J Kohonen

Observe the exceptional exercise session times around Easter. Group H03 has 5B on Wednesday 27.3. Other groups have it after Easter. See SISU for details.

5B Graphs

About terminology. In these exercises, and on this course, a *graph* is a pair $G = (V, E)$, where V (vertices) is any set, and E (edges) is a set of some *two-element sets* $\{u, v\}$, where $u, v \in V$ and $u \neq v$.

Thus between any two distinct vertices, there is either one edge or none (no multiple edges). There are no edges from a vertex to itself (no loops).

Beware that there are many variants, and different people use different terminology. For some people, a "graph" can have loops (expressed as one-element sets $\{u\}$), and they say "simple graph" if loops are not allowed. For us, loops are forbidden by default, and we say "graph with loops" or "loopy graph" if we want to allow them.

In this exercise set, edges do not have direction (all graphs are "undirected"). A related, and important notion is *directed graph* or *digraph*, where edges are ordered pairs (u, v), but more about them later.

5B1 (Collaboration graph) Paul Erdős (1913–1996) was a Hungarian mathematician who was very productive in discrete mathematics. He is also well-known for having collaborated with many people.

In the *Erdős graph*, vertices are people, and two persons have an edge between them if they have been coauthors in a research paper. The distance between two persons is the length of the shortest path between them. The *Erdős number* of a person is his or her distance to Erdős. Thus Erdős has number 0, and his direct coauthors have number 1. If you have not coauthored with Erdős, but you have coauthored with someone who has, then you have Erdős number 2, and so on.

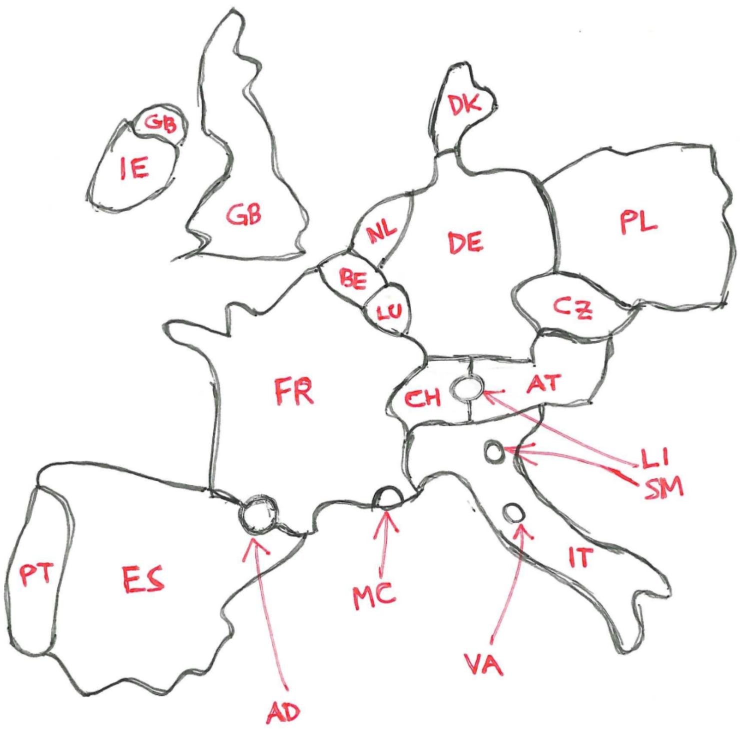
(a) A and B are not direct coauthors, but both have been coauthors with C. What is the largest possible difference between the Erdős numbers of A and B? What is the smallest possible?

(b) D and E have Erdős numbers 3 and 4, respectively. What is the largest possible distance between D and E? What is the smallest possible?

A database of collaboration distances between mathematicians can be found at `https://mathscinet.ams.org/mathscinet/freetools/collab-dist`. Try searching for some mathematicians whose names you know.

Similar collaboration metrics have been constructed in other fields. The Bacon number measures distances of film actors to Kevin Bacon, with a link (edge) between actors if they have appeared in the same film.

5B2 (Country map) Below is a map of some European countries with simplified border shapes, named using their ISO codes. Small countries are drawn out of proportion. Two countries (GB and IE) are separated from all others by sea, and they border only each other. The sea is not a country.



(a) Draw a graph where each country is expressed as a vertex (a point), with an edge between two countries if they have a land border in the map. Draw the graph so that any two edges do not cross (a *planar drawing*). Be careful with France, Spain and Andorra (FR, ES, AD): in our graphs, any two vertices have either zero edges or one edge between them. In the map, the United Kingdom (GB) consists of two disconnected pieces, but in the graph it should be represented as a single vertex, with edges to its land neighbor(s).

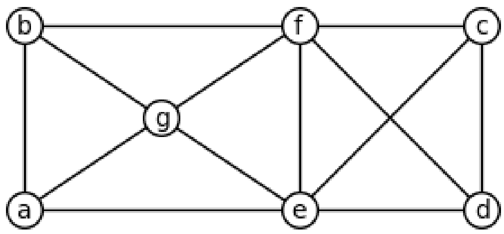
(b) Color the vertices of the graph (or the countries on the map) with four colors, so that neighboring countries do not use the same color.

Hint: If you start by coloring one country, and then proceed sequentially to other countries, what are good strategies? Should you start from countries that have very few neighbors, or leave those countries last?

(c) Find a simple reason why this map cannot be colored with three colors. (Hint: Numbers of neighbors of some countries.)

