## 16th Lecture: 2/3

**Theorem 16.1. (Euler's theorem for digraphs)** *Let $G = (V,\ E)$ be a di(multi)graph. Then*

*(i) $G$ possesses an Euler trail from a vertex $v$ to a different vertex $w$ if and only if*
*(a) outdeg$(v)$ = indeg$(v) + 1$,*
*(b) indeg$(w)$ = outdeg$(w) + 1$,*
*(c) for all other vertices $x$, outdeg$(x)$ = indeg$(x)$.*

*(ii) $G$ possesses an Euler circuit if and only if, for every every vertex $x$, outdeg$(x)$ = indeg$(x)$.*

The theorem is proven in exactly the same way as Theorem 15.2 and yields exactly the same greedy trail/circuit-finding algorithm. The reader is left to meditate on this him/herself.

**De Bruijn graphs.** Let $\mathcal{A}$ be a finite set and $k$ a positive integer. The elements of the $k$-fold Cartesian product $\mathcal{A}^k$ can be identified with strings $a_1 a_2 \dots a_k$, $a_i \in \mathcal{A}$. In this situation it is common to speak of $\mathcal{A}$ as an *alphabet* and of the strings as *words of length* $k$ in the alphabet $\mathcal{A}$.

A digraph $G = (V,\ E)$ is called a *De Bruijn graph* if there is a finite alphabet $\mathcal{A}$ and a positive integer $k$ such that

(i) $V$ consists of all words of length $k$ in the alphabet $\mathcal{A}$,

(ii) there is a directed edge from the word $a_1 a_2 \dots a_k$ to the word $a'_1 a'_2 \dots a'_k$ if and only if $a'_i = a_{i+1}$ for all $i = 1,\ 2,\ \dots,\ k-1$.

Note that (ii) implies that there will be a loop at each vertex of the form $aa\dots a$, $a \in \mathcal{A}$. If we ignore the loops then, for a vertex $\boldsymbol{v} = a_1 a_2 \dots a_k$ one has

$$\text{Outdeg}(\boldsymbol{v}) = \text{Indeg}(\boldsymbol{v}) = \begin{cases} |\mathcal{A}| - 1, & \text{if } a_1 = a_2 = \dots = a_k, \\ |\mathcal{A}|, & \text{otherwise.} \end{cases}$$

Hence, from Theorem 16.1(ii) it follows that a De Bruijn graph always possesses an Euler circuit. Note that this applies even if we include the loops - we can imagine performing an Euler circuit in the loopless graph and executing each loop the first time we visit the corresponding vertex.

**Example 16.2. (The Keycode Problem)** In Sweden, apartment buildings are usually equipped with electronic door locks and to get into the building one must punch in the correct sequence of four decimal digits. Usually it is the case that it suffices to punch in the correct four digits consecutively. So, for example, if the code is $1234$ and you begin by erroneously punching $121$, then it suffices to continue with $234$ to gain entrance, you don't need to "start all over" and punch in the $1$ again. This feature means that a robber, seeking to gain entrance but who has no clue what the correct code is, does not in the worst case (for him !) need to punch in $4 \times 10^4 = 40,000$ digits to be absolutely sure of gaining entrance. Indeed, a sequence of just $10^4 + 3 = 10,003$ digits contains $10^4$ different codes, so the question arises whether there exists such a sequence of $10^4 + 3$

digits which includes every 4-digit code exactly once (and thus makes the robber's job easier by a factor of 4) ?

The answer is yes ! For consider the De Bruijn graph whose nodes are words of length 3 in the alphabet $\{0, 1, \ldots, 9\}$. Every directed edge in this graph corresponds to a 4-digit code and it is easy to see that an Euler circuit, including the 10 loops, corresponds to a sequence of $10^4 + 3$ digits which includes every 4-digit code exactly once.

We can "see" how this works by taking a simpler example, say $\mathcal{A} = \{0, 1\}$ and $k = 2$. The De Bruijn graph is shown in Figure 16.1(i). It contains $|\mathcal{A}|^k = 2^2 = 4$ nodes. The sequence of edges in an Euler circuit (found via the usual greedy search) is shown in Figure 16.1(ii). The corresponding sequence of $|\mathcal{A}|^{k+1} + |\mathcal{A}| = 2^3 + 2 = 10$ binary digits is

$$0001110100$$

and one may check that this includes each of the $2^3 = 8$ three-digit binary words exactly once.

