Olympiad Combinatorics

Pranav A. Sriram

August 2014

Copyright notices

About the Author

Pranav Sriram graduated from high school at The International School Bangalore, India, and is currently pursuing undergraduate studies at Stanford University.

8. GRAPH THEORY

Introduction

Graphs rule our lives: from Google Search to molecular sequencing, flight scheduling to Artificial Intelligence, graphs are the underlying mathematical abstraction fueling the world's most advanced technology. Graphs are also pervasive in models of speech, group dynamics, disease outbreaks and even the human brain, and as such play a crucial role in the natural and social sciences. Where does this versatility come from? Problems in several of the above-mentioned fields involve, at their core, entities existing in complex relationships with each other: hyperlinks between web pages, subject-object relationships between words, flights between cities and synapses between neurons. The power of graph theory stems from the simplicity and elegance with which graphs can model such relationships. Once you've cast a problem as a problem on graphs, you have at your disposal powerful machinery developed by mathematicians over the centuries. This is the power of abstraction.

This chapter is by no means an exhaustive reference on the subject – graph theory deserves its own book. However, we will see several powerful lemmas and techniques that underlie a vast majority of Olympiad and classical graph theory problems, and hopefully build plenty of graph theoretic intuition along the way.

In the final section of this chapter, we will leverage the power of graphs mentioned in the first paragraph to solve Olympiad problems that initially appear to have nothing to do with graphs.

We now recall some results we have proven in earlier chapters, that prove to be extremely useful both in Olympiad problems and in classical graph theory problems. We also advise the reader to go over their proofs again, because the proof techniques for these results are as important as the results themselves.

Some Useful Results

(i) In a graph $\boldsymbol{G}$ with n vertices, suppose no vertex has degree greater than Δ. Then one can color the vertices using at most $\Delta+1$ colors, such that no two neighboring vertices are the same color. [Chapter 1, example 1]

(ii) In a graph $\boldsymbol{G}$ with V vertices and E edges, there exists an induced subgraph $\boldsymbol{H}$ with each vertex having degree at least E/V. (In other words, a graph with average degree d has an induced subgraph with minimum degree at least $d/2$) [Chapter 1, example 3]

(iii) Given a graph $\boldsymbol{G}$ in which each vertex has degree at least $(n-1)$, and a tree $\boldsymbol{T}$ with n vertices, there is a subgraph of $\boldsymbol{G}$ isomorphic to $\boldsymbol{T}$. [Chapter 2, example 3]

(iv) In a graph $\boldsymbol{G}$, if all vertices have degree at least δ, then there exists a path of length at least $\delta+1$. [Chapter 4, example 6]

(v) The vertex set $\boldsymbol{V}$ of a graph $\boldsymbol{G}$ on n vertices can be partitioned into two sets $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ such that any vertex in $\boldsymbol{V}_1$ has at least as many neighbors in $\boldsymbol{V}_2$ as in $\boldsymbol{V}_1$ and vice versa. [Chapter 4, example 8]

(vi) A *tournament* on n vertices is a directed graph such that for any two vertices u and v, there is either a directed edge from u to v or from v to u. A *Hamiltonian path* is a path passing through all the vertices. Every tournament has a Hamiltonian path. [Chapter 4, example 10]

More Useful Results and Applications

Dominating Sets

In a graph $\boldsymbol{G}$ with vertex set $\boldsymbol{V}$, a subset $\boldsymbol{D}$ of $\boldsymbol{V}$ is set to be a *dominating set* if every vertex v is either in $\boldsymbol{D}$ or has a neighbor in $\boldsymbol{D}$. The next lemma tells us that under certain simple conditions, there exists a fairly small dominating set.

Lemma 8.1: If $\boldsymbol{G}$ has no isolated vertices, then it has a dominating set of size at most $\frac{|V|}{2}$.

Proof: By **(v)**, there exists a bipartition $\boldsymbol{V} = \boldsymbol{V}_1 \cup \boldsymbol{V}_2$ so that every vertex in $\boldsymbol{V}_1$ has at least as many neighbors in $\boldsymbol{V}_2$ and vice versa. Since each vertex has degree at least 1, this implies that every vertex in $\boldsymbol{V}_1$ has at least one neighbor in $\boldsymbol{V}_2$ and vice versa. Thus both $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ are dominating sets. One of them has at most $\frac{|V|}{2}$ vertices and we are done.

Remark: In the next chapter we will show that if the minimum degree in an n-vertex graph $\boldsymbol{G}$ is $d > 1$, then $\boldsymbol{G}$ has a dominating set containing at most $n\,\frac{1+\ln(d+1)}{d+1}$ vertices.

Spanning Trees

Recall that a *spanning subgraph* of a graph $\boldsymbol{G}$ is a subgraph of $\boldsymbol{G}$ containing all of the vertices of $\boldsymbol{G}$. A *spanning tree* in $\boldsymbol{G}$ is a spanning subgraph that is a tree (that is, it is acyclic). Note that if

G is **not** connected, it cannot have a spanning tree (because otherwise there would be a path between every pair of vertices along edges in this tree, contradicting disconnectedness). Do all connected graphs have spanning trees?

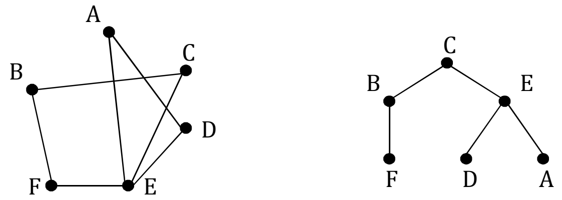

*Figure 8.1. A graph **G** and a spanning tree of **G***

Lemma 8.2: Every (finite) connected graph ***G*** = (***V***, ***E***) has a spanning tree.

Proof: Delete edges from ***G*** as follows. As long as there is at least one cycle present, take one cycle and delete one edge in that cycle. Notice that this procedure cannot destroy connectivity, so the graph obtained at each stage is connected. This process cannot continue indefinitely (we are dealing with finite graphs), so eventually we get a connected graph with no cycles. This is the required spanning tree (note that all the vertices of ***V*** are still present since we only deleted edges).

