Graph Theory

Basics

Instructor: *Nicolò Cesa-Bianchi* version of July 8, 2020

The material in this handout is taken from: Reinhard Diestel, Graph Theory (5th edition), Springer, 2017.

Given a set S and any $k \in \{1, \ldots, |S|\}$, $[S]^k$ is the collection of all k-element subsets of S.

A graph $G = (V, E)$ has a finite **vertex set** V and a finite **edge set** $E \subseteq [V]^2$. We use i, j, u, v, w, x to denote vertices in V. The number $|V|$ of vertices is the **order** of G. A graph of order zero is empty, while graphs of order at most 1 are called trivial. An element of E is denoted by e or (i, j). If $(i, j) \in E$, then i, j denote the endpoints of the edge (the order does not matter). A vertex i is **incident** with an edge e if $e = (i, j)$. Two vertices i, j are **adjacent** if $(i, j) \in E$. If $E \equiv [V]^2$, then G is the complete graph (or **clique**) on n vertices, denoted by K_n. Note that G has no self-loops (i, i) because $(i, i) \notin [V]^2$. Moreover, there can be at most one edge in G between any two pair of vertices. Such graphs are often called *simple*.

If $G' = (V', E')$ is such that $V' \subseteq V$ and $E' \subseteq [V']^2 \cap E$, then G' is a **subgraph** of G. If a subgraph G' is such that $E' \equiv [V']^2 \cap E$, then G' is called the subgraph induced by V'.

Degrees. The **neighborhood** of a vertex v of G is the set $N(v)$ of vertices that are adjacent to v. The **degree** $d(v)$ of v is the cardinality of $N(v)$. A vertex with degree zero is **isolated**. The numbers $\delta(G) = \min\{d(v) : v \in V\}$ and $\Delta(G) = \max\{d(v) : v \in V\}$ are the minimum and maximum degree of G. If $\delta(G) = \Delta(G) = k$ then G is k**-regular**.

The **average degree** of G is

$$d(G) = \frac{1}{|V|} \sum_{v \in V} d(v)$$

and we obviously have $\delta(G) \le d(G) \le \Delta(G)$. A related quantity is the **edge density** $\varepsilon(G) = |E|/|V|$. Note that

$$|E| = \frac{1}{2} \sum_{i \in V} d(v) = \frac{1}{2} d(G)|V|$$

implying $\varepsilon(G) = d(G)/2$.

Fact 1 *The number of vertices of odd degree is always even in any graph.*

PROOF. Since $|E|$ is integer and $|E| = \frac{1}{2}\sum_{v \in V} d(v)$, then $\sum_{v \in V} d(v)$ must be even. Therefore, the number of vertices with odd degree must be even. □

We already know that the edge density is half the average degree. Now note that the minimum degree can be larger than the edge density. For instance, in K_2 we have $\delta(G) = 1$ and $\varepsilon(G) = \frac{1}{2}$.

Fact 2 *Every G with at least one edge has an induced subgraph H such that $\delta(H) > \varepsilon(H) \ge \varepsilon(G)$.*

PROOF. Construct a sequence of nested subgraphs $G \equiv G_0, G_1, \ldots$ induced by the vertex sets $V = V_0 \supseteq V_1 \supseteq V_2 \cdots$ as follows. If V_i has a vertex v_i of degree $d(v_i) \le \varepsilon(G_i)$ then $V_{i+1} \equiv V_i \setminus \{v_i\}$. Otherwise, stop and set $H = G_i$.If G_{i+1} is created, then

$$\varepsilon(G_{i+1}) = \frac{|E_{i+1}|}{|V_{i+1}|} = \frac{|E_i| - d(v_i)}{|V_i| - 1} \ge \frac{|E_i| - \varepsilon(G_i)}{|V_i| - 1} = \frac{|E_i|}{|V_i|} = \varepsilon(G_i)$$

The procedure stops before emptying the graph because $\varepsilon(K_1) = 0 < \varepsilon(G)$. When the procedure stops (say at $H \equiv G_k$ for some $k \ge 0$), it must be that $\delta(H) > \varepsilon(H) \ge \varepsilon(G)$, concluding the proof. □

Paths and cycles. A **path** in $G = (V, E)$ of length $k \ge 0$ is a subgraph P_k containing $k+1$ distinct vertices $v_0, \ldots, v_k \in V$ and k edges $e_1, \ldots, e_k \in E$ such that $e_i = (v_{i-1}, v_i)$ for $i = 1, \ldots, k$. If $k = 0$ then $P_0 = K_1$. A **cycle** C_k in G of length $k \ge 3$ is formed when a path P_{k-1} can be extended in G to include the edge $(v_{k-1}, v_0) \in E$. The length of a shortest cycle in G is the **girth** $g(G)$, while the length of a longest cycle in G is the **circumference**. A **chord** is any edge between two vertices of a cycle which is not itself an edge of the cycle.

