

Lecture 6: Graph Properties

Rajat Mittal

IIT Kanpur

In this lecture, we will take a look at various graph properties and constructs around them. Initially we will discuss independent sets. The bulk of the discussion will be about properties like coloring, matching and planarity.

1 Independent sets

Suppose we are given a graph where vertices correspond to students in the class and edges correspond to friendships. What is the maximum number of students we can put in the exam room, if no two friends should be in the room?

If no two friends are allowed to sit in the exam room, there will be no edges between the people sitting in the exam room. This essentially defines the concept of an *Independent set*.

An independent set of a graph $G = (V, E)$ is a subset S of vertices V . In this subset, no two elements u, v of S have an edge between them, $(u, v) \notin E$. It is also called a *stable set*. The stability number $\alpha(G)$ of a graph G is the maximum possible size of an independent set in the graph.

Exercise 1. What is the stability number of the graph given below?

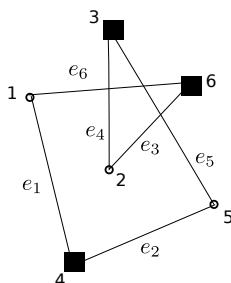


Fig. 1. Independent set in a graph. Is that the biggest possible?

We can also define an associated term, *maximal independent set*, as an independent set where adding any other vertex will make the subset not independent.

Exercise 2. Construct a connected graph where there is a maximal independent set whose cardinality is lesser than the stability number.

A *clique* S of a graph G , on the other hand, is the subset of vertices where there is an edge between every possible pair of vertices. In other words, G is a complete graph when restricted to the vertices of S .

We can define the *complement* of a graph G by taking edge (u, v) if and only if (u, v) is not an edge in G . The complement of a graph G is denoted by \bar{G} . So,

$$E_{\bar{G}} = \{(u, v) : (u, v) \notin E_G\}.$$

Exercise 3. What are the number of edges in \bar{G} if there are m edges and n vertices in G .

From the definitions, it is clear that an independent set in the graph is same as a clique in the complement graph.

Another connected notion is of a *vertex cover*. A vertex cover S is a subset of the vertex set, s.t., every edge has at least one vertex in common with S .

Exercise 4. Show that S is a vertex cover iff $V - S$ is an independent set.

So finding the best vertex cover (minimum) is same as finding the best independent set (maximum).

It is easy to find an approximate vertex cover. We can take a greedy approach. Pick an edge (include both the vertices of the edge in the vertex cover) and delete all the edges connected to it. In the remaining graph, again pick an edge and so on. This algorithm gives a vertex cover which is at most twice the size of the optimal vertex cover. You will prove that in the assignment.

We believe that finding the maximum independent set in a graph is a hard problem. If there is an efficient algorithm for it then there will be efficient algorithms for large class of very interesting problems.

But finding an approximation algorithm for independent set is really hard. Convince yourself that the if you take the approximation algorithm for vertex cover and take its complement. It need not be a good independent set.

2 Coloring

Consider the following problem. For an examination, we need to put students in classrooms such that no pair of friends sit in the same classroom. We are given the social graph of the class, with students as vertices and edges representing friendships. How many minimum classrooms do we need for the examination?

For any valid assignment, assume that there is a distinct color for every classroom. Then assign a vertex the color corresponding to its classroom. By this process, no two adjacent vertices will get the same color. This is called a coloring of the graph.

More formally, given a graph G , a *valid coloring* (or just coloring) is a map from a set of colors to set of vertices V , s.t., no two vertices of the same color are connected. The minimum number of colors needed to have a valid coloring of a graph G is known as its *chromatic number* $\chi(G)$.

The above question about examination classrooms can be reformulated as, what is the chromatic number of the social graph. There are many other applications of coloring,

- Color the map of countries so that no two countries with shared border have the same color.
- Schedule the examinations so that no student has two exams in the same slot.

Exercise 5. Formulate all the above questions as graph coloring problems.

Exercise 6. When can you color a cycle with just two colors?

For a graph coloring, the first thing to notice is that every color class (set of vertices with that color) form an independent set. So a coloring is equivalent to partitioning the vertex set into disjoint independent sets.

Theorem 1. Given a graph G with n vertices. Then $\alpha(G)\chi(G) \geq n$.