**Definition 16.3.** A path in a graph $G = (V, E)$ is called a *Hamilton path* if it visits every vertex exactly once. A path which visits every vertex exactly once and then returns via an edge to the starting vertex is called a *Hamilton cycle*.

The problem of deciding whether or not a graph contains a Hamilton path or cycle is known to be much more difficult than the corresponding problem for Euler trails/circuits, which we resolved completely in the previous lecture. Indeed, it is one of the oldest known examples of a so-called *NP-complete problem*. It is beyond the scope of this course to explain what this means, but it is a central notion in the subject of *complexity theory (of algorithms)* and one reason why the subject of graph theory is so important for theoretical computer science is that it is a rich source of concrete NP-complete problems - we will be seeing further examples in the coming lectures. Philosophically, large classes of "difficult" algorithmic problems can be encoded as problems in graph theory, and the decision problem for Hamilton paths/cycles is a classic example.

A common way of popularising the decision problem for Hamilton paths/cycles is to consider it as a special case of the *Travelling Salesman Problem*. Here one thinks of the nodes in a graph as cities and the edges as representing those pairs of cities for which there exists a flight connection. The travelling salesman wishes to visit every city but has no reason to visit a place more than once, if he can avoid it. Whether or not he can achieve his goal is equivalent to asking if the graph has a Hamilton path and, assuming he'd like to end up back home where he started, whether it has a Hamilton cycle[1].

Intuitively, one can see why the decision problem for Hamilton paths is harder than the corresponding problem for Euler trails, by considering that a Hamilton path in an $n$-vertex graph uses $n - 1$ edges, while such a graph can in principle have anything

[1] In the full TSP, each edge $e$ comes equipped with a non-negative weight $w(e) \in \mathbb{R}_+$, representing the cost of the flight between those two cities. The problem is then to find a path in the graph which visits every city at least once and for which the total cost is minimised. Our special case above is to set $w(e) = 1$ for every $e$ and ask if there exists a path of total cost $|V| - 1$, or a cycle of total cost $|V|$. We will be returning to weighted graphs in Lecture 18.

from $0$ to $\frac{n(n-1)}{2}$ edges. Thus, for most graphs, a Hamilton path would use only a small fraction of the total number of available edges, though not a negligible fraction, we still have to use $n-1$ edges after all[2]. In contrast, an Euler trail must use every edge exactly once. This is a very stringent requirement, which leads to a very sharp (and restrictive !) characterisation of those graphs for which it is possible.

This intuition would, however, naturally lead one to expect that, the *denser* the graph, by which we mean the greater the quotient $|E|/\binom{n}{2}$, the greater the likelihood that Hamilton paths or cycles exist. Indeed, in the extreme case, consider $K_n$. This possesses Euler circuits if and only if $n \geq 3$ is odd, by Theorem 15.2(i). On the other hand, for each $n \geq 3$, $K_n$ possesses $\frac{n!}{n} = (n-1)!$ Hamilton cycles, since every permutation of the $n$ vertices corresponds to a Hamilton path and there are $n$ possible starting points for a given cycle.

However, high density on its own is not enough to guarantee Hamilton paths. Consider, for example, a graph which is the union of $K_{n-1}$ and an isolated vertex. It contains $\frac{n-2}{n-1}$ of all possible edges but obviously no Hamilton path. This might suggest that, in addition to having lots of edges we would like them to be "spread evenly around". There is, in fact, a theorem which makes this precise:

**Theorem 16.4. (Dirac's Theorem)** *Let $G = (V, E)$ be a graph with $|V| = n > 2$. If deg$(v) \geq n/2$ for every $v \in V$, then $G$ possesses a Hamilton cycle.*

PROOF: The proof is by contradiction. Fix an $n > 2$ and suppose the theorem is false for this value of $n$, in other words, suppose there is an $n$-vertex graph which contradicts the theorem. Then there must be such a graph with the maximum possible number of edges. Pick any such graph and call it $G$. Thus we're assuming that

(i) deg$(v) \geq n/2$ for every $v \in V(G)$,

(ii) $G$ possesses no Hamilton cycle,

(iii) adding any edge to $G$ will create a Hamilton cycle.

