Basic Graph Theory
Michael Albert

1 Introduction

The concept of a *graph* is one of the most widely occurring ones in discrete mathematics. Basically, in any situation where we are given a set of objects, and a relationship between pairs of objects, we can model the problem in terms of a graph, and this may help with understanding its solution, and almost always does help with presenting a solution! In Olympiad style problems, the "objects" are often things like people or cities, and the "relationship" something like friendship (for people), or being connected by a road (for cities). Sometimes there might be two or more type of objects (students and questions on a test) and a relationship that holds only between objects of one type and of the other (the questions that a particular student can do.) Before carrying on we should establish the standard terminology.

A *graph* consists of a set of objects called *vertices* and a collection of pairs of vertices called *edges*. In these notes we will deal only with *simple graphs* – in a simple graph the edges are *unordered pairs* of distinct vertices, and no two edges connect the same pair of vertices. That is, there are no loops (edges from a vertex to itself), nor multiple edges between a given pair of vertices. All of these requirements can be (and sometimes must be) relaxed (loops might be allowed, multiple edges might be allowed, edges might have a direction associated with them), but most of the theory can be developed for the case of simple graphs and modified as necessary in a given problem. A useful standing convention will be that capital letters stand for sets of things (so G and H are graphs), while lower case letters stand for individual things (so v and w are vertices, and e and f are edges.)

Let a graph G be given. Two vertices v and w are *adjacent* if $\{v, w\}$ is an edge, in this case we also say that w is a *neighbour* of v (and vice versa.) They are *independent* otherwise. The *degree*, $d(v)$ of a vertex v is the number of edges of which it is an endpoint. A graph is *regular* if all of its vertices have the same degree. A *walk of length n* is a sequence of vertices $v_0, v_1, \ldots, v_n$ such that for each $0 \leq i < n$, v_i and v_{i+1} are adjacent. If a walk contains no repeated vertices it is called a *path*. If $v_0 = v_n$, but otherwise there are no repeated vertices, then it is called a *cycle*. A graph is called *connected* if, for any pair of vertices v and w there is a walk from v to w. Generally, the *component* of a vertex v is the set of vertices w for which there is a walk from v to w (note that this includes the vertex v itself, owing to the trivial walk of length 0 from v to v.)

Exercise: Show that if there is a walk from v to w in G then there is also a path from v to w in G (Hint: consider a shortest walk connecting these two vertices.) In this case show that the components of v and w are the same.

A graph H, obtained from a graph G by deleting some of the vertices and edges is called a *subgraph* of G. If H is obtained only by deleting vertices, that is, for any pair of vertices $v, w \in H$, if $\{v, w\}$ is an edge of G then it is also an edge of H then H is called an *induced subgraph* of G. Two graphs G and H that are identical except for the names of their vertices are called *isomorphic* (formally: G and H are isomorphic if there is a bijection f, between the vertices of G and the vertices of H such that $\{v, w\}$ is an edge of G if and only if $\{f(v), f(w)\}$

is an edge of H.) Generally speaking, when dealing with graphs we are only concerned about their structure, and so we can treat isomorphic graphs as identical.

2 Degrees and Connectedness

The most fundamental result about graphs (and one of the most frequently useful ones in Olympiad style problems) is:

Theorem 1 (Handshake Lemma) *The sum of the degrees of the vertices in a graph is equal to twice the number of edges.*

The proof is embarrassingly easy. If we add up the degrees of the vertices, then each edge gets counted twice (once per endpoint). Another way to count each edge twice is to simply take twice the number of edges. So, these two numbers must be equal. The most usual use of this theorem is in observing that the sum of the degrees of the vertices in a graph must be even, or equivalently that the number of vertices of odd degree must be even. In particular a regular graph of odd degree must have an even number of vertices.

Exercise: Show that this, together with the trivial requirement that the degree of a vertex must be smaller than the number of vertices is the only restriction on regularity. That is, if $n > k$ and nk is even, then there is a regular graph with n vertices, each of degree k.

Obviously there is some minimum number of edges required in order for a graph to be connected. Since a connected graph has exactly one component, that, and a bit more, is captured in the next theorem:

Theorem 2 *A graph G with n vertices and k edges has at least $n - k$ components.*

Proof: Let such a graph be given, and label the edges e_1 through e_k arbitrarily. Consider the subgraphs G_i of G for $0 \leq i \leq k$ where the vertices of G_i are the same as those of G, but G_i only has edges e_1 through e_i inclusive. That is, consider building G "one edge at a time". Since G_{i+1} is obtained from G_i by adding a single edge, if two vertices were already in the same component of G_i then they are still in the same component in G_{i+1}. But, if the endpoints of the edge were in different components, then those two components (and only those two) merge into a single component – that is, the number of components decreases by at most 1. Since the number of components of G_0 is n, the number of components of $G = G_k$ is at least $n - k$. ∎

A connected graph that contains no cycles is called a *tree.*

Theorem 3 *A connected graph with n vertices is a tree if and only if it has exactly $n - 1$ edges.*

Proof: Let G be a connected graph on n vertices. Suppose that G has n or more edges. Then, in the development of G considered previously, at least one edge must join two vertices which were already in the same component. Then it, and the edges that previously connected those vertices forms a cycle, and so G is not a tree. Conversely, suppose that G has exactly $n - 1$ edges. If any edge belonged to a cycle, we could delete it and obtain a connected graph with n vertices and $n - 2$ edges, which is a contradiction. So, no edge belongs to a cycle, and G is a tree. ∎

An *Euler Circuit* of a graph G is a closed walk in G that contains each edge exactly once.

