Catalan numbers, Eulerian Circuits, Hamiltonian Cycles:
Math 454 Lecture 13 (7/19/2017)

Cole Franks
July 22, 2017

Keller and Trotter Graph Theory chapter:

More is in Brualdi 11.2 (cycles, circuits) and 8.1 (Catalan numbers). If you can find a copy of Modern graph theory by Bollobas it is also a very good resource.

Contents

1 Counting Dyck Words: More Catalan 1
  1.1 Counting Dyck words via a recurrence and a generating function . . . . . . 2
  1.2 Counting Dyck words awesomely ............................................. 4

2 Walks, Cycles, Circuits ................................................................. 6

3 Eulerian Circuits ........................................................................... 6
  3.1 Connected even-degree graphs always have Eulerian circuits; different from class! ................................................................. 7

4 Hamiltonian Cycles ....................................................................... 10

1 Counting Dyck Words: More Catalan

Definition 1.1. A sequence of $n$ $U$’s and $n$ $D$’s (really doesn’t have to be $U$ and $D$, could be anything) such that every prefix has at least as many $U$’s as it has $D$’s is called a Dyck word (pronounced deek). Note that this definition switches $U$’s and $D$’s.

The following theorem summarizes what we saw last lecture:
Figure 1: The graph of a Dyck word. One can view Dyck words as mountain ranges that never dip below the horizon.

**Theorem 1.1.** There is a one-to-one correspondence between Dyck words of length $2n$ and ordered trees with $n$ edges.

Now let’s count the number of Dyck words of length $2n$. There are two ways:

1.1 Counting Dyck words via a recurrence and a generating function

Let $C_n$ be the number of Dyck words of length $2n$. Think of $C_0 = 1$ as usual. Let $w$ be a Dyck word. Think of the graph of $w$ as in Figure 1. Let $2 \leq 2k \leq 2n$ be the first point at which $w$’s graph hits the $x$ axis (we are justified in calling it $2k$ because we know it is even). The number of words that first hit the horizon at $k$ is $C_{k-1}C_{n-k}$, because the possible words between 1 and $2k-1$ are the Dyck words of length $2k-2$ and the possible words between $2k$ and $2n$ are the Dyck words of length $2n-2k$. Further, for different $k$, these sets of words are disjoint. This means that

$$C_n = \sum_{k=1}^{n} C_{k-1}C_{n-k}, \text{ for } n \geq 1$$

or

$$C_n = \sum_{k=0}^{n-1} C_kC_{n-k-1}, \text{ for } n \geq 1$$

**Theorem 1.2.** The number of Dyck words of length $n$, $C_n$, satisfies the recurrence

$$C_{n+1} = \sum_{k=0}^{n} C_kC_{n-k}, \text{ for } n \geq 0.$$
Let \( h(x) = \sum_{n \geq 0} C_n \) be the generating function for the sequence \((C_n)_{n \geq 0}\). The sum

\[
\sum_{i=0}^{n} C_i C_{n-i}
\]

should look familiar from our formula for multiplying generating functions. If \( f(x) = \sum_{i \geq 0} a_i \), and \( g(x) = \sum_{j \geq 0} b_j \), then

\[
f(x)g(x) = \sum_{n \geq 0} \left( \sum_{i=0}^{n} a_i b_{n-i} \right) x^n.
\]

This means if we multiply \( h(x) \) by itself, we get

\[
h(x)h(x) = \sum_{n \geq 0} \left( \sum_{i=0}^{n} C_i C_{n-i} \right) x^n = \sum_{n \geq 0} C_{n+1} x^n = \frac{1}{x} \left( \sum_{n \geq 1} C_n x^n \right).
\]

Since \( C_0 = 1 \),

\[
\frac{1}{x} \left( \sum_{n \geq 1} C_n x^n \right) = \frac{h(x) - 1}{x},
\]

so

\[
h(x) = 1 + xh(x)^2.
\]