If a graph has a large minimum degree, then it contains long paths and cycles.

Fact 3 *Every graph G with $\delta(G) \ge 2$ contains a path of length $\delta(G)$ and a cycle of length at least $\delta(G) + 1$.*

PROOF. Let $v_0, \ldots, v_k$ be the vertices on any longest path P_k in G. Then $N(v_k)$ all belong to P_k (otherwise P_k is not the longest path). Therefore, $k \ge d(v_k) \ge \delta(G)$. Now let v_i the vertex of P_k with smallest index i such that $(v_i, v_k) \in E$. Then the vertices $v_i, \ldots, v_k$ form a cycle of length at least $\delta(G) + 1$ because the degree of v_k is at least $\delta(G)$. □

The **distance** $d(i, j)$ betweeen two vertices i, j is the length of the shortest path between them (if no path exists between the two vertices, then the distance is infinite). The **diameter** $\operatorname{diam}(G)$ of G is the largest distance between any two vertices in G (note that the diameter can be infinite, for example when the graph has an isolated vertex). The **radius** $\operatorname{rad}(G)$ of a graph G is the smallest distance d such that there exists a vertex whose distance from any other vertex in G is at most d. Formally,

$$\operatorname{rad}(G) = \min_{i \in V} \max_{j \in V} d(i, j)$$

Clearly, $\operatorname{rad}(G) \le \operatorname{diam}(G)$. Also, let $x \in V$ such that $d(x, v) \le \operatorname{rad}(G)$ for all $v \in V$. Pick any two vertices $u, v \in V$ then $d(u, v) \le d(u, x) + d(x, v) \le 2 \operatorname{rad}(G)$. This shows that $\operatorname{diam}(G) \le 2 \operatorname{rad}(G)$.

Girth and diameter are related as follows.

Fact 4 *Every graph G containing at least a cycle satisfies $g(G) \le 2 \operatorname{diam}(G) + 1$.*

PROOF. Let C be the shortest cycle in G. If $g(G) \ge 2 \operatorname{diam}(G) + 2$, then C contains at least $2 \operatorname{diam}(G) + 2$ edges. Take any two vertices x, y at opposite extremes of C. Then x, y are connected by two paths in C, say P_1 and P_2, whose each length is at least $\operatorname{diam}(G) + 1$. On the other hand, the distance between x and y in G can be at most $\operatorname{diam}(G)$ by definition of diameter. Let P be a path joining x to y in G. Note that not all the edges of P can be in C (otherwise, C is not the shortest cycle in G). Then P togheter with the shortest between P_1 and P_2 forms a cycle shorter than C, and we have a contradiction. $\square$

The next result, which we state without proof, shows that the order of G is lower bounded by a function $n_0(d, g)$ of its average degree d and girth $g \geq 3$, where

$$n_0(d,g) = \begin{cases} 1 + d \displaystyle\sum_{i=0}^{(g-1)/2-1} (d-1)^i & \text{if } g \text{ is odd} \\ 2 \displaystyle\sum_{i=0}^{g/2-1} (d-1)^i & \text{if } g \text{ is even.} \end{cases}$$

Theorem 5 (Alon, Hoory, and Linial, 2002) *For any graph $G = (V, E)$, $|V| \geq n_0(d, g)$ for all $d = d(G) \geq 2$ and $g = g(G) \geq 3$.*

Using the well-known formula

$$\sum_{i=0}^{n-1} a^i = \frac{a^n - 1}{a - 1} \qquad a, n \in \mathbb{N} \quad a > 1,\, n \geq 1$$

we have that $n_0(d, g) = d^{\Theta(g)}$ for $d \geq 2$. Therefore, the above theorem says that $|V| = d^{\Omega(g)}$. Also, using again the same formula,

$$n_0(3,g) = 2 \sum_{i=0}^{g/2-1} 2^i = 2(2^{g/2} - 1) = 2^{g/2} + 2^{g/2} - 2 > 2^{g/2} + 4 - 2 > 2^{g/2} \qquad \text{(for } g \geq 3 \text{ even)}$$

and

$$\begin{aligned} n_0(3,g) &= 1 + 3 \sum_{i=0}^{(g-1)/2-1} 2^i = 1 + 3(2^{(g-1)/2} - 1) = \frac{3}{\sqrt{2}} 2^{g/2} - 2 \\ &> \left(\frac{3}{\sqrt{2}} - 1\right) 2^{3/2} - 2 + 2^{g/2} > 2^{g/2} \qquad \text{(for } g \geq 3 \text{ odd)} \end{aligned}$$

Therefore, $n_0(3, g) \geq 2^{g/2}$. So when $\delta(G) \geq 3$ we have $|V| \geq 2^{g(G)/2}$, which in wurn implies $g(G) < 2 \log_2 |V|$.