Proof. Let's say that the optimal coloring divides the vertex set V into color classes V_1, V_2, \dots, V_k ($\chi(G) = k$). That means two vertices in the same V_i have same color and two vertices in different V_i 's have different colors. Then,

$$n = \sum_{i=1}^k |V_i|.$$

But every V_i is an independent set and hence $|V_i| \leq \alpha(G)$ for all i . So,

$$n \leq k\alpha(G) \leq \chi(G)\alpha(G).$$

Hence proved. □

Exercise 7. Construct a connected graph with at least 6 vertices, s.t., $\chi(G)\alpha(G) = n$.

An upper bound on chromatic number can be given by the degree.

Theorem 2. *If G has maximum degree k , then $\chi(G) \leq k + 1$. Moreover, if G is also connected and there is at least one vertex with degree strictly less than k then $\chi(G) \leq k$.*

Proof. We will first prove that $k + 1$ colors suffice to color a graph with degree k . Let's apply induction on number of vertices in the graph. For the base case, clearly if the number of vertices are less than or equal to $k + 1$ then graph can be colored with $k + 1$ colors.

For the general case, consider any particular vertex v . Say G' be the graph obtained by deleting v and all edges which include v . G' has maximum degree k and has less number of vertices. So G' can be colored with $k + 1$ colors.

Now consider the neighbors of v . There are k of them, pick the $(k + 1)$ -th color for v . Hence G can be colored with $k + 1$ colors.

For the Moreover part, we will again use induction. There exists a vertex with degree less than k , say v . Remove v and all edges which include v . Suppose we get connected components H_1, H_2, \dots, H_k . All H_i 's have to be connected to v , since G is connected.

Exercise 8. At least one vertex in every H_i has degree less than k .

Applying induction hypothesis on all H_i 's we get a coloring for all of them. Since they are disconnected, the entire coloring is consistent. Now there is at least one color remaining for v as there are at most $k - 1$ neighbors. Hence proved. □

Exercise 9. Construct a graph with maximum degree d which is not colorable using d colors.

3 Matching

Suppose you are the organizer of a friendly tennis tournament between India and Pakistan. You have been provided with a graph with vertices as the participants. There are edges between the participants if there is a match possible between them (the players have the same level). How many maximum matches can you hold simultaneously?

This is a matching problem. You want to match participants in pairs so that all pairs are disjoint and number of pairs are maximized. For the following discussion on matching, we will assume that the graph is bipartite. The two set of vertices are denoted by X and Y , s.t., $|X| \leq |Y|$.

A *matching* M in a bipartite graph $G = (X \cup Y, E)$ is a subset of edges, s.t., no two edges share a vertex. A *maximum matching* is a matching M , s.t., for any matching N in G , $|M| \geq |N|$.

Exercise 10. Is the matching shown by bold edges maximum? (Fig. ??)

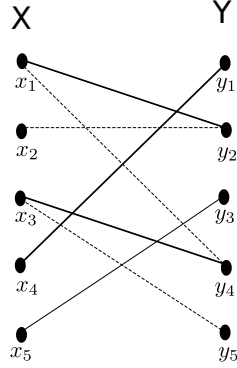


Fig. 2. An alternating path. Thick edges are the edges of existing matching.

Suppose there is a path which starts from a non-matched vertex, alternates between a matched and a non-matched edge and finishes at a non matched vertex. The path $x_2, y_2, x_1, y_4, x_3, y_5$ is such an example. By putting $(x_2, y_1), (x_1, y_4), (x_3, y_5)$ in the matching and removing $(x_1, y_2), (x_3, y_4)$ we can increase the size of the matching.

An *alternating path* in a graph with a given matching M is a path,

- The first and the last vertex are not matched.
- The edges alternate between unmatched and matched edges.

You should convince yourself that if such a path exist then we can increase the size of matching. Even the converse is true.

Theorem 3. A matching M is a maximum matching iff there are no alternating paths.

Proof. As noticed, if there is an alternating path then the matching is not maximum. We will now prove that if the matching is not maximum then there is an alternating path.

Suppose M' is the maximum matching. Consider the graph G' which has only edges from M and M' (if they share an edge, we will keep two edges from them). Clearly every vertex has degree at most 2. The connected components in G' will look like isolated vertex, cycle or paths.

Exercise 11. Prove that the only connected components where number of edges from M and M' differ are,

- An alternating path for M . The number of edges from M are less than the number of edges from M' .
- An alternating path for M' . The number of edges from M' are less than the number of edges from M .

Since the number of edges in M' are more than the number of edges in M . At least one alternating path for M will exist. Actually the number of alternating path for M will be greater than the number of alternating path for M' , but it is sufficient for us to have one alternating path for M . \square

Since $|X| \leq |Y|$, if all vertices of X are included in M then it is called a *complete matching*. For the above example, if there is a complete matching, then all participants (of Pakistan, their population is less) can play at the same time with someone of their own level.