Spanning trees arise very often in the study of graphs, especially in optimization problems. I like to think of them as the "skeleton" of the graph, since they are in a sense the minimal structure that is still connected on its own. Their main usage on Olympiad problems is that instead of focusing on general graphs ***G*** which may have a complicated structure, we can sometimes find what we are looking for just by taking a spanning tree. For our purposes, all you really need to know about spanning trees is

a) They are trees
b) They exist (unless ***G*** isn't connected).

We'll now use our arsenal of lemmas to reduce some rather challenging Olympiad problems to just a few lines.

Example 1 [Based on ELMO Shortlist 2011, C7]
Let $\boldsymbol{T}$ be a tree with t vertices, and let $\boldsymbol{G}$ be a graph with n vertices. Show that if $\boldsymbol{G}$ has at least $(t-1)n$ edges, then $\boldsymbol{G}$ has a subgraph isomorphic to $\boldsymbol{T}$.

Answer:
By (ii), $\boldsymbol{G}$ has a subgraph $\boldsymbol{H}$ such that all vertices in $\boldsymbol{H}$ have degree at least $t-1$ (in $\boldsymbol{H}$). Applying (iii), $\boldsymbol{H}$ has a subgraph isomorphic to $\boldsymbol{T}$. ■
Remark: This is probably the shortest solution to a problem in this book. Note that it would actually be quite a difficult problem if you didn't know the super useful lemmas listed above!

Example 2 [Based on ELMO Shortlist 2011, C2]
Let $\boldsymbol{G}$ be a directed graph with n vertices such that each vertex has indegree and outdegree equal to 2. Show that we can partition the vertices of $\boldsymbol{G}$ into three sets such no vertex is in the same set as both the vertices it points to.

Answer:
Take a partition that maximizes the number of "crossing edges", that is, edges between distinct sets. If some v belongs to the same set as both of its out-neighbors, moving v to one of the other two sets (whichever has fewer in-neighbors of v) will add 2 crossing edges but destroy at most 1 old one. Then we get a partition with even more crossing edges, contradiction. Thus the original partition indeed works. ■
Remark: This is essentially the same idea used to prove (v).

Example 3 [Russia 2001]
A company with $2n+1$ people has the following property: For each group of n people, there exists a person amongst the remaining $n+1$ people who knows everyone in this group. Show that there exists a person who knows all the people in the company. (As

usual, knowing is mutual: A knows B if and only if B knows A).

Answer:
Assume to the contrary that no one knows everyone else. Construct a graph $\boldsymbol{G}$ with $2n+1$ vertices representing the people, and an edge between two vertices if and only if those two people **do not** know each other. Our assumption implies that every vertex has degree at least 1. Now applying lemma 8.1, there exists a dominating set of $\boldsymbol{G}$ containing n vertices. This means that each of the other $n+1$ vertices has a neighbor in this set of n vertices. In other words, no person outside this set of n people knows everyone in this set, contradicting the problem statement. This contradiction establishes the result. ■

The Extremal Principle

We've already encountered the extremal principle several times in various forms. The true power of this technique lies in its ubiquitous use in graph theory. In each of the next five examples, the step in which we use the extremal principle is marked in bold letters.

Example 4 [IMO Shortlist 2004, C3]
The following operation is allowed on a finite graph: choose any cycle of length 4 (if one exists), choose an arbitrary edge in that cycle, and delete this edge from the graph. For a fixed integer $n \geq 4$, find the least number of edges of a graph that can be obtained by repeated applications of this operation from the complete graph on n vertices (where each pair of vertices is joined by an edge).

Answer:
Clearly the answer cannot be less than $n-1$, since the graph obtained at each stage will always be connected. We claim that the

graph obtained at each stage is also non bipartite. This will imply that the answer is at least n (since a graph with n-1 vertices is a tree which is bipartite).

K_n is non-bipartite (it has a triangle), so suppose to the contrary that at some stage the deletion of an edge makes the graph bipartite. Consider the **first** time this happens. Let edge AB from the 4 cycle ABCD be the deleted edge. Since the graph is non bipartite before deleting AB but bipartite afterwards, it follows that A and B must lie on the same side of the partition. But since BC, CD, DA are edges in the now-bipartite graph, it follows that C and A are on one side and B and D are on the other side of the bipartition. Contradiction.

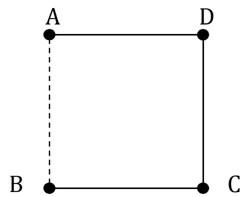

Figure 8.2.

To show n can be achieved, let the vertices be V_1, V_2, ..., V_n. Remove every edge V_iV_j with $3 \le i < j < n$ from the cycle $V_2\, V_i\, V_j\, V_n$. Then for $3 \le i < n$ delete edges $V_2\, V_i$ and $V_i\, V_n$ from cycles $V_1\, V_2\, V_i\, V_n$ and $V_1\, V_i\, V_n\, V_2$ respectively. This leaves us with only n edges: $V_1\, V_i$ for $2 \le i \le n$ and $V_2\, V_n$. ■

Remark: You may have been tempted to guess that the answer is $(n-1)$, since the problem looks like the algorithm for obtaining a spanning tree. While guessing and conjecturing is an important part of solving problems, it is important to verify these guesses by experimenting a bit before trying to prove the guess. The process of realizing your guess was wrong may give you a clue as to how to proceed with the proof. In this example, you may have noticed that the graph you end up with always had an odd cycle, which would lead to the correct claim that the graph obtained is never bipartite.

Example 5 [Croatian TST 2011]
There are n people at a party among whom some are friends. Among any 4 of them there are either 3 who are all friends with each other or 3 who aren't friends with each other. Prove that the people can be separated into two groups $\boldsymbol{A}$ and $\boldsymbol{B}$ such that $\boldsymbol{A}$ is a clique (that is, everyone in $\boldsymbol{A}$ is friends with each other) and $\boldsymbol{B}$ is an independent set (no one in $\boldsymbol{B}$ knows anyone else in $\boldsymbol{B}$). (Friendship is a mutual relation).

Answer
Construct a graph $\boldsymbol{G}$ with vertices representing people and edges between two people if they are friends. The natural idea is to let $\boldsymbol{A}$ be the **largest clique in G**, and the remaining people as $\boldsymbol{B}$. We prove that this works.

