Appendix C

Combinatorial Graphs

C.1 Terminology and Examples

In this appendix, we will go over the terminology and some elementary theory of combinatorial graphs. An excellent and complete introduction to the topics appears in [37].

Definition C.1 *A* **combinatorial graph** *or* **graph** *G is a set $V(G)$ of* **vertices** *together with a set $E(G)$ of* **edges**. *Edges are unordered pairs of vertices and, as such, may be thought of as arcs connecting pairs of vertices. The two vertices that make up an edge are its* **ends** *and are said to be* **adjacent**.

An example of a graph appears in Figure C.1. The graph has 16 vertices and 32 edges. In spite of their simplicity, graphs have a boatload of terminology. Prepare to remember.

Definition C.2 *If a vertex is part of the unordered pair that makes up an edge we say that the edge and vertex are* **incident**.

Definition C.3 *The number of edges incident with a vertex is the* **degree** *of the vertex.*

Definition C.4 *If all vertices in a graph have the same degree, we say the graph is* **regular**. *If that degree is k, we call the graph* **k-regular**.

The example of a graph given in Figure C.1 is a 4-regular graph. It is, in fact, the graph of vertices and edges of a 4-dimensional hypercube.

Definition C.5 *A graph is said to be* **bipartite**, *if its vertices can be divided into two sets, called a* **bipartition**, *such that every edge has an end in each set.*

Definition C.6 *A* **subgraph** *of a graph G is a graph H whose vertex and edge sets are both subsets of $V(G)$ and $E(G)$.*

Definition C.7 *A graph is said to be* **connected**, *if it is possible to start at any one vertex and then follow a sequence of pairwise adjacent vertices to any other.*

Definition C.8 *A graph is* **k-connected**, *if the deletion of less than k edges cannot disconnect the graph.*

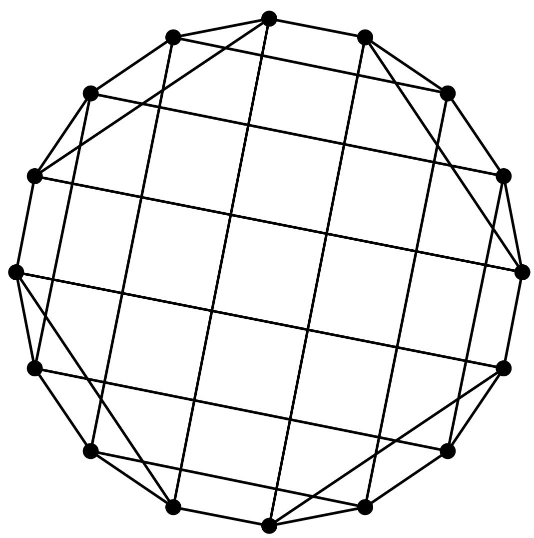

Figure C.1: An example of a graph

The example of a graph given in Figure C.1 is bipartite. Following is the so-called first theorem of graph theory.

Theorem C.1 *The number of vertices of odd degree in a graph is even.*

Proof:

Count the number of pairs of incident vertices and edges. Since each edge is incident on two vertices, the sum is a multiple of two. Since each vertex contributes its degree to the sum, the total is the sum of all the degrees. A sum of integers with an even total has an even number of odd summands, and, so, the number of odd degrees is even. □

This theorem and its proof are included for two reasons. The first is to demonstrate the beautiful technique involved: count something two different ways and then deduce something from the equality of the two answers. The second is to show that even in a very general structure like graphs there are some constraints. Suppose, for example, that you have an evolutionary algorithm that is evolving 3-regular graphs. If you have a mutation that adds vertices, then it must add them in pairs, as a 3-regular graph has an even number of vertices. In some of the other examples, we will see other constraints on graphs. There are quite a lot of named families of graphs. Here are some that are used in this text.

