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.

Solution.

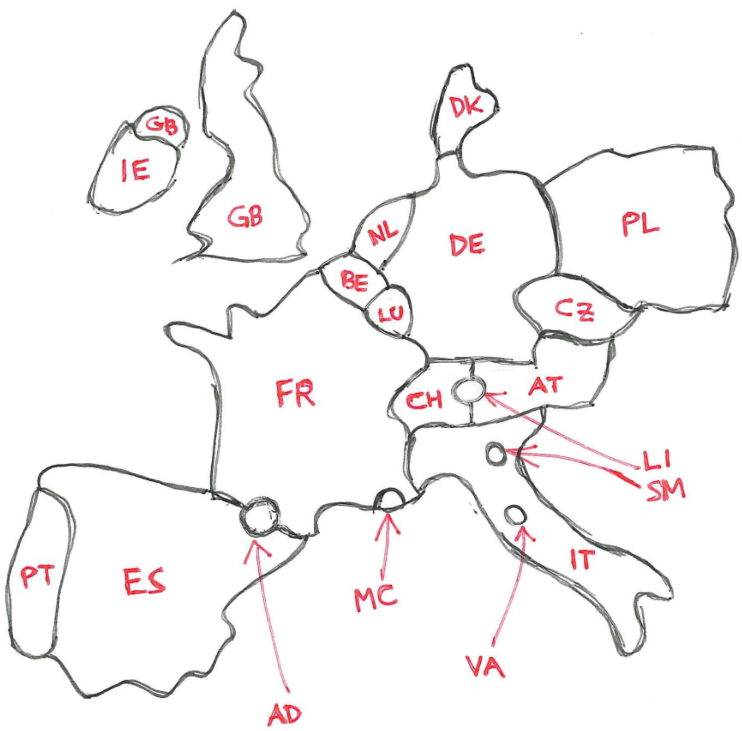
(a) If A has a larger Erdős number than B, then a path connecting A to Erdős is t most two edges longer, as then path between B and Erdős can be appended by first connecting B to C and then C to A. This same reasoning can be applied vice versa. The maximum difference between the Erdős numbers of A and B is two.

In case C has a smaller Erdős number than both A and B, then the Erdős numbers of A and B are both one greater then the Erdős number of C. This means the Erdős numbers of A and B are equal and thus their difference is zero.

(b) The maximum distance possible is seven. This is achieved when all vertices on the path from D to Erdős have their shortest paths to vertices on the path from E to Erdős via Erdős.

The minimum distance between D and E is one. This occurs when D lies on the path from E to Erdős.

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.



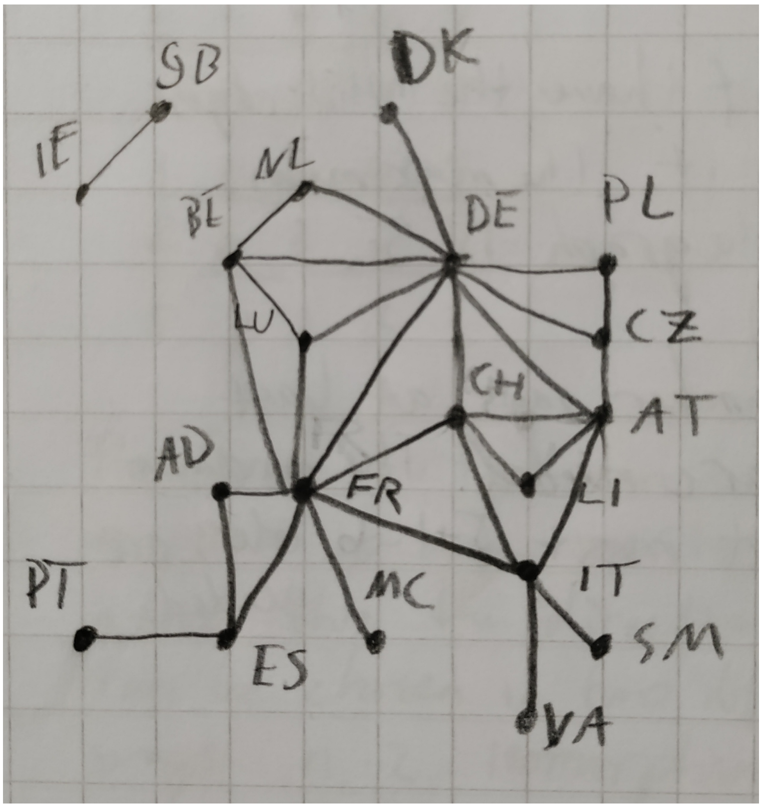
(a) Draw a graph where each country is expressed as a vertex (a point), with an edge between two countries if they have a land border in the map. Draw the graph so that any two edges do not cross (a *planar drawing*). Be careful with France, Spain and Andorra (FR, ES, AD): in our graphs, any two vertices 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 graph (or the countries) with four colors.