If $\boldsymbol{A}$=$\boldsymbol{G}$ or $|\boldsymbol{A}|$=1 we are trivially done, so assume that $n > |\boldsymbol{A}| \geq 2$. We only need to show that $\boldsymbol{B}$ is independent, that is, $\boldsymbol{G}$-$\boldsymbol{A}$ is independent. Assume to the contrary v_1, v_2 belong to $\boldsymbol{G}-\boldsymbol{A}$ and v_1v_2 is an edge. Since $\boldsymbol{A}$ is the largest clique, there exists u_1 in $\boldsymbol{A}$ such that v_1u_1 is not an edge (otherwise we could add v_1 to $\boldsymbol{A}$, forming a bigger clique).

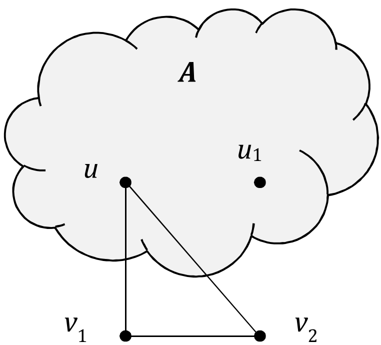

Figure 8.3.

If u_1v_2 is not an edge, then let u be any other vertex in $\boldsymbol{A}$. Since v_1v_2 and uu_1 are edges, by the condition of the problem there must be a triangle amongst these four vertices. The only possibility is uv_1v_2 since u_1v_2 and u_1v_1 are not edges. Then uv_1 and uv_2 are edges for all u in $\boldsymbol{A}$, so $\boldsymbol{A}$ U $\{v_1, v_2\}\setminus \{u_1\}$ is a larger clique, contradiction.

Similarly, if v_2u_1 is an edge, then for all u in $\boldsymbol{A}$ either v_2u_1u or v_2v_1u must be a triangle. In either case v_2u is an edge for all u in $\boldsymbol{A}$. Thus $\boldsymbol{A}$ U $\{v_2\}$ is a larger clique, contradiction. ■

Example 6 [Degree vectors]

A vector $\boldsymbol{v}$ = [d_1 d_2 … d_k] with $d_1 \geq d_2 \geq \ldots \geq d_k$ is said to be a *graphical* vector if there exists a graph $\boldsymbol{G}$ with k vertices $x_1, x_2, \ldots, x_k$ having degrees d_1, d_2, …, d_k respectively. Note that there could be multiple graphs $\boldsymbol{G}$ with degree vector $\boldsymbol{v}$. Let $\boldsymbol{v}'$ be the vector obtained from $\boldsymbol{v}$ by deleting d_1 and subtracting 1 from the next d_1 components of $\boldsymbol{v}$. Let $\boldsymbol{v}_1$ be the non-increasing vector obtained from $\boldsymbol{v}'$ by rearranging components if necessary. (For example, if $\boldsymbol{v}$ = [4 3 3 2 2 2 1 1] then $\boldsymbol{v}'$ = [**2 2 1 1** 2 1 1] and $\boldsymbol{v}_1$ = [2 2 2 1 1 1 1].) Show that $\boldsymbol{v}_1$ is also a graphical vector.

Answer:

Let S be the sum of the indices of the neighbors of x_1 (for instance, if x_1 is adjacent to x_3, x_4 and x_8, then $S = 15$). Take the graph $\boldsymbol{G}$ with degree vector v such that the **S is as small as possible**.

Now we claim that there do not exist indices $i < j$ such that x_1x_i is **not** an edge and x_1x_j is an edge in $\boldsymbol{G}$. Suppose the contrary. Since $d_i \geq d_j$, there must be some vertex x_t such that x_ix_t is an edge but x_jx_t is not an edge. Now in $\boldsymbol{G}$, delete edges x_1x_j and x_ix_t and replace them with edges x_1x_i and x_jx_t. Note that all degrees remain unchanged, but the sum of indices of neighbors of $\boldsymbol{v}_1$ has decreased by $(j-i)$, contradicting our assumption on $\boldsymbol{G}$. This proves our claim.

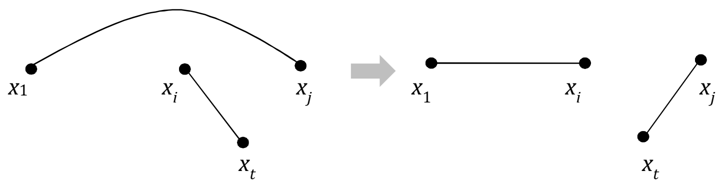

Figure 8.4. Illustration of a swap that decreases S

Our claim implies that x_1 is adjacent to the next d_1 vertices, namely x_2, x_3, …, x_{d_1+1}. Hence $\boldsymbol{v}_1$ is nothing but the graphical vector of the graph obtained from $\boldsymbol{G}$ by deleting x_1, since then the degrees of its neighbors all reduce by one. Hence $\boldsymbol{v}_1$ is graphical. ■

Remark 1: How did we come up with our rather strange extremal condition in the first paragraph? The problem provides a hint: it says we delete d_1 and subtract 1 from the next d_1 components to obtain $\boldsymbol{v}'$. Hence we wanted a graph such that x_1 is connected to the next d_1 vertices, since in this case simply removing vertex x_1 would have precisely this effect on the degree vector (this is our reasoning in the last paragraph of the proof). Now to prove such a graph exists, we needed a simple extremal property satisfied by such a graph. This naturally leads to our definition of S and the extremal condition that $\boldsymbol{G}$ minimizes S.

Remark 2: This is quite a useful lemma for testing whether a given degree sequence is graphical (see exercise 27).

Example 7 [MOP 2008]
Prove that if the edges of K_n, the complete graph on n vertices, are colored such that no color is assigned to more than $n-2$ edges, there exists a triangle in which each edge is a distinct color.

Answer
Assume to the contrary that there exists no such triangle. Define a *C-connected component* to be a set of vertices such that for any

two vertices in that set, there exists a path between them, all of whose edges are of color C. Now let $\boldsymbol{X}$ be the **largest** C-connected component of the graph for any color C, and say the color of $\boldsymbol{X}$ is red.