Connectivity. A non-empty graph G is **connected** if any two of its vertices are linked by a path in G. A maximal connected subgraph of G is a **component** of G. Any non-empty graph corresponds to a set containing at least one component. G is k**-connected** if $|V| > k$ and for all $X \subset V$ with $|X| < k$, the subgraph induced by $V \setminus X$ is connected. Every non-empty graph is 0-connected, and the 1-connected graphs are precisely the non-trivial connected graphs (K_1 is not 1-connected because it does not have order 2).

The largest integer k such that G is k-connected is the **connectivity** $\kappa(G)$ of G. Thus $\kappa(G) = 0$ if and only if G is disconnected or $G \equiv K_1$, and $\kappa(K_n) = n - 1$ for all $n \geq 1$.

We now relate connectivity to minimum degree and to the existence of a set of edges whose removal disconnects the graph.

Theorem 6 *If G is non-trivial, then $\kappa(G) \leq |F| \leq \delta(G)$ where F is any smallest set of edges whose removal causes the graph to disconnect.*

PROOF. Let $G = (V, E)$ be non-trivial and let v be any vertex with minimum degree $\delta(G)$. Then $|F| \leq \delta(G)$ because v can be disconnected by removing the edges that are incident with $N(v)$. We now show that $\kappa(G) \leq |F|$ by a case analysis. Let $G' = (V, E \setminus F)$.

Case 1. G has a vertex v that is not incident with an edge in F. Let C be the component of G' that contains v and consider the set V_C of vertices of C that are incident with an edge of F. If we remove these vertices, then v is disconnected from the other component of G. Hence $\kappa(G) \leq |V_C|$. On the other hand, no edge in F can have both ends in C (otherwise, F is not minimal). Therefore, $|V_C| \leq |F|$.

Case 2. All vertices of G are incident with some edge in F. Pick an arbitrary vertex v and let C be the component of G' that contains v. Some $u \in N(v)$ are such that $(v, u) \in F$. The others nodes in the neighborhood of v must belong to C and are incident with distinct edges of F (otherwise, F is not minimal). Therefore $d(v) \leq |F|$, which implies $d(v) = |F| = \delta(G)$, because we already know that $|F| \leq \delta(G)$. As removing $N(v)$ disconnects v, we conclude $\kappa(G) \leq \delta(G) = |F|$. □

Trees and forests. An acyclic graph, one not containing any cycles, is called a forest. A connected forest is called a tree. (Thus, a forest is a graph whose components are trees.) The vertices of degree 1 in a tree are its leaves, the others are its inner vertices. Every non-trivial tree has a leaf—consider, for example, the ends of a longest path. The next theorem (whose proof is left as exercise) characterizes which graphs are trees.

Theorem 7 *The following assertions are equivalent for a graph T:*

1. *T is a tree;*
2. *Any two distict vertices x, y of T are linked by a unique path in T (denoted by xTy);*
3. *T is minimally connected, i.e., removing any edge disconnects T;*
4. *T is maximally acyclic, i.e., adding an edge between any two non-adjacent vertices create a cycle.*

A **spanning tree** of a connected graph $G = (V, E)$ is a subgraph $T = (V, E')$ that is a tree (note that T and G share the same vertex set). A common application of the theorem above is to prove that every connected graph contains a spanning tree: take a minimal connected spanning subgraph and use 3, or take a maximal acyclic subgraph and apply 4. Spanning trees can be found in time $\mathcal{O}(|E|)$ via breadth-first or depth-first search of G.

Corollary 8 *A connected graph with n vertices is a tree if and only if it has $n - 1$ edges.*

PROOF. First, we prove that a tree has $n - 1$ edges. Note that there always exists a permutation $v_1, \ldots, v_n$ of the vertices of a tree so that every v_i with $i \geq 2$ has a unique neighbour in $\{v_1, \ldots, v_{i-1}\}$. Induction on $i = 2, \ldots, n$ shows that the subgraph spanned by the first i vertices has $i - 1$ edges. For $i = n$ this proves the claim.
Conversely, let G be any connected graph with n vertices and $n - 1$ edges. Let G' be a spanning

tree in G. Since G' has $n-1$ edges by the first implication, it follows that $G = G'$. □

A tree $T = (V, E)$ with a fixed root r is a **rooted tree**. Writing $x \leq y$ for $x \in rTy$ then defines a partial order on V, the **tree-order** associated with T and r.