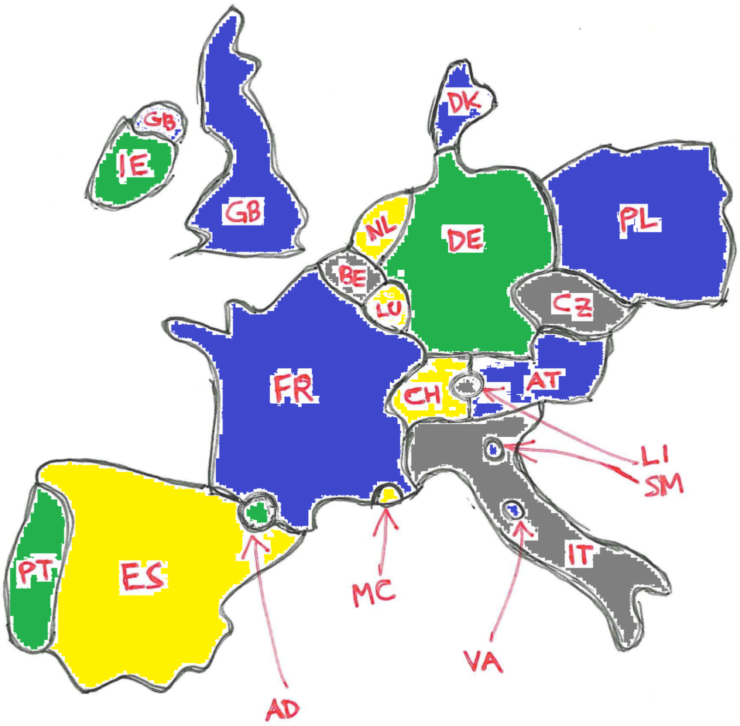
Hint: If you start by coloring one country, and then proceed sequentially to other countries, what are good strategies? Should you start from countries that have very few neighbors, or leave those countries last?

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

Solution.

(a) 

(b) One example, there are many many ways to color this map with four colors.

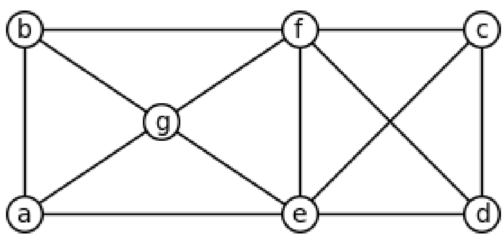


Some thoughts about strategy: Any vertex of degree 3 or less can always be properly colored with one of the colors 1 to 4, no matter how its neighbors have been colored, so perhaps it is best to leave them last. A good strategy might be to start from the K_4 clique (FR, DE, BE, LU), then proceed to other high-degree vertices (CH, AT, IT). If we get this far, the rest will be easy.

If we started with the low-degree vertices, we might make choices that lead to trouble once we reach the high-degree vertices.

(c) France, Germany, Belgium, and Luxembourg are all connected to each other (a clique) and thus each of them needs a unique color. This proves at least four colors are needed.

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.)

Solution.

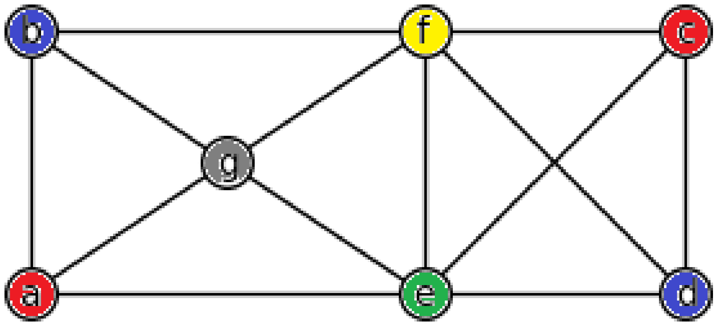
(a) The largest clique in the graph is formed by vertices c, d, e, and f. The clique number the graph is therefore four.

(b) Vertices e and f have the most edges connecting to them. The maximum degree of the graph is five.

(c) From the clique number we can tell at least four colors are needed, as the largest clique cannot be colored with less than four colors. From the maximum degree we can tell the maximum number of colors needed is six. Six is needed if all of the vertices connecting to e or f have a unique color, and no other vertex has this issue with more colors.

