
(a set is $\mathcal{C}$-independent if none of its subsets are circuits)