Definition C.9 *The* **complete graph** *on n vertices, denoted K_n has n vertices and all possible edges. An example of a complete graph with 12 vertices is shown in Figure C.2.*

Definition C.10 *The* **complete bipartite graph** *with $n + m$ vertices, denoted $K_{n,m}$ has vertices divided into disjoint sets of n and m vertices and all possible edges that have one end in each of the two disjoint sets. An example of a complete bipartite graph with 8 (4+4) vertices is shown in Figure C.2.*

Definition C.11 *The n-***cycle***, denoted C_n has vertex set $\mathbb{Z}_n$. Edges are pairs of vertices that differ by 1 $(\text{mod } n)$ such that the vertices form a ring with each vertex having two neighbors. A* **cycle in a graph** *is a subgraph that happens to be a cycle.*

Definition C.12 *A* **path** *on n vertices is a graph with n vertices that results from deleting one edge from an n-cycle. A* **path in a graph** *is a subgraph that happens to be a path.*

Definition C.13 *The n-***hypercube***, denoted H_n has the set of all n-character binary strings as its set of vertices. Edges consist of pairs of strings that differ in exactly one position. A 4-hypercube is shown in Figure C.2.*

Definition C.14 *The $n \times m$-***torus***, denoted $T_{n,m}$ has vertex set $\mathbb{Z}_n \times \mathbb{Z}_m$. Edges are pairs of vertices that differ either by 1 $(\text{mod } n)$ in their first coordinate or by 1 $(\text{mod } m)$ in their second coordinate, but not both. These graphs are $n \times m$ grids that wrap (as tori) at the edges. A 12×6 torus is shown in Figure C.2.*

Definition C.15 *The* **generalized Petersen graph** *with parameters n and k is denoted $P_{n,k}$. It has two sets of n vertices. The two sets of vertices are both considered to be copies of $\mathbb{Z}_n$. The first n vertices are connected in a standard n-cycle. The second n vertices are connected in a cycle-like fashion, but the connections jump in steps of size $k(\text{mod } n)$. The graph also has edges joining corresponding members of the two copies of $\mathbb{Z}_n$. The graph $P_{32,5}$ is shown in Figure C.2.*

Definition C.16 *A sequence of pairwise adjacent vertices that is allowed to repeat vertices is called a* **walk**.

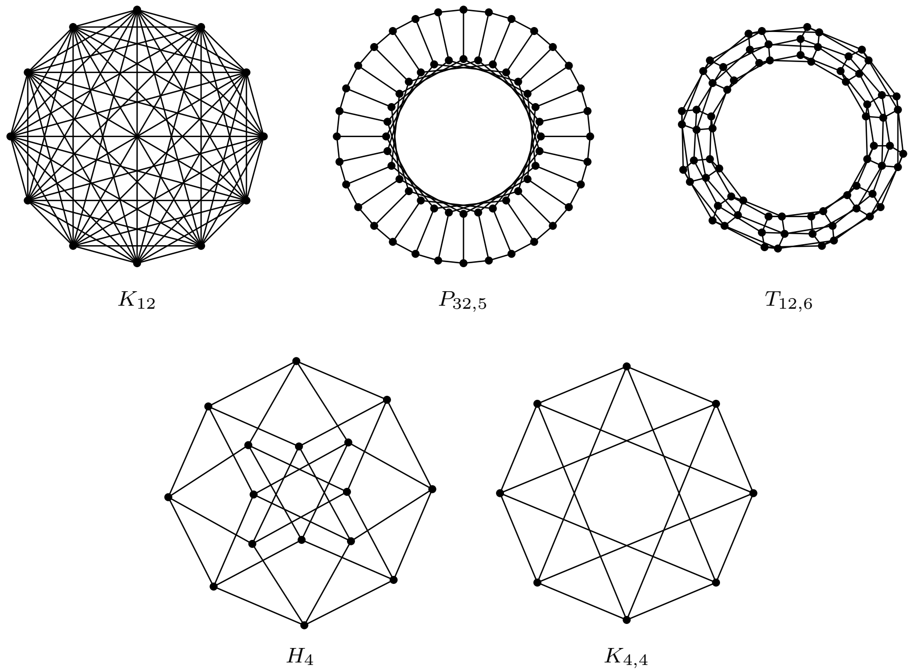

Figure C.2: Examples of complete, Petersen, torus, hypercube, and complete bipartite graphs (These examples are all smaller that the graphs actually used, but are members of the same family of graphs.)