J Kohonen

5B3 (Radio network) The vertices of the following graph represent radio stations. Every station must choose a channel to use (a positive integer). Two vertices have an edge if they are physically too close to use the same channel.



(a) Find the clique number of the graph.

(b) Find the maximum degree of the graph (the largest of all vertex degrees).

(c) Based on (a) and (b), what can be said about how many channels are needed for this network?

(d) Assign channels to the vertices by applying the greedy coloring algorithm (lecture notes §3.4.4) in order a, b, c, d, e, f, g.

(e) Based on (a)—(d), what can be said about how many channels are needed for this network?

(f) Find a channel assignment that uses as few channels as possible. Use any method you like, e.g. manual experimentation until you have a solution, such that you are sure that it is not possible to reduce the channel count any further. Or apply the greedy algorithm with the vertices in a different order.

(g) If we apply the greedy algorithm in every possible order of the vertices, how many times are we applying it?

(h) (OPTIONAL, not required, no points.) Apply the greedy algorithm on this graph in every possible order, and find what proportion of them use more channels than necessary. (You can use a computer.)

5B4 (All isomorphisms) How many different isomorphisms are there from G to H, when G and H are ...

(a) ... two path graphs, each containing n vertices numbered $1, 2, \dots, n$? For $n = 5$ list all the isomorphisms.

(b) ... two cycle graphs, each containing n vertices numbered $1, 2, \dots, n$? For $n = 4$ list them all. Compare to exercise 5A4 and explain possible connections.

(c) ... two complete graphs, each containing n vertices numbered $1, 2, \dots, n$?

5B5 (Listing graph structures)

(a) For each of $n \in \{1, 2, 3, 4, 5, 6\}$, draw all unlabeled connected graphs that have no cycles, and count the graphs. Be careful that you do not list the same structure twice (i.e. do not draw two graphs that look different but are in fact isomorphic).

Hint: It pays off to be systematic. For example, break into cases by the largest degree in the graph. If the graph has 4 vertices, what is the maximum possible degree that any vertex might have? If there is a vertex with this degree, how do you proceed? If there is no such vertex, consider the next smaller degree. If there is such a vertex, how do you proceed?

For checking: The counts should be $1, 1, 1, 2, 3, 6$.

(b) For each of $n \in \{2, 3, 4, 5, 6, 7\}$, draw all unlabeled connected graphs that have no cycles, and where every vertex has degree at most four.

These are possible arrangements of carbon atoms in alkanes (acyclic saturated hydrocarbons).

Up to $n = 6$ the sequence may look familiar, but don't be fooled! It is not what you think.

(c) (OPTIONAL) Continue the lists from (a) and (b) for as high n as you like. You can check your counts from OEIS. If you search for the counts $1, 1, 1, 2, 3, 6$ you got from (a), there are probably several different integer sequences that begin that way. Browse them until you find an entry that seems to be thing that you are looking for.

5B6 (Hypercube) The n-dimensional *hypercube* (in graph theory) is a graph where each vertex is string of n bits, that is, a tuple $(b_1, b_2, \ldots, b_n)$ where each b_k is zero or one. Two vertices have an edge between them, if their bits are otherwise identical but exactly one bit has different value. For example, in the 3-dimensional hypercube, there is an edge between 000 and 010, but no edge between 000 and 011.

(a) Draw the n-dimensional hypercubes for $n = 1, 2, 3, 4$ as graphs.

In graph theory, the exact geometric shape of a graph drawing is not that important. You can place your vertices any way you like, and the drawings do not have to "look like cubes". It is enough that the edges are between the correct vertices.

But it may be possible and helpful to attempt a geometric-looking drawing. For $n = 4$ it is probably easiest to start with two 3-dimensional hypercubes side by side, and then add the required edges between them. Edges in a graph drawing can be curved if it is helpful.

(b) How many vertices does the n-dimensional hypercube have (in general, for any $n \geq 1$)?

(c) How many edges does the n-dimensional hypercube have?

One possible (and allowed) method is to count the edges from (a) and enter the numbers to OEIS. You will find something, but you still need to explain why the formula is correct.

(d) What is the distance (shortest path length) between the opposite corners $(0, 0, \ldots, 0)$ and $(1, 1, \ldots, 1)$ of an n-dimensional hypercube?

(e) Consider the following scenario: A supercomputer contains 1 024 computing nodes. Each node is understood as a vertex of a suitable n-dimensional hypercube. Two nodes have a direct network connection if they have an edge in the hypercube.

How many direct network connections do we need to build? What is the maximum path length between any two computing nodes in this network?

If we wanted that every two nodes in the computer have a direct connection between them, how many connections would we have to build?

Of course the nodes are physically in our usual 3-dimensional world. The n-dimensional hypercube (where presumably $n > 3$) expresses the *topology* of the network, that is, the connections.

(f) (OPTIONAL) How many colors are needed to color a hypercube?

(g) (OPTIONAL, perhaps difficult) How many different isomorphisms are there from an n-dimensional hypercube to another?

Hypercubes pop up surprisingly often in discrete mathematics, and not necessarily in any apparently geometric context. The subsets of an n-element set, ordered by the $\subseteq$ relation, form a hypercube. So do the divisors of an integer that is a product of distinct primes, such as $2 \cdot 3 \cdot 5 = 30$ or $2 \cdot 3 \cdot 5 \cdot 7 = 210$, ordered by divisibility.

Supercomputer nodes are really sometimes arranged as a hypercube. Part (e) probably explains why: There are not too many direct connections to build, but the distance from any node to any other node is still not very long.