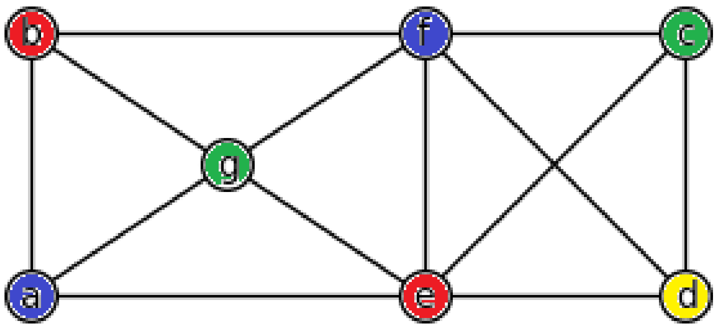
(d)

$$
\begin{aligned}
a &\mapsto 1 \\
b &\mapsto 2 \\
c &\mapsto 1 \\
d &\mapsto 2 \\
e &\mapsto 3 \\
f &\mapsto 4 \\
g &\mapsto 5
\end{aligned}
$$



(e) Coloring with the greedy algorithm yielded a coloring with five colors. This information brings the upper limit for minimum colors needed down to five.

(f) One example for a coloring with four colors: Start by assigning $e \mapsto 1$ and $f \mapsto 2$. Now g must differ from both e and f, so make $g \mapsto 3$. a is now connected to colors 1 and 3 and b is connected to colors 2 and 3, and a and b are connected to each other. Assigning $b \mapsto 1$ and $a \mapsto 2$ is avoids any new colors. Three colors have been used so far and the remaining vertices are a part of a clique of four, the third color to the clique shares a color with g, $c \mapsto 3$. Three colors used and one vertex left, none of the colors used are available for d so $d \mapsto 4$.



(g) For the greedy algorithm the vertices have to be ordered. No special rules apply so there are $n!$ different possible permutation to go through. With $n = 7$ this is 5040 permutations.

(h) Out of the 5040 orders, 888 (17.6%) give five colors, and 4152 (82.4%) give four colors.

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 $1, 2, \ldots, n$? For $n = 5$ list all the isomorphisms.

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

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

Solution.

(a) In a path graph there are two vertices with degree 1 and the rest are of degree 2. In an isomorphism the endpoints, the vertices with degree one must be mapped onto themselves or each other to preserve neighborhoods, the other vertices are then determined. For any $n \geq 2$ there are two isomorphisms. With $n = 5$ $(1, 2, 3, 4, 5) \mapsto (1, 2, 3, 4, 5)$ and $(1, 2, 3, 4, 5) \mapsto (5, 4, 3, 2, 1)$

(b) In these cycle graphs any one vertex can first be mapped onto any of the n vertices in the graph, as all vertices are similar. The next vertex, that is next to the first vertex selected, must be mapped so that it maps onto a vertex next to the image of the first vertex, there are two such vertices. After these two are set, the entire mapping is determined. n choices and two choices, with $n = 4$ there are a total of eight isomorphisms.

$$
\begin{aligned}
(1, 2, 3, 4) &\mapsto (1, 2, 3, 4)\\
(1, 2, 3, 4) &\mapsto (2, 3, 4, 1)\\
(1, 2, 3, 4) &\mapsto (3, 4, 1, 2)\\
(1, 2, 3, 4) &\mapsto (4, 1, 2, 3)\\
(1, 2, 3, 4) &\mapsto (1, 4, 3, 2)\\
(1, 2, 3, 4) &\mapsto (2, 1, 4, 3)\\
(1, 2, 3, 4) &\mapsto (3, 2, 1, 4)\\
(1, 2, 3, 4) &\mapsto (4, 3, 2, 1)
\end{aligned}
$$

As in the dihedral group D_4 that was studied in 5A4, here we also have four rotations and a reflection for each. Rotations correspond to how the mapping of the first vertex is chosen, and reflection to the direction of the cycle.

(c) All vertices are neighbours of each other. The first vertex can be mapped in n different ways. The second can be mapped to any of the remaining $n - 1$ vertices. The third to any of the remaining $n - 2$, etc. A total of $n!$ isomorphisms.

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 $1, 1, 1, 2, 3, 6$.