Suppose there is a vertex v not in $\boldsymbol{X}$. Consider two vertices u_1 and u_2 that are joined by a red edge. Neither edge vu_1 nor edge vu_2 can be red (otherwise v could be added to $\boldsymbol{X}$). So vu_1 and vu_2 are the same color (otherwise vu_1u_2 would have three distinct colors, contradicting our assumption). It follows that v is joined to all elements of $\boldsymbol{X}$ by the same color edge, say blue. But then $\boldsymbol{X} \cup \{v\}$ is a larger connected component (of color blue), contradiction.

It follows that there cannot be any vertex v outside $\boldsymbol{X}$, so all n vertices are in $\boldsymbol{X}$. Now, since $\boldsymbol{X}$ is red-connected and has n vertices, there must be at least $n-1$ red edges, contradiction. ■

Example 8 [Generalization of USAMO 2007-4]
Given a connected graph $\boldsymbol{G}$ with V vertices, each having degree at most D, show that $\boldsymbol{G}$ can be partitioned into two connected subgraphs, each containing at least $\frac{V-1}{D}$ vertices.

Answer:
We induct on E, the number of edges of $\boldsymbol{G}$. When $E = 1$, there are only two vertices and the partition consists of two isolated vertices. For the induction step, note that if we can delete an edge and $\boldsymbol{G}$ remains connected, we are done by induction. Hence we only consider the case when $\boldsymbol{G}$ is a tree.

Pick the root x such that the size of the **largest subtree T is minimized**. Clearly $|\boldsymbol{T}| \geq \frac{V-1}{D}$, since there are at most D subtrees and V–1 vertices amongst them. Also, $|\boldsymbol{T}| \leq \frac{V}{2}$. This is because if $|\boldsymbol{T}| > \frac{V}{2}$, then instead of rooting the tree at x we could root the tree at the first vertex y of $\boldsymbol{T}$. This would decrease $|\boldsymbol{T}|$ by 1, but T\{y} would still be the largest subtree since it would still have at least

$\frac{V-1}{2}$ vertices. This would contradict our assumption on x, as we would have a smaller largest subtree.

Thus $\frac{V}{2} > |\boldsymbol{T}| \geq \frac{V-1}{D}$. Take $\boldsymbol{T}$ to be one subgraph and $\boldsymbol{G} - \boldsymbol{T}$ to be the other. Both have size at least $\frac{V-1}{D}$ by the bounds on $\boldsymbol{T}$ and are connected, so we have found a valid partition. ■

Remark: The "induction" in the first paragraph is really just a formal way of saying "Well, the *worst* case for us is when $\boldsymbol{G}$ is a tree, so let's just forget about general graphs and prove the result for trees: if it's true for trees, it's true for everyone". Another way of reducing the focus to just trees is to take a spanning tree of $\boldsymbol{G}$.

Hall's Marriage theorem

Given N sets (not necessarily distinct), we say that the family of N sets has a ***system of distinct representatives*** (SDR) if it is possible to choose exactly one element from each set such that all the chosen elements are distinct. For example, if we have 4 sets {1, 2, 3}, {2, 4}, {2, 3, 4} and {1, 3} then this family has (1, 2, 4, 3) as a system of distinct representatives. Under what conditions does a family have a system of distinct representatives? One obvious necessary condition is that for any subfamily of k sets, the union of these k sets must have at least k elements. It turns out that this condition, known as the marriage condition, is also sufficient.

Example 9 [Hall's Marriage Theorem]
Show that the marriage condition is sufficient for the existence of an SDR.

Proof: Let the marriage condition hold for the family $A_1, A_2, \ldots, A_n$. Keep deleting elements from these sets until a family $F' = A_1', A_2', \ldots, A_n'$ is reached such that the deletion of any element will cause the marriage condition to be violated. We claim that at this stage each set contains exactly one element. This would imply the result, since these elements would be distinct by the marriage condition, and would hence form the required SDR.

Suppose our claim is false. Then some set contains at least 2 elements. WLOG this set is A_1 and let x and y be elements of A_1. Deleting x or y would violate the marriage condition by the definition of F'. Thus there exists subsets $\boldsymbol{P}$ and $\boldsymbol{Q}$ of $\{2, 3, 4, \ldots, n\}$ such that $\boldsymbol{X}=(A_1'-x) \cup (\bigcup_{i \in P} A_i')$ and $\boldsymbol{Y}=(A_1'-y) \cup (\bigcup_{i \in Q} A_i')$ satisfy $|\boldsymbol{X}| \le |\boldsymbol{P}|$ and $|\boldsymbol{Y}| \le |\boldsymbol{Q}|$. Adding gives

$$|\boldsymbol{X}| + |\boldsymbol{Y}| = |\boldsymbol{X} \cap \boldsymbol{Y}| + |\boldsymbol{X} \cup \boldsymbol{Y}| \le |\boldsymbol{P}| + |\boldsymbol{Q}|.$$

Now, $\boldsymbol{X} \cup \boldsymbol{Y} = \boldsymbol{A}_1' \cup (\bigcup_{i \in P \cup Q} \boldsymbol{A}_i')$ and $\boldsymbol{X} \cap \boldsymbol{Y} = \bigcup_{i \in P \cap Q} \boldsymbol{A}_i'$.

Thus the marriage condition implies that

$$|\boldsymbol{X} \cup \boldsymbol{Y}| \ge 1 + |\boldsymbol{P} \cup \boldsymbol{Q}|, \text{ and } |\boldsymbol{X} \cap \boldsymbol{Y}| \ge |\boldsymbol{P} \cap \boldsymbol{Q}|.$$

Adding gives

$$|\boldsymbol{X} \cap \boldsymbol{Y}| + |\boldsymbol{X} \cup \boldsymbol{Y}| \ge 1 + |\boldsymbol{P} \cup \boldsymbol{Q}| + |\boldsymbol{P} \cap \boldsymbol{Q}| = |\boldsymbol{P}| + |\boldsymbol{Q}| + 1,$$

contradicting our earlier bound. ■

Remark: The key idea in this proof was the fact that the marriage condition holds for the sets $\boldsymbol{A}_1', \boldsymbol{A}_2', \ldots \boldsymbol{A}_n'$ but **not** for $\boldsymbol{A}_1 \setminus \{x\}, \boldsymbol{A}_2', \ldots, \boldsymbol{A}_n'$ and $\boldsymbol{A}_1 \setminus \{y\}, \boldsymbol{A}_2', \ldots, \boldsymbol{A}_n'$. This proof illustrates an important idea: it's useful to exploit conditions given to us, but it's even more useful to exploit situations when the conditions **don't** hold.