This says \( h(x) \) is a solution of the equation \( xh(x)^2 - h(x) + 1 \), or \( h(x) = \frac{1 \pm \sqrt{1-4x}}{2x} \). By our definition, we should have \( h(0) = 1 \), but \( \frac{1 \pm \sqrt{1-4x}}{2x} \) tends to infinity as \( x \to 0 \), let’s try the other root \( \frac{1 + \sqrt{1-4x}}{2x} \). Since this has a power series expansion, its coefficients must satisfy the same recurrence as \( h(x) \), and it has first coefficient 1 in its power series, so the coefficients must actually match \( x \). We can apply Newton’s binomial theorem to obtain

\[
h(x) = \frac{1}{2x} - \frac{1}{2x} (1 - 4x)^{1/2} = \frac{1}{2x} - \frac{1}{2x} \sum_{k \geq 0} \binom{1/2}{k} (-4)^k x^k = -\frac{1}{2} \sum_{k \geq 1} \binom{1/2}{k} (-4)^k x^{k-1}.
\]

The reason for the last equality is that \( \binom{1/2}{0} \) is just 1, so the \( 1/2x \) terms cancel out. For
For us, everything miraculously cancels:

\[ h(x) = -\frac{1}{2} \sum_{k \geq 1} \binom{1/2}{k} (-4)^{k} x^{k-1} \]

\[ = -\frac{1}{2} \sum_{k \geq 1} \frac{(-1)^{k-1}}{k} \frac{(2k-2)}{(k-1)!} x^{k-1} \]

\[ = \sum_{k \geq 1} \frac{(-1)^{k}}{k \cdot 2^{2k-1}} \frac{(2k-2)}{(k-1)!} x^{k-1} \]

\[ = \sum_{k \geq 1} \frac{1}{k} \left( \frac{2k-2}{k-1} \right) x^{k-1}, \]

\[ = \sum_{k \geq 0} \frac{1}{k+1} \binom{2k}{k} x^{k}, \]

So

\[ C_{n} = \frac{1}{n+1} \binom{2n}{n}. \]

These are called the Catalan numbers, and they count lots of other quantities in combinatorics.

1.2 Counting Dyck words awesomely

We want to count the number of sequences of \( n \) U's and D's whose graph does not cross the x axis. To do this, we can also count

the total number of sequences - those that do cross the x axis.
The first term is, of course, \( \binom{2n}{n} \). The second is a little trickier. We count these using a one-to-one and onto function. Let \( S \) be the set of sequences of \( n \) U’s and \( n \) D’s whose graph crosses the \( x \)-axis. If \( s \in S \), let \( f(s) \) be the sequence where \( U \)’s and \( D \)’s are switched \textit{after the first time} the graph of \( s \) crosses the \( x \)-axis.

Since the graph of \( s \) just crossed the \( x \) axis, there must be \( k + 1 \) U and \( k \) D’s in \( s \) after the point at which it crossed the axis (we have to get back to the \( x \) axis by the end of \( s \)). Then \( f(s) \) must have \( n - 1 \) U’s and \( n + 1 \) D’s, because swapping \( U \)’s and \( D \)’s after crossing increases the number of \( U \)'s by one and decreases the number of \( D \)'s by one. Thus \( f(s) \) is a sequence of \( n - 1 \) U’s and \( n + 1 \) D’s.

Now our goal is to show that
\[
f : \{ \text{sequences that do cross the x axis} \} \rightarrow \{ \text{permutations of } \{(n - 1) \cdot U, (n + 1) \cdot D\} \}
\]
is one-to-one and onto. Once we do that, we are done, because the right hand side has size \( \binom{2n}{n+1} \).

\( f \) is one-to-one because this process is reversible - in particular there is a function
\[
g : \{ \text{permutations of } \{(n - 1) \cdot U, (n + 1) \cdot D\} \} \rightarrow \{ \text{sequences that do cross the x axis} \}
\]
such that \( f(g(r)) = r \) for any permutations \( r \) of \( \{(n - 1) \cdot U, (n + 1) \cdot D\} \).

