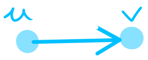
Math 38 — Graph Theory
Directed graphs

Nadia Lafrenière
04/13/2022

We introduce directed graphs and their terminology. Applications include Markov chains, automata and De Bruijn graphs.

A directed graph or digraph is made of two sets: the vertices, and a set of edges defined as ordered pairs of two vertices: a tail and a head. For one edge, the tail and the head are both endpoints, and we say the edge is from its tail to its head. We sometimes use the word arrow for the edges of a directed graph.

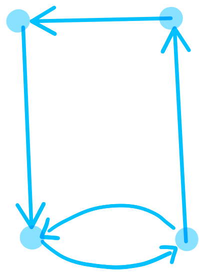


Edge from u to v

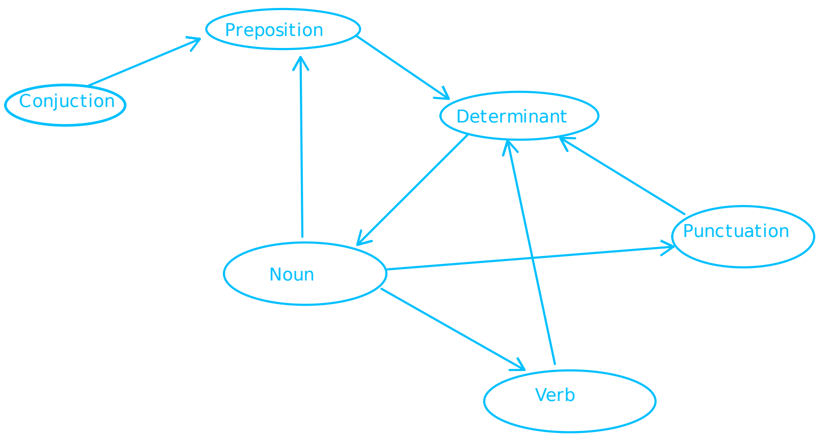
Some examples

Trail etiquette

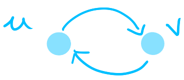
Car traffic around the Green



Parts of speech in the sentence: "While at the beach, the dog eats the biscuits in the box."



Like in undirected graphs, a loop is an edge with its two endpoints being equal. Multiple edges are edges having the same tail and the same head.

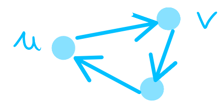


Not a multiple edge

A directed graph is simple if there is no loop nor multiple edges.

Example: All the directed graphs above are simple.

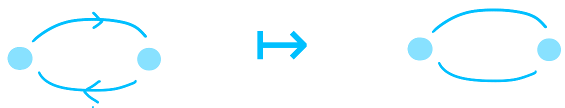
In a simple digraph, we write the edge from u to v as uv (and so this is not the same as vu). If uv is in the graph, v is a successor of u and u is a predecessor of v.



A simple digraph is a path if its vertices can be ordered so that v follows u in the vertex ordering if and only if there is an edge from u to v. The only vertex that can be repeated is the first and the last vertices, if they are equal; the path is then a cycle.
Equivalently, we can define walks and trails (walks without repeated edges) in the same way as in undirected graphs.

The underlying graph of a digraph D is the undirected graph G in which we removed the orientation of the edges. Hence, uv=vu, and if uv and vu both appear in D, uv is a multiple edge of G.

Remark: The underlying graph of a simple digraph is not always a simple graph.



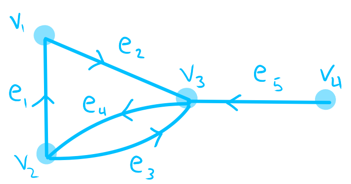
Subgraphs and isomorphisms are defined in the same way as for undirected graphs.

Adjacency and incidence matrix

In a digraph, the adjacency and incidence matrices are not defined in the same way as in graphs.

The adjacency matrix A(D) of a loopless digraph D has u,v−entry the number of edges from u to v. The incidence matrix has v,e−entry +1 if v is tail of e, −1 if it is its head, and 0 if v is not an endpoint.

The digraph on the left has the following adjacency and incidence matrices:



$$A(D) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

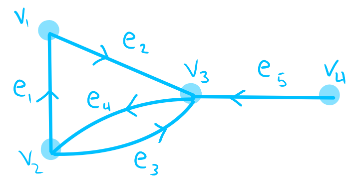
The adjacency matrix is no more symmetric!

$$M(G) = \begin{pmatrix} -1 & +1 & 0 & 0 & 0 \\ +1 & 0 & +1 & -1 & 0 \\ 0 & -1 & -1 & +1 & -1 \\ 0 & 0 & 0 & 0 & +1 \end{pmatrix}$$

Connectedness: weak and strong

A digraph is weakly connected if its underlying graph is connected. It is strongly connected if there is a path from u to v, for every two vertices u and v.

The graph below is weakly connected, but not strongly connected, as there is no path from vertex 3 to vertex 4.



The strong components of a digraph are its maximal strongly connected subgraphs.