Hall's marriage theorem was phrased above in the language of set theory, but we can also interpret it in graph theoretical terms. Consider a bipartite graph $\boldsymbol{G}$ with vertex set $\boldsymbol{V} = \boldsymbol{V}_1 \cup \boldsymbol{V}_2$. A *complete matching* of the vertices of $\boldsymbol{V}_1$ is a subset of the edges of $\boldsymbol{G}$

such that:

(i) Every vertex of $\boldsymbol{V}_1$ is incident on exactly one edge
(ii) Each vertex of $\boldsymbol{V}_2$ is incident on at most one edge

In other words, it is a pairing such that every vertex in $\boldsymbol{V}_1$ is paired with a vertex in $\boldsymbol{V}_2$ and no two vertices in $\boldsymbol{V}_1$ are paired with the same vertex of $\boldsymbol{V}_2$. The vertices in a pair are joined by an edge.

If the vertices of $\boldsymbol{V}_1$ represent the sets $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, …, $\boldsymbol{A}_n$ and the vertices of $\boldsymbol{V}_2$ represent elements in $\bigcup_{i=1}^{n} \boldsymbol{A}_i$, then a complete matching of $\boldsymbol{V}_1$ gives us a system of distinct representatives (namely the vertices of $\boldsymbol{V}_2$ to which the vertices of $\boldsymbol{V}_1$ are matched).

Example 10 [Canada 2006-3]

In a rectangular array of nonnegative reals with m rows and n columns, each row and each column contains at least one positive element. Moreover, if a row and a column intersect in a positive element, then the sums of their elements are the same. Prove that $m = n$.

Answer:

Create a bipartite graph, with the left side representing rows and the right side for columns. Place an edge between two vertices if and only if the corresponding row and column intersect in a positive element. The idea is to show that there is a matching from rows to columns, so $n \geq m$. By symmetry the same argument will give $m \geq n$, implying $m = n$.

Assume to the contrary there is no such matching. Then the marriage condition must be violated, so there exists some set $\boldsymbol{S}$ of rows having a total set $\boldsymbol{T}$ with $|\boldsymbol{T}| < |\boldsymbol{S}|$ of columns in which positive entries appear. Let the sums of the $|\boldsymbol{S}|$ rows be s_1, …, s_k. By the property, each of the $|\boldsymbol{T}|$ columns has sum equal to one of the s_i. So the total sum of the elements in the $\boldsymbol{S}$ rows, when calculated from the column point of view (since entries outside are all nonnegative) is at most a sum of a subset of the s_i. Yet from the row point of view, it is the full sum. As all $s_i > 0$, this is a contradiction. ■

Remark: This duality between arrays of numbers (i.e. matrices) and graphs (especially bipartite graphs) comes up very often. Keep an eye out for this trick, since it can prove very useful. In fact, analyzing and algebraically manipulating these matrices allows graph theory to be studied from an algebraic viewpoint, and fast algorithms for multiplying matrices form the basis of a class of algorithms for large graphs known as algebraic graph algorithms.

Unexpected Applications of Graph Theory

Most problems in previous sections suggested a natural graph theoretic interpretation. In this section, we will leverage the power of graphs to model complex relationships in nonobvious ways. Carefully constructed graphs can reduce unfamiliar, complex problems to familiar graph theoretic ones.

Example 11 [Taiwan 2001]
Let $n \geq 3$ be an integer and let $\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_n$ be n distinct subsets of $\boldsymbol{S} = \{1, 2, \ldots, n\}$. Show that there exists an element $x \in \boldsymbol{S}$ such that the subsets $\boldsymbol{A}_1 \setminus \{x\}, \boldsymbol{A}_2 \setminus \{x\}, \ldots, \boldsymbol{A}_n \setminus \{x\}$ are also distinct.

Answer:
We construct a graph $\boldsymbol{G}$ with vertices $A_1, A_2, \ldots, A_n$. For each element y, if there exist distinct sets $\boldsymbol{A}_i$ and $\boldsymbol{A}_j$ such that $\boldsymbol{A}_i \setminus \{y\} = \boldsymbol{A}_j \setminus \{y\}$, we select **exactly one** such pair $(\boldsymbol{A}_i, \boldsymbol{A}_j)$ and join them by an edge (even if there are multiple such pairs, we select only one for each y). Suppose to the contrary there doesn't exist x as stated in the problem. Then all elements of $\boldsymbol{S}$ contribute at least one edge to the graph. Moreover, it is impossible for two different elements

to contribute the same edge since if $\boldsymbol{A}_i \setminus y_1 = \boldsymbol{A}_j \setminus y_1$ and $\boldsymbol{A}_i \setminus y_2 = \boldsymbol{A}_j \setminus y_2$ for distinct y_1 and y_2, this would force $\boldsymbol{A}_i = \boldsymbol{A}_j$.

Thus $\boldsymbol{G}$ has at least n edges, and hence has a cycle, WLOG $A_1A_2 \ldots A_kA_1$ for some $k \geq 3$. Then there exists some distinct $x_1, x_2, \ldots, x_k$ such that $\boldsymbol{A}_1 \setminus \{x_1\} = \boldsymbol{A}_2 \setminus \{x_1\}$; $\boldsymbol{A}_2 \setminus \{x_2\} = \boldsymbol{A}_3 \setminus \{x_2\}, \ldots, \boldsymbol{A}_k \setminus \{x_k\} = \boldsymbol{A}_1 \setminus x_k$. Now x_1 is in exactly one of $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ (otherwise $\boldsymbol{A}_1 = \boldsymbol{A}_2$). WLOG it is in $\boldsymbol{A}_2$ but not in $\boldsymbol{A}_1$. But then x_1 must also be in $\boldsymbol{A}_3$ since $\boldsymbol{A}_2 \setminus \{x_2\} = \boldsymbol{A}_3 \setminus \{x_2\}$, and similarly must be in $\boldsymbol{A}_4$, and so on. We finally get that $x_1 \in \boldsymbol{A}_1$, a contradiction. ■

The next problem, like several others in this book, underscores the usefulness of induction. Problems with around "2^n" objects practically beg you to induct: all you need to do is find an appropriate way to split the set of objects into two parts, and apply the induction hypothesis to the larger or smaller part as applicable. But how does this connect to graph theory?