We will have a contradiction if we can prove that $G$ had a Hamilton cycle all along. We start by using (iii). Pick a pair of vertices $x$, $y$ such that the edge $\{x, y\}$ is not in $G$. Adding it must create a Hamilton cycle and we may assume any such cycle includes the edge $\{x, y\}$, as otherwise it would already have been present in $G$. So we can pick such a cycle and let $x$ be the "first" and $y$ the "last" vertex, i.e.: the cycle reads

$$v_1 = x \to v_2 \to v_3 \to \cdots \to y = v_n \to x.$$

Now define the subsets $S$ and $T$ of $\{1, 2, \ldots, n-1\}$ as follows:

$$S = \{i : \{x, v_{i+1}\} \in E(G)\},$$
$$T = \{i : \{v_i, y\} \in E(G)\}.$$

Note that $|S| = \text{deg}(x)$ and $|T| = \text{deg}(y)$. Hence, $|S| \geq n/2$ and $|T| \geq n/2$. But both are subsets of $\{1, 2, \ldots, n-1\}$, so $|S \cup T| \leq n-1$. It follows that $S \cap T \neq \phi$.

[2] A more rigorous way of developing this intuition is to firstly imagine the graph being chosen (uniformly) at random by inserting each of the $\binom{n}{2}$ possible edges with probability $1/2$, independent of all other edges, and then to search for a Hamilton path by taking a "random walk" from a randomly chosen starting vertex. It's beyond the scope of the course to delve into this further, though we will come back to this idea of choosing an $n$-vertex graph at random in Theorem 17.15.

Let $i \in S \cap T$, so both $\{v_i, y\}$ and $\{x, v_{i+1}\}$ are edges in $G$. We can now construct a Hamilton cycle in $G$ as follows (see Figure 16.2):

$$v_1 = x \to v_2 \to \cdots \to v_i \to y = v_n \to v_{n-1} \to \cdots \to v_{i+1} \to x.$$

This is a contradiction, completing the proof.

**Remark 16.5.** **(i)** The theorem doesn't hold for $n = 2$, since $K_2$ satisfies the requirement that every vertex has degree at least $2/2 = 1$, but obviously it has no Hamilton cycle (though it does have a Hamilton path).
**(ii)** Dirac's theorem gives a *sufficient* condition for existence of a Hamilton cycle, but it's a million miles away from being a *necessary* one. For example, the cycle $C_n$ *is* a Hamilton cycle for every $n \geq 3$, but doesn't satisfy Dirac's condition once $n \geq 5$.

**Euler's "other" theorem.** There is one other theorem which is attributed to Euler and which lies at the origins of graph theory as a subject. It is of a completely different nature to Theorem 15.2. In particular, while the latter is an *algorithmic* result for an *optimization* problem, his other result is *geometrical/topological* in nature.

**Definition 16.6.** A (multi)graph $G$ is said to be *planar* if it is possible to draw the graph on a plane surface without any two edges crossing.

Any such drawing of a planar graph is called a *plane* graph.

**Example 16.7.** The usual representations of $K_n$, $1 \leq n \leq 5$, are given in Figure 16.3. For $n \leq 3$ they are already plane. But $K_4$ is also planar since one of the diagonals can be moved outside the square (see Figure). It turns out, however, that $K_5$ cannot be untangled and is not planar (see Homework 3, Exercise 7). Hence, $K_n$ is not planar for any $n > 5$ either, because $K_n$ contains copies of $K_m$ whenever $n > m$. So $K_n$ is planar if and only if $n \leq 4$.

**Definition 16.8.** A graph $G = (V, E)$ is said to be *bipartite* if there exist subsets $V_1$, $V_2$ of $V$ such that

(i) $V_1 \neq \phi$, $V_2 \neq \phi$
(ii) $V_1 \cap V_2 = \phi$
(iii) $V = V_1 \cup V_2$
(iv) if $\{v, w\} \in E$, then one of $v$ and $w$ is in $V_1$ and the other is in $V_2$.

It is normal to write $G = (V_1, V_2, E)$ for a bipartite graph. Pictorially, a bipartite graph has two "sides" and every edge crosses from one side to the other.

**Example 16.9.** Let $m, n \in \mathbb{Z}_+$. The *complete bipartite graph* $K_{m,n}$ is the unique bipartite graph $G = (V_1, V_2, E)$, up to isomorphism, for which $|V_1| = m$, $|V_2| = n$, $\{v_1, v_2\} \in E$ for all $v_1 \in V_1$ and $v_2 \in V_2$. Note that $K_{m,n}$ has a total of $mn$ edges.