Definition C.17 *A graph which has no cycles as subgraphs is said to be* **acyclic**. *An acyclic, connected graph is called a* **tree**.

Paths are examples of trees. There are a large number of constructions possible on graphs, a few of which are given here.

Definition C.18 *The* **complement** *of a graph G, denoted $\overline{G}$, is a graph with the same vertex set but a complementary set of edges.*

The complement of a 5-cycle is, for example, another, different, 5-cycle; the complement of a 4-cycle is two disconnected edges.

Definition C.19 *If we take a vertex of degree k and replace it with a copy of K_k so that each member of $V(K_k)$ is adjacent to one of the neighbors of the replaced vertex, we say we have* **simplexified** *the vertex. Simplexification of a graph is defined as simplexification of all its vertices.*

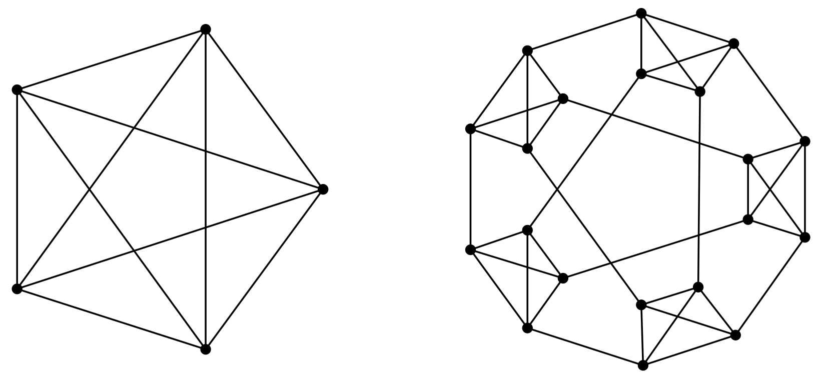

Figure C.3: K_5 and K_5-simplexified

Simplexification is not a construction used much in elementary graph theory, but it is useful for the graph-based evolutionary algorithms discussed in Chapter 13. A picture of a graph and its simplexification are given in Figure C.3.

Definition C.20 *A* **random graph** *is the result of sampling a particular graph from a random process that produces graphs.*

There are more types of random graphs than you can shake a stick at. We, again, give a few examples.

Definition C.21 *A* **random graph with edge probability** α *is generated by examining each possible pair of vertices and, with probability* α*, placing an edge between them. The number of vertices is determined in advance.*

Definition C.22 *A* **random regular graph** *can be generated by a form of random walk, as follows, with thanks to Mike Steel for the suggestion. Begin with a regular graph. A large number of times (think at least twice as many times as the graph has edges) perform the following edge swap operation. Pick two edges that have the property that (i) their ends form a set of 4 vertices and (ii) those 4 vertices have exactly two edges,* $\{a, b\}$ *and* $\{c, d\}$*, between then in the graph. Delete those two edges and replace them with the edges* $\{a, c\}$ *and* $\{b, d\}$*. Again, the number of edges is chosen in advance.*

Definition C.23 *To get a* **random toroidal graph with connection radius** β*, place vertices at random in the unit square. Connect with edges all pairs of vertices at distances at most* β *in the torus created by wrapping the edges of the unit square.*

Definition C.24 *A* **random simplicial graph** *is created by first choosing a number* n *of vertices and a collection of allowed sizes, e.g.,* $\{3\}$ *or* $\{7, 8, 9, 10\}$*. The graph is generated* *by performing the following move k times. A size m is selected at random from the list of allowed sizes. A set of m vertices is selected at random. All pairs of vertices in the selected set not already joined by edges are joined by edges.*

Definition C.25 *A* **simplexification driven random graph** *is created by picking an initial graph and repeatedly choosing a vertex at random and simplexifying it. Since simplexification adds a number of vertices equal to the degree of the vertex it acts on less one, some planning is needed.*

C.2 Coloring Graphs

There are a plethora of problems that involve coloring the vertices or edges of a graph.