Example 12 [USA TST 2002]

Let n be a positive integer and let $\boldsymbol{S}$ be a set of (2^n+1) elements. Let f be a function from the set of two-element subsets of $\boldsymbol{S}$ to $\{0, 1, \ldots, (2^{n-1}-1)\}$. Assume that for any elements (x, y, z) of $\boldsymbol{S}$, one of $f(\{x, y\})$, $f(\{y, z\})$ and $f(\{z, x\})$ is equal to the sum of the other two. Show that there exist a, b, c in $\boldsymbol{S}$ such that $f\{a, b\}$, $f\{b, c)$ and $f\{c, a\}$ are all equal to 0.

Answer

Step 1: The basic strategy

Our idea is to find a subset $\boldsymbol{S}'$ of $\boldsymbol{S}$ such that $|\boldsymbol{S}'| \geq 2^{n-1} + 1$ and for all x, y in $\boldsymbol{S}'$ **$f(\{x, y\})$ is even**. Then if we let $g(\{x, y\}) = \frac{f(\{x,y\})}{2}$ for all x, y in $\boldsymbol{S}'$, we would have a function from pairs in $\boldsymbol{S}'$ to $\{0, 1, \ldots, 2^{n-2}-1\}$ satisfying the same conditions as f and we could apply the induction hypothesis to get the result. It remains to show that such a set $\boldsymbol{S}'$ exists.

Step 2: Constructing the graph and a new goal

Construct a graph $\boldsymbol{G}$ with 2^n+1 vertices representing elements in $\boldsymbol{S}$ as follows: there is an edge between a and b if and only if $f(\{a, b\})$ is odd. We now need to find an independent set in $\boldsymbol{G}$ of size at least $2^{n-1}+1$. Our hope is that $\boldsymbol{G}$ is bipartite: then we can just take the larger side of the bipartition, which will have size at least $\lceil (2^n + 1)/2 \rceil = 2^{n-1}+1$. Some experimentation confirms our hope – but of course we still need a proof.

Step 3: A key observation

Note that for any 3 vertices a, b, c there must be 0 or 2 edges amongst them, since $f(\{a, b\}) + f(\{b, c\}) + f(\{c, a\})$ is even (since one of these terms is the sum of the other two).

Step 4: Proving G is bipartite

If $\boldsymbol{G}$ is not bipartite, it has an odd cycle, so consider its *smallest* odd cycle $v_1v_2...v_{2k+1}$. Consider vertices v_1, v_3, v_4. There must be an even number of edges amongst them. As v_3v_4 is an edge, v_1v_3 or v_1v_4 must be an edge. v_1v_3 is not an edge since otherwise amongst vertices v_1, v_2 and v_3 there would be three edges, contradicting our earlier observation. Hence v_1v_4 is an edge. But then $v_1v_4...v_{2k+1}$ is a smaller odd cycle, contradicting our assumption. ■

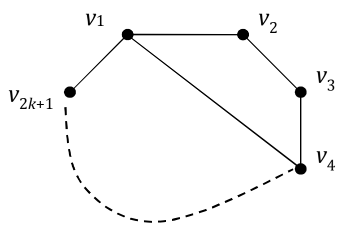

Figure 8.5. The edge v_1v_4 creates a smaller odd cycle

Remark 1: In step 4 we essentially proved the following result: If $\boldsymbol{G}$ is a graph in which for any three vertices there are either 0 or 2 edges between them, ***G*** is bipartite. This is a very handy lemma to keep in mind, especially since so many Olympiad problems boil down to proving that a certain graph is bipartite.

Remark 2: The idea of taking a shortest cycle arises very often.

Example 13 [IMO Shortlist 2002, C6]
Let n be an even positive integer. Show that there is a permutation $(x_1, x_2, \ldots, x_n)$ of $(1, 2, \ldots, n)$ such that for every $i \in (1, 2, \ldots, n)$, the number x_{i+1} is one of the numbers $2x_i$, $2x_i - 1$, $2x_i - n$, $2x_i - n - 1$. Here we use the cyclic subscript convention, so that x_{n+1} means x_1.

Answer:
Let $n = 2m$. We define a directed graph with vertices 1, 2, …, m and edges numbered 1, 2,…,$2m$ as follows. For each i ≤ m, vertex i has two outgoing edges numbered $2i-1$ and $2i$, and two incoming edges labeled i and $i+m$. All we need is an Eulerian circuit, because then successive edges will be of one of the forms $(i, 2i-1)$, $(i, 2i)$, $(i+m, 2i)$ or $(i+m, 2i-1)$. Then we can let $x_1, x_2, \ldots, x_n$ be the successive edges encountered in the Eulerian circuit and they will satisfy the problem's conditions.

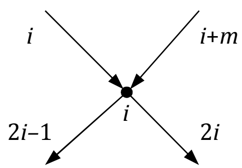

Figure 8.6.

Now, each vertex's indegree is equal to its outdegree, so we just need to show weak connectivity to establish that there is an Eulerian circuit. We do this by strong induction. There is a path from 1 to k: since there is a path from 1 to j where $2j = k$ or $2j-1 = k$, and an edge from j to k, there is a path from 1 to k. Thus ***G*** is weakly connected and hence has an Eulerian circuit. ■

Exercises

Hall's Theorem and Related Problems

1. **[Konig's marriage theorem]**
   Show that a *k-regular bipartite graph* $\boldsymbol{G}$ (all vertices have degree k) has a perfect matching (a matching covering all its vertices).

2. **[Konig's theorem]**
   A *matching* in a graph $\boldsymbol{G}$ is a set $\boldsymbol{M}$ of edges such that each vertex in $\boldsymbol{G}$ is incident to at most one edge in $\boldsymbol{M}$. A *vertex cover* in $\boldsymbol{G}$ is a set of vertices $\boldsymbol{C}$ such that each edge is incident to **at least** one vertex in $\boldsymbol{C}$. Using Hall's marriage theorem, show that in a bipartite graph $\boldsymbol{G}$, the maximum possible size of a matching is equal to the minimum possible size of a vertex cover.