Theorem 4 *A graph G with no vertices of degree 0 has an Euler Circuit if and only if it is connected and every vertex has even degree.*

Proof: The necessity of the two conditions is clear. The existence of a circuit containing each edge (and hence visiting each vertex as there are no vertices of degree 0) immediately shows that there is a walk between any two vertices and so the graph is connected. Each time an Euler Circuit arrives at a vertex on some edge, it must leave on a different edge. This pairing establishes that the degree of each vertex must be even. So, it remains to prove that the two conditions are sufficient.

Let a connected graph G with n vertices in which every vertex is of even degree be given. Suppose that the result is true for all graphs with fewer edges (it holds trivially for the first three cases.) Since the sum of the degrees of G is at least $2n$, G has at least n edges and hence at least one cycle, C. Let H be the graph obtained from G by deleting the edges of C and any resulting vertices of degree 0. The degrees of the vertices of H are all even. H may not be connected, but each component of H is and satisfies the conditions of the theorem, hence has an Euler Circuit. Furthermore, each component of H must contain a vertex in C (otherwise, it would already have been a separate component in G, but G was connected). So, we can "glue" an Euler Circuit of each component of H on to C when C first visits that component. The resulting closed walk is an Euler Circuit of G as required. $\blacksquare$

3   Ramsey's Theorem

Another popular source of Olympiad style problems is a result called Ramsey's Theorem (and its extensions or specializations). Put in its simplest form it expresses the idea that in any very large group of people, there must be a fairly large group of acquaintances, or a fairly large group of mutual strangers. To state it, just a couple more words of terminology. In a graph G a subset, K, of the vertices is called a *clique* if every pair of vertices from K are adjacent. A subset, I, of the vertices is called an *independent set*, if every pair of vertices from I are independent. We begin with the first non-trivial case.

Example 1 *Every graph with 6 vertices contains either a clique or an independent set of size 3.*

Proof: Let a graph G with 6 vertices be given. Choose a vertex v, and consider its degree. If its degree is less than three, then there are at least three vertices that are independent of v. If any two of them are independent, then those two, together with v form an independent set of size three. On the other hand if they all form a clique, then we have a clique of size at least three. Alternatively, if the degree of v is at least three, then consider the neighbours of v. If any two of them are adjacent then, with v, we have a clique of size three. If they are all independent then we have an independent set of size at least three. Regardless, we're done. $\blacksquare$

Now for the general version.

Theorem 5 *Let k and i be positive integers. There is a number $R(k, i)$ with the property that any graph having at least $R(k, i)$ vertices contains either a clique of size k or an independent set of size i.*

Proof: We can take (trivially) $R(1, i) = R(k, 1) = 1$ (or, slightly less trivially, $R(2, i) = i$, $R(k, 2) = k$.) Suppose that the result holds for all pairs k' and i' with $k' + i' < k + i$. Consider a graph G with $R(k - 1, i) + R(k, i - 1)$ vertices. Choose an arbitrary vertex v in G. If $d(v) \geq R(k - 1, i)$, then there is either

a clique of size $k-1$ or an independent set of size i among the neighbours of v. In the first case add v to get a clique of size k in G, and in the second case we're already done. If $d(v) < R(k-1, i)$ then v is independent of at least $R(k, i-1)$ vertices and we proceed similarly. ∎

Note that the proof does not necessarily tell us the smallest possible value for $R(k, i)$, it just tells us that $R(k-1, i) + R(k, i-1)$ is small enough. Determining exact values of $R(k, i)$, or even of $R(k, k)$ for all but the smallest values of k seems to be a very difficult problem. Erdős is said to have suggested that "If aliens arrived on Earth and said they would destroy the world in 10 years unless we could tell them $R(6, 6)$, then everyone should work on figuring out $R(6, 6)$; but if they ask for $R(7, 7)$ then everyone should try to figure out how to destroy the aliens."

4   Turan's theorem

Another part of extremal graph theory deals with the question of how many edges a graph on n vertices can have before certain subgraphs *must* appear. Alternatively put, we want to find the maximum possible number of edges in a graph where certain subgraphs are forbidden. Let's consider a specific example:

Example 2 *How many edges can a graph with 2009 vertices have if it contains no triangle (clique of size 3)?*

Suppose that we have a graph G with 2009 vertices, no triangles, and as many edges as possible. Consider two vertices v and w that are not adjacent. Let X be the set of neighbours of v and Y the set of neighbours of w. These two sets must have the same size, as otherwise (say $|X| > |Y|$) we could form a new graph G' just like G except that the edges from w to Y have been eliminated and replaced by edges from w to X. This would increase the number of edges, but not create any triangles, which would contradict our choice of G. In fact, we can make this change anyhow, and after a sequence of such changes we'd obtain a new graph, with the same number of edges, in which any vertex independent of v has the same neighbours as v. Since these neighbours form an independent set (no triangles), we see that the graph looks like two sets X, and $G \setminus X$, with every possible edge between them. Therefore, the total number of edges is $|X||G \setminus X|$, and we seek to maximize that. It's well known (and easy to show) that we should have $|X|$ as near to $(1/2)|G|$ as possible – that is, the maximum possible number of edges is 1004×1005.

This argument generalizes – the maximum number of edges in a graph on n vertices with no triangles is achieved by a *complete bipartite graph* of this type, and is $\lfloor n/2 \rfloor \lceil n/2 \rceil$. In fact, it generalizes even more – the maximum number of edges in a graph having no clique of size r is obtained by dividing the vertices up into $r-1$ sets, as near to equal sizes as possible, and connecting every pair of vertices that belong to different sets.