Though it is not necessary that a graph has a complete matching. You might notice that if there is an isolated vertex (degree zero vertex) then there is no matching. The guess would be if all vertices have high degree then there should be a complete matching.

Exercise 12. Find a bipartite graph such that all degrees are higher than 3, which does not have a complete matching. Show that 3 can be replaced by any constant.

Given a subset S of vertices X , define $N(S)$ to be the set of vertices in Y adjacent to S .

$$N(S) = \{u \in Y : \exists v \in X : (u, v) \in E\}.$$

The set $N(S)$ is called the neighborhood of S . Suppose the cardinality of $N(S)$ is smaller than $|S|$ for a graph G . Then we won't be able to match every vertex in S . In other words, if there exists a subset S , s.t., $|N(S)| < |S|$ then there is no complete matching. The condition that $|N(S)| \geq |S|$ is called *Hall's condition*.

We would like to show that the converse is also true. That means, if Hall's condition is true for every subset then there is a complete matching. The idea would be to grow the matching if it is not complete. We will finally show that if Hall's conditions are satisfied then we can always grow the matching.

To show that Hall's condition is sufficient for a perfect matching. We will show that if matching M is not complete and Hall's condition is satisfied, then alternating path for M exists.

Theorem 4. *Hall's theorem: Suppose there is a bipartite graph $G = (X \cup Y, E)$. There is a complete matching in the graph G iff for all $S \subseteq X$, $|N(S)| \geq |S|$.*

Proof. Again, if there is a complete matching then the size of the neighborhood $N(S)$ for any $S \subseteq X$ has to be larger than or equal to the size of S .

For the opposite direction, suppose M is a maximum matching which is not complete. We will show the existence of an alternating path.

Consider any vertex x_0 in X which is not matched. $N(\{x_0\})$ will have at least one element, say y_1 . If y_1 is not matched we are done else it is matched to x_2 . For the set $\{x_0, x_2\}$, again we can find a y_3 which is connected to either x_0 or x_1 (not equal to y_1). Continuing this way, we should reach a vertex y_r which is not matched.

This can be traced back. y_r is connected to some x_i , x_i to matching y_{i-1} and so on. This gives an alternating path.

$$y_r, x_i, y_{i-1}, x_j, y_{j-1}, \dots, x_0.$$

But there can't be an alternating path for a maximum matching by Thm. ???. This gives a contradiction. □

Exercise 13. Do we need the converse part of Thm. ??? for Hall's theorem?

4 Planarity ¹

A *planar representation* of a graph G is a drawing of the graph on the piece of paper such that no two edges intersect each other except at the end points. Take a look at two representations of the graph K_4 .

A graph G is called *planar* if it has a planar representation. Notice that a planar graph can also have non-planar representations. To show that a graph is planar, we can just show a planar representation. But showing that a graph is non-planar takes a lot of effort. Let's take an example.

Theorem 5. *The graph K_5 is non-planar.*

Proof. Any planar drawing of a graph divides the plane into regions. Look at the following example of K_4 .

In any planar drawing of the graph K_5 , pick 4 vertices and call them v_1, v_2, v_3, v_4 . Since $(v_1, v_2), (v_2, v_3), (v_3, v_4)$ and (v_4, v_1) are all connected. They will divide the plane into two regions. We can call them *inside* and *outside*.

Case 1: v_5 falls in the region *inside*. Then both the edges (v_1, v_3) and (v_2, v_4) have to fall in the region *outside*. But then they will have to cross each other.

Case 2: v_5 falls in the region *outside*. The proof in this case is similar and left as an exercise. □

¹ This section is taken from the book by Rosen [?].

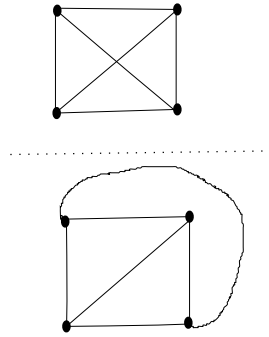


Fig. 3. Two different representations of K_4 . Second one is planar.

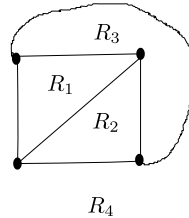


Fig. 4. Regions in a graph.

Exercise 14. Show that $K_{3,3}$ is non-planar.