\( g \) simply takes a sequence \( r \) of \( n - 1 \) U’s and \( n + 1 \) D’s to the sequence of \( n \) U’s and \( n \) D’s obtained by swapping \( U \) and \( D \) after the first time \( r \)'s graph crosses the \( x \)-axis - which it must, because there are \( n - 1 \) U’s and \( n + 1 \) D’s. This equalizes the number
of $U'$s and $D'$s in $g(r)$, and you can see that if you apply $f$ to $g(r)$ you will get $r$ back.

Finally, we can say $C_n = \binom{2n}{n} - \binom{2n}{n-1} = \frac{1}{n+1} \binom{2n}{n}$.

## 2 Walks, Cycles, Circuits

We will be talking about looped, directed multigraphs here.

**Definition 2.1.** A walk of length $l$ in a directed, looped, multigraph $G$ a sequence of $l + 1$ vertices $(x_0, \ldots, x_l)$ such that $x_{i-1}x_i$ is an edge (arc in directed graphs) of $G$ for $i \in [l]$. A closed walk of length $l$ is a walk of length $l$ $(x_0, \ldots, x_l)$ where $x_0 = x_l$.

**Example 1.** Careful about following directions.

Top left is the walk $(3, 5, 1, 2, 3, 4, 1)$. Top right is the walk $(D, C, D, A, D, B, D)$. Bottom left is not a walk, because it violates the direction of the edge $(1, 4)$. 


3 Eulerian Circuits

Definition 3.1. A trail of length \( l \) in a graph \( G \) is a walk of length \( l \) that does not repeat edges, and a circuit is a closed walk that does not repeat edges.

Example 2. Example in directed and in undirected.

Definition 3.2. A Eulerian trail in a graph \( G \) is a trail that uses every edge, and a Eulerian circuit is a circuit that uses every edge.

3.1 Connected even-degree graphs always have Eulerian circuits; different from class!

Remark 3.1. Warning: we need to assume the graphs have no isolated vertices for the rest of the section to make sense. Frustratingly, the following circuit in a 7 vertex graph satisfies the definition of an Eulerian circuit:

This means the theorem in class was slightly off. For the rest of this section we assume our graphs don’t have isolated vertices; that is, vertices with no neighbors.

Suppose \( G \) is an undirected multigraph with no isolated vertices. What conditions do we need for there to be an Eulerian circuit?

• First of all, we need to be able to get from each edge to every other edge in \( G \), because the circuit visits every edge. This means \( G \) must be connected.

• Secondly, the circuit will enter and leave each vertex the same number of times, so the degrees must be even.

What if \( G \) is a directed multigraph? Then we will still need to be able to get from each arc to every other arc in \( G \), because the circuit visits every arc. However, this time we need to be able to do visit all arcs while following the directions of the arcs. For this we need

---

2By Cmglee - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=60330862
1.3 Hamilton Cycles and Euler Circuits

The seven bridges on the Pregel in Königsberg.

![Image of the seven bridges on the Pregel in Königsberg.](image)

Figure 2: Top is the very old 7 bridges of Königsberg puzzle. The residents of Königsberg asked if there was a walk that could take you across each bridge exactly once. Bottom is a graph in which the blobs of land have been treated as vertices. An Eulerian trail in this graph would solve the Königsberg puzzle in the affirmative.

Figure 3: An Eulerian circuit in the directed graph $DB_4$, and the corresponding de Bruijn sequence.\(^2\)
a condition called strongly connected: a directed multigraph \( G \) is strongly connected if for any two vertices \( x \) and \( y \) there is a walk from \( x \) to \( y \) in \( G \) (respecting directions of arcs). It turns out, however, that we don’t need this!