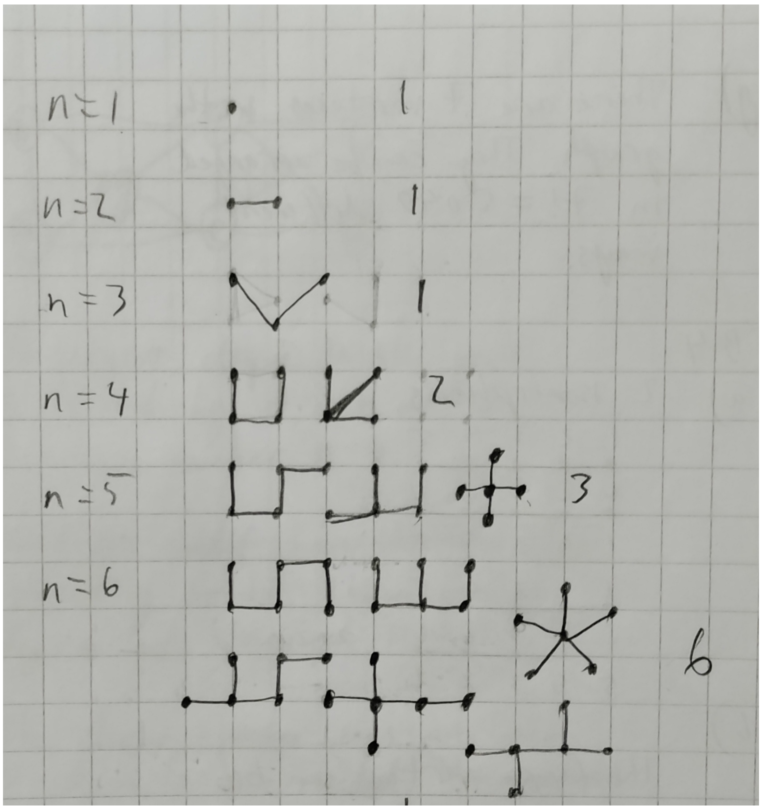
(b) For each of $n \in \{2, 3, 4, 5, 6, 7\}$, draw all unlabeled connected graphs that have no cycles, and where every vertex has degree at most four.

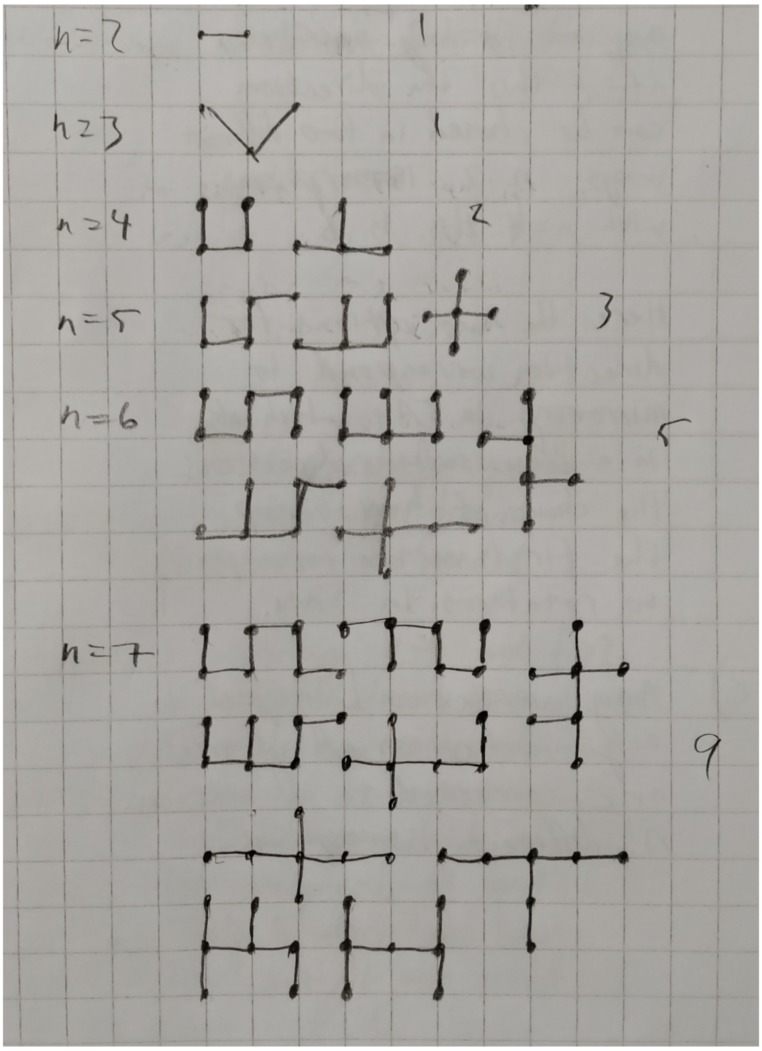
These are possible arrangements of carbon atoms in alkanes (acyclic saturated hydrocarbons).