The graphs $K_{1,n}$, $K_{2,n}$ and $K_{3,3}$ are drawn in Figure 16.4. The drawing for $K_{1,n}$ is plane. That for $K_{2,n}$ isn't, but we can move one vertex to the right and get a plane drawing (see Figure), so $K_{2,n}$ is also planar. However, it turns out that $K_{3,3}$ is not planar (see Homework 3, Exercise 7) and hence that $K_{m,n}$ is planar if and only if $\min\{m, n\} < 3$.

Non-planarity of $K_5$ and $K_{3,3}$ can be deduced from the fundamental result on plane graphs proven by Euler.

**Theorem 16.10.** *Let $G = (V, E)$ be a connected plane graph. Then*

$$v - e + r = 1, \tag{16.1}$$

*where $v = |V|$ is the number of vertices, $e = |E|$ is the number of edges and $r$ is the number of minimal enclosed regions.*

**Example 16.11.** For the plane graph in Figure 16.5 one has $v = 23$, $e = 33$ and $r = 11$, the regions being numbered as in the Figure. Hence $v - e + r = 1$, as the theorem says.

PROOF OF THEOREM 16.10: The easiest way to prove the theorem is by induction on the number of edges.

*Base case:* If $G$ has one edge, then it must be a $K_2$, hence $v = 2$, $e = 1$ and $r = 0$, so yes, $v - e + r = 1$ in this case.

*Induction step:* Suppose the theorem holds for all connected, plane graphs on $n$ edges and let $G$ be a connected, plane graph on $n+1$ edges. We can certainly draw $G$ one edge at a time, in such a way that it is always plane and connected. Let $G'$ represent the drawing when one edge remains to be added. By the induction assumption, $v' - e' + r' = 1$, where the primes represent the various quantities in $G'$ and $e' = e - 1 = n$.

When we now add the last edge, since $G$ is connected two possibilities arise:

*Case 1:* This last edge joins two existing vertices. Thus no new vertex is created at this last step and $v = v'$. By joining two existing vertices we will create a new minimal enclosed region. However, we must create exactly one new such region, since $G$ is plane so the new edge does not cross any existing edge. Hence $r = r' + 1$. So $v - e + r = v' - (e' + 1) + (r' + 1) = v' - e' + r' = 1$, v.s.v.

*Case 2:* This last edge joins an existing vertex to a new vertex. Thus one new vertex is created at this last step and $v = v' + 1$. The new vertex cannot subdivide any existing edge, as otherwise we'd create two new edges at this last step, not one. Nor can the last edge create a new enclosed region, as this would have to involve it crossing an existing edge. Hence $r = r'$ in this case. Thus, $v - e + r = (v' + 1) - (e' + 1) + r' = v' - e' + r' = 1$, v.s.v.

**Remark 16.12.** Sometimes you will see Euler's Theorem for plane graphs written as $v - e + r = 2$. Here one counts the "outside" of the graph as a region. Alternatively, one thinks of the graph being drawn instead on a sphere. This version of Theorem 16.10 is sometimes easier to worth with - see, for example, Demo5, Exercise 3.

Before leaving the subject of planarity, I wish to state the fundamental result about which graphs are planar. It is well beyond the scope of this course to prove the theorem below, so the rest of this section is not examinable.

**Definition 16.13.** Let $G = (V, E)$ and $G' = (V', E')$ be graphs. We say that $G'$ is a *one-step subdivision* of $G$ if[3]

(i) $V' = V \sqcup \{x\}$, for some single vertex $x$

(ii) there is an edge $\{v, w\} \in E$ such that $E' = [E \cup \{\{v, x\}, \{x, w\}\}] \backslash \{\{v, w\}\}$.

In words, $G'$ is gotten from $G$ by inserting an extra vertex along one of its edges and thus dividing that edge into two.

More generally, we say that $G'$ is a *subdivision* of $G$ if $G'$ can be obtained from $G$ by a finite sequence of one-step subdivisions.

**Theorem 16.14. (Kuratowski's Theorem)** *A graph $G$ is planar if and only if it contains no subgraph which is a subdivision of $K_5$ or $K_{3,3}$.*

Note that in Examples 16.7 and 16.9 we have already discussed the "easy half" of this theorem, namely to show that neither $K_5$ nor $K_{3,3}$ is planar and hence that $G$ is not planar if it possesses a subgraph which is a subdivision of either of them. The much harder part is to prove that these are the only two minimal patterns which prevent a graph being planar.

[3] $\sqcup$ denotes disjoint union.