**Definition 3.3** (Unofficial definition:). If \( G \) is a directed graph, the undirected version of \( G \) is the graph obtained by replacing each arc (ordered pair) by the corresponding edge (ordered pair).

We just need that the undirected version of \( G \) is connected; strong connectivity comes for free from the degree condition. Hence, we just need the following conditions.

- The undirected version of \( G \) is connected.
- Secondly, the circuit will enter and leave each vertex the same number of times, so the number of arcs leaving every vertex \( v \), called the in-degree \( d_-(v) \) of \( v \), must be the same as the number of arcs entering every vertex, called the out-degree \( d_+(v) \).

Summarizing,

**Theorem 3.1** (Eulerian circuits for looped undirected multigraphs). A looped multigraph with no isolated vertices has an Eulerian circuit if and only if it is connected and all its degrees are even.

**Theorem 3.2** (Eulerian circuits for looped directed multigraphs). A looped directed multigraph \( G \) with no isolated vertices has an Eulerian circuit if and only if

\[
d_+(v) = d_-(v)
\]

for each vertex \( v \in G \), and (different from class!) the undirected version of \( G \) is connected.

This is similar to the proof from class, but less inductive and more constructive.

*Proof for undirected graphs.* We already saw why \( G \) must be connected with even degrees if it has an Eulerian circuit. Now suppose \( G = (V, E) \) is a connected multigraph with all degrees even and no isolated vertices. Then \( G \) must have a circuit \( C \); the endpoint \( x_l \) of the longest path \( x_1 \ldots x_l \) in \( G \) is of degree at least two and has no neighbors off the path (else longer path!), so must be adjacent to some vertex among \( \{x_1, \ldots, x_{l-1}\} \), forming a cycle (which is also a circuit).

Let this circuit \( C \) be our starting point. If \( C \) is all the edges, we are done; else, \( G(V, E \setminus E(C)) \) must have some connected component \( H \) with more than one vertex. Here \( E(C) \) means the edge-set of \( C \). Since \( G \) is connected, \( C \) must join the connected components of \( G(V, E \setminus E(C)) \), so it must visit \( H \).
Also note that all degrees of the vertices of \( G(V, E \setminus E(C)) \) are even (because they decreased by an even number when we removed \( E(C) \)), so the degrees of \( H \) are all at least two. By the reasoning from before, this means we can find a circuit \( D \) in \( H \) through the vertex it shares with the circuit \( C \). Let \( C_2 \) be the concatenation of \( C \) and \( D \); that is, if \( C = (x_1, \ldots, x_l) \) and \( D = (y_1, \ldots, y_k) \) then \( C_2 = (x_1, \ldots, x_l, y_1, \ldots, y_k) \). Since we can assume \( y_k = y_1 = x_l = x_1 \) and \( C \) and \( D \) share no edges, \( C_2 \) is still a circuit. \( C_2 \) is also larger than \( C \).

Repeat this process with \( C_2 \) and \( G(V, E \setminus E(C_2)) \), and so on. Eventually we will run out of edges, so at that stage we will have completed an Eulerian circuit.

Here is a sketch of how to fix the proof to work for directed multigraphs.

**Proof for directed graphs:** The proof is essentially the same, but after finding the circuit \( C \) we split into the components of the undirected version of \( G(V, E \setminus E(C)) \). Also, when we concatenate \( C \) and \( D \) we must go round \( D \) in the correct direction.

\[ \square \]

4 Hamiltonian Cycles

**Definition 4.1.** A Hamiltonian cycle in a graph \( G \) is a cycle that visits every vertex exactly once. Recall that cycles do not repeat vertices or edges.

**Lemma 4.1.** Suppose a set \( M \) of disjoint edges disconnects the graph \( G \). Then any Hamiltonian cycle in \( G \) must pass through the edges of \( M \) an even number of times.

**Example 3.**
A set of $M$ disjoint edges connecting a graph, where the pink parts are the (at least) two connected components of $G = (V, E \setminus M)$. 