Degree and neighborhood, in and out

Let v be a vertex in a directed graph. Its outdegree is the number of edges that have v has a tail, and is noted $d^+(v)$. The indegree is the number of edges that have v has a head, $d^-(v)$.

The number of edges is $\sum_{v\in V} d^+(v) = \sum_{v\in V} d^-(v)$.
The out-neighborhood of v is the set of vertices $\{u : vu \text{ is an edge}\}$.
The in-neighborhood is defined similarly.

Eulerian graphs

A digraph is Eulerian if it contains an Eulerian circuit, i.e. a trail that begins and ends in the same vertex and that walks through every edge exactly once.

Obviously, a graph will not be Eulerian if it has more than one nontrivial component or if the sum of the in and out degree of some vertex is odd. The following theorem gives a classification of Eulerian digraphs.

Theorem

A digraph is Eulerian if and only if it there is at most one nontrivial strong component and, for every vertex v, $d^+(v)=d^-(v)$.

Proof

($\Rightarrow$) If there is an Eulerian circuit, it visits all the vertices in a nontrivial component, so there is at most one of them. Also, the Eulerian circuit goes in and out of v the same number of times, which means the in- and outdegrees must be equal.

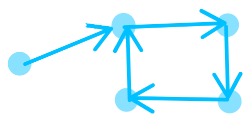
(⇐) We prove by induction on the number of edges that if the in- and outdegrees are the same at every vertex in a strongly connected graph, there is an Eulerian circuit.

Base case: When there is no edge, the empty circuit is Eulerian.

Induction hypothesis: Suppose that, whenever there is at most m edges, every graph that has, at each vertex, the same in- and outdegree, and that has at most one non-trivial strong component, is Eulerian.

- For a graph with m+1 edges, we first prove the following lemma:

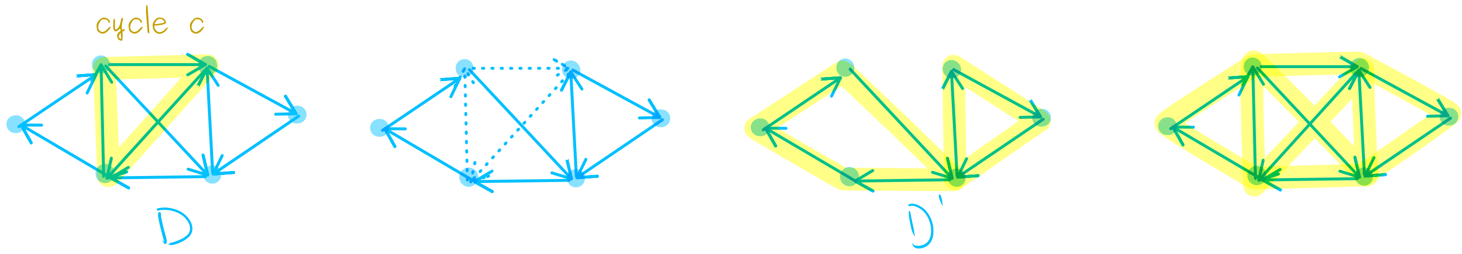
> Lemma
>
> If the outdegree of every vertex is at least 1, then the digraph has a cycle.
>
> Proof (of the lemma)
>
> Let v be a vertex. Since it has outdegree at least 1, there is a walk starting at v. Since every vertex has outdegree at least 1, the walk can always be extended. Since the number of vertices in D is finite, the walk will go back to a vertex it already visited. The first time this happens, the part of the walk between the two occurrences of a vertex is a cycle.



Proof of the theorem (continued)

For a graph with m+1 edges, consider the unique nontrivial strong component. The lemma applies to it, so there is a cycle c. Removing the edges of c to the digraph preserves the equality of the in- and outdegrees. Let D' be that reduced graph. We can apply the induction hypothesis to get an Eulerian circuit in each strong component of D'. Each such component shares at least one vertex with c, since they are in the same strong component of D. To build an Eulerian circuit, we travel through c. Each time we get to a vertex that has neighbors not in c, we visit all the edges in its strong component: we know it is possible since the component is Eulerian. That process gives an Eulerian circuit in the original digraph.

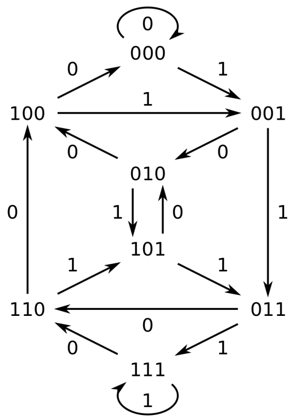
∎



Application: De Bruijn cycles

Let D_{n+1} be the following digraph:

- vertices are binary sequences of length n
- there is an edge from a sequence s to another s' if the $n-1$ last letters of s are the $n-1$ first letters of s'.



As an example, D_4 is illustrated on the left.

Proposition

The De Bruijn graphs are Eulerian.

For the proof, use the previous theorem and verify the equality of the out- and indegrees.

Problem: What is the minimum length for a sequence containing all the binary sequences of length n? To solve this problem, use the above proposition and your homework.

Reference: Douglas B. West. Introduction to graph theory, 2nd edition, 2001. Section 1.4