3. **[Vietnam TST 2001]**
   A club has 42 members. Suppose that for any 31 members in this club, there exists a boy and a girl among these 31 members who know each other. Show that we can form 12 disjoint pairs of people, each pair having one boy and one girl, such that the people in each pair know each other.

4. **[IMO Shortlist 2006, C6]**
   Consider an upward equilateral triangle of side length n, consisting of n^2 unit triangles (by upward we mean with vertex on top and base at the bottom). Suppose we cut out n upward unit triangles from this figure, creating n triangular holes. Call the resulting figure a *holey triangle*. A *diamond* is a 60-120 unit rhombus. Show that a holey triangle T can be tiled by diamonds, with no diamonds overlapping, covering a hole or sticking out of T if and only if the following holds: every

upward equilateral of side length k in T has at most k holes, for all $1 \le k \le n$.

5. **[Dilworth's theorem]**
A *directed acyclic graph* (DAG) is a directed graph with no directed cycles. An *antichain* in a DAG (with some abuse of notation) is a set of vertices such that no two vertices in this set have a directed path between them. Show that the size of the maximum antichain of the DAG is equal to the minimum number of disjoint paths into which the DAG can be decomposed.

6. **[Romanian TST 2005]**
Let $\boldsymbol{S}$ be a set of n^2+1 positive integers such that in any subset $\boldsymbol{X}$ of $\boldsymbol{S}$ with $n+1$ integers, there exist integers $x \neq y$ such that $x|y$. Show that there exists a subset $\boldsymbol{S}'$ of $\boldsymbol{S}$ with $\boldsymbol{S}' = \{x_1, x_2, \ldots, x_{n+1}\}$ such that $x_i|x_{i+1}$ for each $1 \le i \le n$.

Coloring Problems

7. **[Welsh-Powell theorem] (U*)**
A *proper coloring* of the vertices of a graph $\boldsymbol{G}$ is an assignment of one color to each vertex of a graph $\boldsymbol{G}$ such that no two adjacent vertices in $\boldsymbol{G}$ have the same color. Let $\boldsymbol{G}$ be a graph with vertices having degrees $d_1 \ge d_2 \ge \ldots \ge d_n$. Show that there exists a proper coloring of $\boldsymbol{G}$ using at most $\max_i$ (min$\{i, d_i+1\}$ colors.

8. **[Dominating sets and coloring] (U*)**
If a graph on n vertices has no dominating set of size less than k, then show that its vertices can be properly colored in $n-k$ colors.

9. **[IMO 1992, Problem 3]**
Let $\boldsymbol{G}$ be the complete graph on 9 vertices. Each edge is either colored blue or red or left uncolored. Find the smallest value

of n such that if n edges are colored, there necessarily exists a monochromatic triangle.

10. [IMO Shortlist 1990]
The edges of a K_{10} are colored red and blue. Show that there exist two disjoint monochromatic odd cycles, both of the same color.

11. [Szekeres-Wilf theorem] (U*)
Show that any graph $\boldsymbol{G}$ can be properly colored using at most 1+max $\Delta(\boldsymbol{G}')$ colors, where the maximum is taken over all induced subgraphs $\boldsymbol{G}'$ of $\boldsymbol{G}$ and $\Delta(\boldsymbol{G}')$ refers to the maximum degree of a vertex in the induced subgraph $\boldsymbol{G}'$.

12. [Generalization of USA TST 2001]
Let $\boldsymbol{G}$ be a directed graph on n vertices, such that no vertex has out-degree greater than k. Show that the vertices of $\boldsymbol{G}$ can be colored in $2k+1$ colors such that no two vertices of the same color have a directed edge between them.

13. [Brook's theorem] (U*)
We know from (i) that $\Delta+1$ colors are sufficient to properly color the vertices of a graph $\boldsymbol{G}$, where Δ is the maximum degree of any vertex in $\boldsymbol{G}$. Show that if $\boldsymbol{G}$ is connected and is neither a complete graph nor an odd cycle, then actually Δ colors suffice.
[Easy version: prove the result above with the added condition that not all vertices have the same degree.]

Turan's theorem and applications

14. [Turan's theorem] (U*)
The *Turan graph* $\boldsymbol{T}(n, r)$ is the graph on n vertices formed as follows: partition the set of n vertices into r equal or almost equal (differing by 1) parts, and join two vertices by an edge if and only if they are in different parts. Note that $\boldsymbol{T}(n, r)$ has no $(r+1)$-clique. Show that amongst all graphs having no $(r+1)$-

clique, the Turan graph has the most edges. Hence, deduce that the maximum number of edges in a K_{r+1}-free graph is $\frac{(r-1)n^2}{2r}$. This generalizes the bound of $n^2/4$ edges in triangle-free graphs that we proved in chapter 6.

[Hint: first prove the following claim: there do not exist three vertices u, v, w such that uv is an edge in $\boldsymbol{G}$ but uw and vw are not. Show this by assuming the contrary and then making some adjustments to $\boldsymbol{G}$ to obtain a graph with more edges, contradicting the fact that $\boldsymbol{G}$ has the maximum possible number of edges amongst all K_{r+1}-free graphs. The claim establishes that if two vertices u and v have a common non-neighbor, then u and v themselves are non-neighbors. This shows that $\boldsymbol{G}$ is k-partite for some k. Now show that the maximum number of edges will occur when $k = r$ and the parts are equal or differ by 1.]

Remark: This method of proving that $\boldsymbol{G}$ is multipartite is extremely important, and is sometimes called Zykov symmetrization.

15. [Poland 1997]

There are n points on a unit circle. Show that at most $n^2/3$ pairs of these points are at distance greater than $\sqrt{2}$.

16. [IMO Shortlist 1989]

155 birds sit on the circumference of a circle. It is possible for there to be more than one bird at the same point. Birds at points P and Q are mutually visible if and only if angle POQ ≤ 10°, where O is the center of the circle. Determine the minimum possible number of pairs of mutually visible birds.