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.

Solution.

(a) 

(b) 

5B6 (Hypercube) The n-dimensional *hypercube* (in graph theory) is a graph where each vertex is string of n bits, that is, a tuple $(b_1, b_2, \dots, b_n)$ 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 $(0, 0, \dots, 0)$ and $(1, 1, \dots, 1)$ 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?

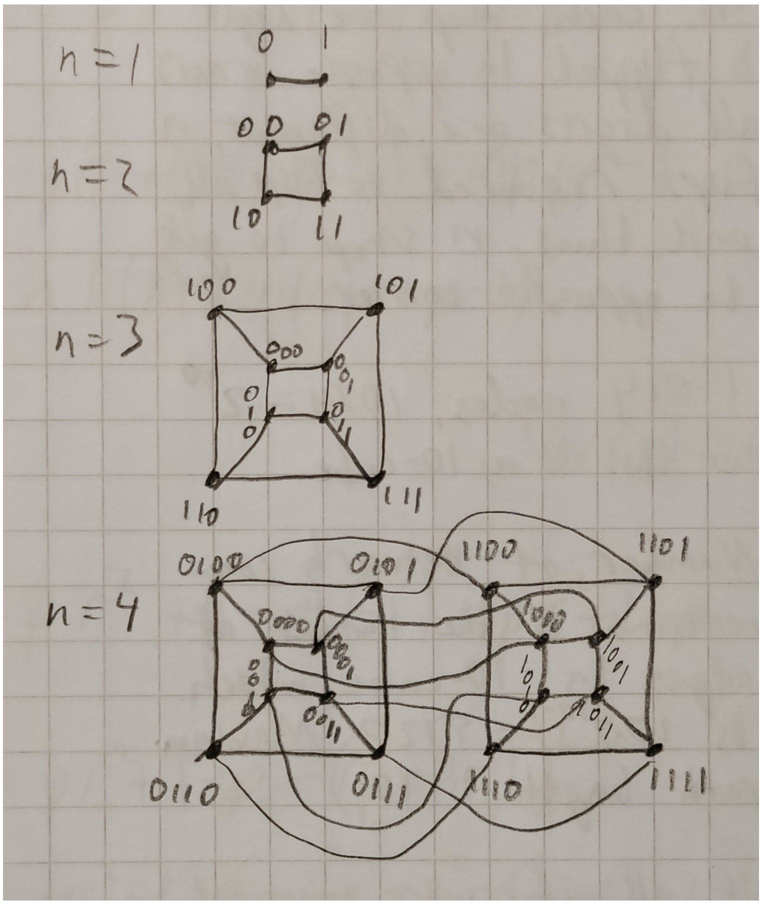
(g) (OPTIONAL, perhaps difficult) How many different isomorphisms are there from an n-dimensional hypercube to another?

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.

Solution.

(a) For $n = 1 \ldots 4$



(b) Each vertex corresponds to a unique binary number of n digits. For each digit there are two options to select from and thus the number of vertices is 2^n.

(c) Adding up the degrees or all vertices gives twice the number of edges. In an n-cube there are 2^n vertices and each has a degree of n. The number of edges in an n-cube is therefore $2^n \cdot n\frac{1}{2} = n2^{n-1}$.

(d) In opposite corners all the n digits of the vertices are different. On each step the value of any one, and exactly one digit is changed, so n steps are needed to change the value of all digits.

(e) The hypercube has 1024 vertices and as $1024 = 2^{10}$ this a 10-cube. The number connections needed is the number of edges in a 10-cube. Using the formula from part (c), $10 \cdot 2^{10-1} = 5120$.

If every node were to be connected, this would be a complete graph of 1024 vertices. Each vertex would have a degree of 1023. Using the same reasoning as in part (c) the number edges is $\frac{1024 \cdot 1023}{2} = 532\ 776$.

(f) Only two colors are needed to color any hypercube. A coloring of two colors can be constructed by coloring all vertices with an even number of ones with color 1 and all vertices with an odd number of ones with color 2. As each edge connects two vertices with exactly one different bit, each vertex is connected only to vertices with one more, or one less one. As the number of ones changes by exactly one, the parity also changes, and therefore no vertex with color 1 is connected to another vertex with color 1 and the same goes for color 2.

(g) First, the origin (00..0) can be freely mapped to any of the 2^n vertices in the other cube. Then, its n neighbors can be mapped into the n neighbors in the other cube in $n!$ ways. After that the mapping is determined. Total $(2^n)(n!)$ ways. For n=1,2,3 this is 2,8,48 (which is enough to find it in OEIS).