Definition C.26 *A* **vertex coloring** *of a graph is an assignment of colors to the vertices of a graph. A vertex coloring of a graph is said to be* **proper**, *if no two adjacent vertices are the same color.*

Definition C.27 *The minimum number of colors in a proper vertex coloring of a graph G is the* **chromatic number** *of a graph, denoted $\chi(G)$.*

Bipartite graphs, for example, have chromatic number 2 (see if you can prove this in one or two lines).

Knowing the chromatic number of a graph is valuable, as can be seen in the following application. Suppose that we have a group of people from which are drawn several committees. Construct a graph with each committee as a vertex and with edges between two vertices if the committees in question share at least one member. Let colors represent time slots for meetings. A proper coloring of the vertices of this graph corresponds to a meeting schedule that allows every member of every committee to be present at each meeting of that committee. The chromatic number is the least number of slots needed for such a schedule.

Definition C.28 *An* **edge coloring** *of a graph is an assignment of colors to the edges of a graph. An edge coloring of a graph is* **proper**, *if no two edges incident on the same vertex are the same color.*

Definition C.29 *The minimal number of colors in a proper edge coloring of a graph G is the* **edge chromatic number** *of a graph, denoted $\chi_E(G)$.*

Proper edge colorings are useful in the development of communications networks. Suppose we have a large number of sites which must send status or other information to all other sites. These sites are the vertices of the graph and the edges represent direct communications links. If we assume each site can communicate with only one other site at a time, then a proper edge coloring of the graph is an efficient algorithm for coordinating communications. If we have a proper edge coloring in n colors, $0, 1, \ldots, n-1$, then processors talk over the edge colored i on each time-step congruent to $i(mod\ n)$. Minimizing the number of colors maximizes usage of the communications links.

There are interesting coloring problems that do not involve proper colorings as well. In *Ramsey Theory*, the goal is to color the edges of a complete graph with some fixed number k of colors, and then find some minimal number of vertices such that any edge coloring in k colors forces a monochromatic subgraph to appear that looks like K_m, $m < k$. For example, if we color the edges of a complete graph on 6 or more vertices red and blue, then there must be a red or a blue triangle (K_3). However, it is possible to bi-edge-color K_5 without obtaining any monochrome triangles. Formally, we say the Ramsey number $R(3,3) = 6$. If you're interested, try to find a red-blue coloring of the edges of K_5 that avoids monochromatic triangles.

Very few Ramsey numbers are known and improving lower bounds on Ramsey numbers is a very hard problem that one can attempt with evolutionary algorithms. Recently Brendan McKay spent 4.3 processor years on UNIX workstations showing that, in order for a complete graph to have either a red K_4 subgraph or a blue K_5 subgraph forced no matter how it was red-and-blue edge colored, the graph must have at least 25 vertices. Formally, $R(4,5) = 25$. This is the hardest of the two-colored Ramsey numbers known so far. There is only one 3-colored Ramsey number known at the time of this writing, $R(3,3,3) = 17$ (neglecting the case in which monochromatic K_2s (edges) are forced). In other words, if we 3-color the edges of a complete graph in 3 colors, then, no matter what coloring we use, we must have a monochromatic triangle, if the complete graph has 17 or more vertices.

The proof that the Ramsey numbers are finite will appear in any good undergraduate combinatorics course, as will several more general definitions of Ramsey numbers and a plethora of Ramsey-style problems. The Ramsey numbers are pervasive in existence proofs in combinatorics and discrete math; so, additional information about a Ramsey number usually turns out to be additional information about many, many other problems as well.

C.3 Distances in Graphs

If we define the distance between two vertices to be the length of the shortest path between them (and define the distance to be infinite if no such path exists), then graphs become *metric spaces.*

Definition C.30 *A* **metric space** *is a collection of points, in this case the vertices of a graph, together with a function d (distance) from pairs of points to the real numbers, which has three properties:*

(i) For all points p, $d(p, p) = 0$,
(ii) For all pairs of points $p \neq q$, $d(p, q) > 0$, *and*
(iii) For all triples of points p, q, r, $d(p, q) + d(q, r) \geq d(p, r)$.