17. [USA TST 2008]

Given two points (x_1, y_1) and (x_2, y_2) in the coordinate plane, their *Manhattan distance* is defined as $|x_1-x_2| + |y_1-y_2|$. Call a pair of points (A, B) in the plane *harmonic* if $1 < d(A, B) \leq 2$.

Given 100 points in the plane, determine the maximum number of harmonic pairs among them.

More Extremal Graph problems

18. [China TST 2012]

Let n and k be positive integers such that $n > 2$ and $n/2 < k < n$. Let $\boldsymbol{G}$ be a graph on n vertices such that $\boldsymbol{G}$ contains no $(k+1)$-clique but the addition of any new edge to $\boldsymbol{G}$ would create a $(k+1)$-clique. Call a vertex in $\boldsymbol{G}$ central if it is connected to all $(n-1)$ other vertices. Determine the least possible number of central vertices in $\boldsymbol{G}$.

19. [China TST 2011]

Let $\boldsymbol{G}$ be a graph on $3n^2$ vertices ($n > 1$), with no vertex having degree greater than $4n$. Suppose further that there exists a vertex of degree one and that for any two points, there exists a path of length at most 3 between them. Show that G has at least $(7n^2-3n)/2$ edges.

20. [Generalization of IMO Shortlist 2013, C6]

In a graph $\boldsymbol{G}$, for any vertex v, there are at most $2k$ vertices at distance 3 from it. Show that for any vertex u, there are at most $k(k+1)$ vertices at distance 4 from it.

21. [IMO Shortlist 2004, C8]

For a finite graph $\boldsymbol{G}$, let $f(\boldsymbol{G})$ denote the number of triangles in $\boldsymbol{G}$ and $g(\boldsymbol{G})$ the number of tetrahedra (K_4s). Determine the smallest constant c such that $g(\boldsymbol{G})^3 \leq c\, f(\boldsymbol{G})^4$ for all graphs $\boldsymbol{G}$.

22. [IMO Shortlist 2002, C7]

In a group of 120 people, some pairs are friends. A weak quartet is a group of 4 people containing exactly one pair of friends. What is the maximum possible number of weak quartets?

Tournaments

23. Show that if a tournament has a directed cycle, then it has a directed triangle.

24. [Landau's theorem]
Call a vertex v in a tournament $\boldsymbol{T}$ a champion if for every vertex u in $\boldsymbol{T}$, there is a directed path from v to u in $\boldsymbol{T}$ of length at most 2. Show that every tournament has a champion.

25. [Moon-Moser theorem]
A directed graph $\boldsymbol{G}$ is called *strongly connected* if there is a directed path from each vertex in $\boldsymbol{G}$ to every other vertex in $\boldsymbol{G}$. A directed graph with n vertices is called *vertex-pancyclic* if every vertex is contained in a cycle of length p, for each $3 \le p \le n$. Show that a strongly connected tournament is vertex pancyclic.

26. [Based on USA TST 2009]
Let $n>m>1$ be integers and let $\boldsymbol{G}$ be a tournament on n vertices with no $(m+1)$-cycles. Show that the vertices can be labeled 1, 2, …, N such that if $a \ge b+m-1$, there is a directed edge from b to a.

Miscellaneous

27. [Generalization of Saint Petersburg 2001]
For each positive integer n, show that there exists a graph on $4n$ vertices with exactly two vertices having degree d, for each $1 \le d \le 2n$.
Remark: The degree vector lemma kills this otherwise difficult problem.

28. [Iran 2001]
In an $n \times n$ matrix, a *generalized diagonal* refers to a set of n

entries with one in each row and one in each column. Let $\boldsymbol{M}$ be an $n \times n$ 0-1 matrix, and suppose $\boldsymbol{M}$ has exactly one generalized diagonal containing all 1s. Show that it is possible to permute the rows and columns of $\boldsymbol{M}$ to obtain a matrix $\boldsymbol{M}'$ such that (i, j) entry in $\boldsymbol{M}'$ is 0 for all $1 \leq j < i \leq n$.

29. [Generalization of Russia 2001]

Let $\boldsymbol{G}$ be a tree with exactly $2n$ leaves (vertices with degree 1). Show that we can add n edges to $\boldsymbol{G}$ such that $\boldsymbol{G}$ becomes 2-connected, that is, the destruction of any edge at this point would still leave $\boldsymbol{G}$ connected.

30. [Russia 1997]

Let m and n be odd integers. An $m \times n$ board is tiled with dominoes such that exactly one square is left uncovered. One is allowed to slide vertical dominoes vertically and horizontal dominoes horizontally so as to occupy the empty square (thereby changing the position of the empty square). Suppose the empty square is initially at the bottom left corner of the board. Show that by a sequence of moves we can move the empty square to any of the other corners.

31. [Generalization of Japan 1997]

Let $\boldsymbol{G}$ be a graph on n vertices, where $n \geq 9$. Suppose that for any 5 vertices in $\boldsymbol{G}$, there exist at least two edges with endpoints amongst these 5 vertices. Show that $\boldsymbol{G}$ has at least $n(n-1)/8$ edges. Determine all n for which equality can occur.

32. [Japan 1997]

Let n be a positive integer. Each vertex of a 2^n-gon is labeled 0 or 1. There are 2^n sequences obtained by starting at some vertex and reading the first n labels encountered clockwise. Show that there exists a labelling such that these 2^n sequences are all distinct.

33. [Based on USA TST 2011]

In an undirected graph $\boldsymbol{G}$, all edges have weight either 1 or 2. For each vertex, the sum of the weights of edges incident to it is odd. Show that it is possible to orient the edges of $\boldsymbol{G}$ such that for each vertex, the absolute value of the difference between its in-weight and out-weight is 1, where in-weight refers to the sum of weights of incoming edges and out-weight refers to the sum of weights of outgoing edges.

34. [IMO Shortlist 1990]

Consider the rectangle in the coordinate plane with vertices (0, 0), (0, m), (n, 0) and (m, n), where m and n are odd positive integers. This rectangle is partitioned into triangles such that each triangle in the partition has at least one side parallel to one of the coordinate axes, and the altitude on any such side has length 1. Furthermore, any side that is not parallel to a coordinate axis is common to two triangles in the partition. Show that there exist two triangles in the partition each having one side parallel to the x axis and one side parallel to the y axis.