Bipartite graphs. Let $r \geq 2$ be an integer. A graph $G = (V, E)$ is called r-partite if V admits a partition into r elements such that every edge has its ends in different elements: vertices in the same partition elements must not be adjacent. Instead of 2-partite one usually says bipartite. An r-partite graph in which every two vertices from different partition elements are adjacent is called complete. Note that a bipartite graph is not necessarily connected.

Clearly, a bipartite graph cannot contain an odd cycle, a cycle of odd length. We now prove that also the converse is true.

Fact 9 *If a graph does not contain an odd cycle, then it is bipartite.*

PROOF. Let $G = (V, E)$ be a graph without odd cycles. If G is bipartite, then all its components are bipartite or trivial. So we may assume that G is connected. Let $T = (V, E_T)$ be a spanning tree in G, pick a root r and denote the associated tree-order on V by $\leq_T$. For each $v \in V$, the unique path rTv has odd or even length. This defines a bipartition of V. We show that G is bipartite with this partition. Let $e = (x, y) \in E$. If $e \in E_T$, say with $x \leq_T y$, then $rTy = rTxy$ and so x and y lie in different partition elements. If $e \notin E_T$, then $C_e = (xTy, e)$ is a cycle (because of 4 in the theorem above), and the vertices along xTy alternate between the two partition elements. Since C_e is even by assumption, x and y again lie in different elements. □

Euler tours. A walk (resp., a closed walk) is a path (resp., cycle) whose vertices may not be all distinct. A closed walk in a graph is an Euler tour if it traverses every edge of the graph exactly once. A graph is **Eulerian** if it admits an Euler tour.

Theorem 10 (Euler, 1736) *A connected graph is Eulerian if and only if every vertex has even degree.*

PROOF. The degree condition is clearly necessary: a vertex appearing k times in an Euler tour (or $k+1$ times, if it is the starting and finishing vertex and as such counted twice) must have degree $2k$. Conversely, we show by induction on $|E|$ that every connected graph $G = (V, E)$ with all degrees even has an Euler tour. The induction starts trivially with $|E| = 0$. Now let $|E| \geq 1$. Since all degrees are even, we can find in G a non-trivial closed walk that contains no edge more than once (see Exercise 4). Let W be such a walk of maximal length, and write F for the set of its edges. If $F = E$, then W is an Euler tour and we are done. Suppose, therefore, that $E' = E \setminus F$ has at least an edge. For every vertex $v \in V$, an even number of the edges $\{(v, u) : u \in N(v)\}$ lies in F (because W is a closed walk containing each edge only once), so the degrees of the subgraph $G' = (V, E')$ are again all even. Since G is connected, G' has an edge e incident with a vertex on W. By the induction hypothesis, the component C of G' containing e has an Euler tour. Concatenating this with W, we obtain a closed walk in G that contradicts the maximal length of W. □

Euler tours can be found in time $\mathcal{O}(|E|)$ using Hierholzer's algorithm.

Hamilton cycles. A Hamilton cycle is a cycle that contains all vertices. A graph is Hamiltonian if it contains a Hamilton cycle.

Unlike Euler tours, only sufficient conditions are known for the existence of Hamilton cycles.

Theorem 11 (Dirac 1952) *Every graph G with $n \geq 3$ vertices and $\delta(G) \geq n/2$ is Hamiltonian.*

Proof. Let $G = (V, E)$ be a graph with $|G| = n \geq 3$ and $\delta(G) \geq n/2$. Then G is connected: otherwise, the degree of any vertex in the smallest component C of G would be less than $|C| \leq n/2$. Let $P_k = v_0, \ldots, v_k$ be a longest path in G. Let us call v_i the left end of the edge (v_i, v_{i+1}), and v_{i+1} its right end. By the maximality of P_k, each of the $d(v_0) \geq n/2$ neighbours of v_0 is the right end of an edge of P, and these $d(v_0)$ edges are distinct. Similarly, at least $n/2$ edges of P are such that their left end is adjacent to v_k. Since P has fewer than n edges, it has an edge (v_i, v_{i+1}) with both properties.

We claim that the cycle $C = v_0, v_{i+1}, \ldots, v_k, v_i, v_{i-1}, \ldots, v_0$ of length $k + 1$ is a Hamilton cycle of G. Indeed, since G is connected, C would otherwise have a neighbour not in C which could be used to extend P_k, violating the maximality of P_k. □

The problem of determining whether an Hamiltonian path exists in a graph is NP-complete.

Exercises.

1. Show that every 2-connected graph contains a cycle.
2. Show that every connected graph $G = (V, E)$ contains a path of length at least
$$\min\left\{2\delta(G), |V| - 1\right\}$$
3. Prove Theorem 7.
4. Show that in every connected graph whose each vertex has even degree there exists a non-trivial closed walk that contains no edge more than once.

Thanks to Ivan Masnari for spotting a mistake in the previous version of this exercise.