The third property is called the *triangle inequality*.

Definition C.31 *The* **diameter** *of a graph is the maximum distance between any two vertices of a graph.*

As we will see in Chapter 13, the diameter is sometimes diagnostic of the behavior of a graph-based evolutionary algorithm.

Definition C.32 *The* **eccentricity** *of a vertex is the largest distance from it to any other vertex in the graph.*

Notice that the diameter is then the maximum eccentricity of a vertex.

Definition C.33 *The* **radius** *of a graph is the minimum eccentricity (and it is not usually half the diameter, graphs aren't circles).*

Definition C.34 *The* **center** *of a graph is the set of vertices that have minimum eccentricity.*

Definition C.35 *The* **periphery** *of a graph is the set of vertices that have maximal eccentricity.*

Definition C.36 *The* **annulus** *of a graph are those vertices that are not in the periphery or the center.*

The several terms given above for different eccentricity-based properties are useful for classifying the vertices of network graphs in terms of their probable importance. Peripheral vertices tend to be lower traffic, while central vertices are often high traffic.

Definition C.37 *A* **dominating set** *in a graph is a set* D *of vertices with the property that every vertex is either in* D *or adjacent to a member of* D.

For graphs representing guards and lines of sight, or vital services and minimal feasible travel times to reach them, small dominating sets can be quite valuable. There may be reasons that we want dominating sets that are only in the periphery of a graph (imagine a town in which affordable land is only at the "edge" of town). Vertices in the center of the graph are more likely to cover lots of other vertices, and so it may be wise to choose them when searching for small dominating sets. The problem of locating minimal dominating sets is thought to be intractable, but evolutionary algorithms may be used to locate tolerably small dominating sets.

C.4 Traveling Salesmen

It is possible to generalize the notion of distance in graphs by placing weights on their edges so that, instead of adjacent vertices being at distance 1, they are at a distance given by the edge weight. In this case the edge weights may represent travel costs or distances.

Definition C.38 *The* **Traveling Salesmen Problem***, starts with a complete graph that has cities as its vertices and the cost of traveling between cities as edge weights. What we desire is an ordered list of all the cities that corresponds to a minimal cost (total of edge weights) cycle in the graph that visits all the cities.*

Finding exact solutions to this problem is almost certain to be intractable (NP-complete for the computer science majors among you), but evolutionary algorithms can be used to find approximate answers (see Section 7.2). The Traveling Salesman Problem is a standard test problem for evolutionary algorithms that operate on genes that are ordered lists without repetition (in this case the list is the salesman's itinerary).

C.5 Drawings of Graphs

Definition C.39 *A drawing of a graph is a placement of the vertices and edges of a graph into some space, e.g., the Cartesian plane.*

There are a number of properties of drawings that can be explored, estimated, or optimized with evolutionary algorithms. In Chapter 3, we discussed evolutionary algorithms that tried to minimize the crossing number of a graph when the edges were drawn as line segments.

Definition C.40 *The* **crossing number** *of a graph is the minimum number of times one edge crosses another in any drawing.*

Definition C.41 *A graph is said to be* **planar***, if it can be drawn with zero edge crossings in the Cartesian plane.*

Another property of a graph related to drawings is the thickness of a graph.

Definition C.42 *The* **thickness** *of a graph is the minimum number of colors in an edge-coloring of the graph that has the property that all the induced monochromatic graphs are planar.*

A planar graph thus has thickness 1. Thickness gives a useful measure of the complexity of a graph. An electrical circuit with a thickness of 3 might need to be put on 3 stacked circuit boards, for example. Many other problems concerning drawing of graphs exist but require a knowledge of topology beyond the scope of this text. If you are interested, look for books on topological graph theory that discuss the genus of a graph or the M-pire (empire) problem. The problem of embedding topological knowledge in a data structure that is to be manipulated by an evolutionary algorithm is a subtle one.