If a graph has $K_{3,3}$ or K_5 as subgraph, then it is non-planar. Surprisingly, it can be shown that a graph is non-planar iff it has a subgraph which is *derived* from $K_{3,3}$ or K_5 . We will not cover the meaning of the term *derived* and the proof. Interested readers can look at *Kuratowski's Theorem*.

Looking at the proof of Thm. ??, we were interested in different regions of the graph. Euler showed that for a graph G , any planar representation have same number of regions and this number is related to vertices and edges.

Theorem 6. *Euler's Formula: Suppose we are given a connected graph G , then number of regions r in a planar representation is $m - n + 2$. Here m is the number of edges and n is the number of vertices.*

Proof. Let's keep a particular planar representation of G in mind. We are going to construct this representation by adding one edge at a time. We will start with any single edge, call this basic graph G_1 .

Given G_n , look at a new edge e ($e \notin E_{G_n}$) which has at least one vertex in G_n . e exists at every step because graph is connected. If both the endpoints of e are already in G_n , to obtain G_{n+1} , we just need to draw the edge. Otherwise, to obtain G_{n+1} , draw the edge and the additional vertex too.

Suppose r_n, e_n, v_n be the regions, edges and vertices respectively in the graph G_n . We will show by induction that $r_n = e_n - v_n + 2$.

Exercise 15. Show the base case.

As mentioned above, there can be two case when adding a new edge e to G_n .

Case 1: Both vertices of e are already present in G_n . Then they should be in the same region (otherwise there will be a crossing). By connecting those two vertices, we have divided the region into two regions. So $r_{n+1} = r_n + 1, e_{n+1} = e_n + 1, v_{n+1} = v_n$ and Euler's formula holds.

Case 2: If only one vertex of e is present. In this case the new edge does not make a new region. So $r_{n+1} = r_n, e_{n+1} = e_n + 1, v_{n+1} = v_n + 1$ and again Euler's formula holds.

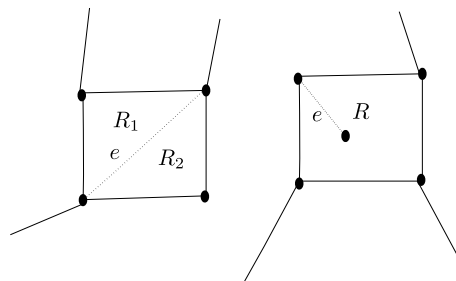


Fig. 5. Two different cases of adding an edge.

□

A very interesting fact is known about any planar graph. It can always be colored by 4 colors. This is known as 4-color theorem and the proof of it required lot of effort. If you are interested please read more about 4-color theorem.

5 Assignment

Exercise 16. Find all graphs on 4 vertices which are isomorphic to their complement.

Exercise 17. Prove that the vertex cover algorithm given in the notes gives a result which is at most twice the size of the best vertex cover.

Exercise 18. Define graph product of G_1 and G_2 as the graph $G_1 \otimes G_2$ with vertex set $V_1 \times V_2$. For the edges, $((u, i), (v, j))$ is an edge in the product graph if,

$$(u = v \text{ and } (i, j) \in E_{G_2}) \text{ or } ((u, v) \in E_{G_1} \text{ and } i = j).$$

Show that G can be colored by t colors iff $\alpha(G \otimes K_t) = |V_G|$.

Exercise 19. What is the chromatic number of a bipartite graph with m vertices in one part and n vertices in another?

Exercise 20. Thm. ?? gives a greedy way to color the vertices of a graph. Write the pseudocode of an algorithm which gives a d -coloring of a non-regular connected graph with maximum degree d .

Exercise 21. Show that a complete matching always exists in a bipartite regular graph $G = (X \cup Y, E)$.

Exercise 22. Prove that every tree is a planar graph. Hence, number of edges in a tree is $|V| - 1$.

Exercise 23. The degree of a region in a planar representation is the number of edges on the boundary of the region. For example, in the Fig. ??, R has degree 6, R_1, R_2 have degree 3. Show that for a connected planar graph,

$$2|E| = \sum_R \deg(R).$$

Where the summation runs over all regions of the planar representation. Hence prove that,

$$|E| \leq 3|V| - 6.$$

Note 1. It shows that, for a planar graph, the number of edges are linear in terms of the number of vertices.

References

1. K. H. Rosen. Discrete Mathematics and Its Applications. *McGraw-Hill*, 1999.
2. N. L. Biggs. Discrete Mathematics. *Oxford University Press*, 2003.