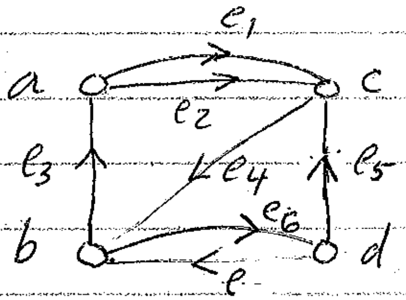
Ch.2 – Connectedness & Distance

§1 Walks, trails, circuits, cylces & paths.

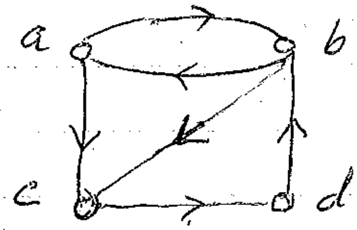
Def. Let G be a digraph-like object. A <u>directed walk</u> in G is a finite alternating sequence $\langle v_0, \vec{e}_1, v_1, \vec{e}_2, \ldots, v_{n-1}, \vec{e}_n, v_n \rangle$ of vertices & directed edges in G such that the initial & terminal endpoints of e_i are v_{i-1} & v_i respectively. The directed walk is then said to be from v_0 to v_n. The sequence $\langle v_0, v_1, \ldots, v_n \rangle$ is called the <u>vertex sequence</u> of the directed walk. When G is a pseudo-digraph the directed walk is completely determined by the vertex sequence alone. The directed walk is <u>closed</u> $v_0 = v_n$ & <u>open</u> if $v_0 \neq v_n$.

Ex.1 Let $G =$



Then $\langle b, \vec{e}_3, a, \vec{e}_1, c, \vec{e}_4, b, \vec{e}_6, d, \vec{e}_5, c \rangle$ is an open directed walk.
Also $\langle a, \vec{e}_1, c, \vec{e}_4, b, \vec{e}_3, a \rangle$ is a closed directed walk.

Def. A <u>directed trail</u> in G is a directed walk in which no directed edge is repeated. A <u>directed circuit</u> is a closed directed trail. A <u>directed path</u> is a directed trail in which no vertex is repeated. A <u>directed cycle</u> is a directed circuit in which all vertices are distinct except that the first vertex & the last vertex are the same.

Ex. 2 Let $G =$  . Then

(a) $\langle a, b, a, c, d \rangle$ is the vertex sequence of an open directed trail.

(b) $\langle a, b, c, d, b, a \rangle$ is the vertex sequence of a directed circuit

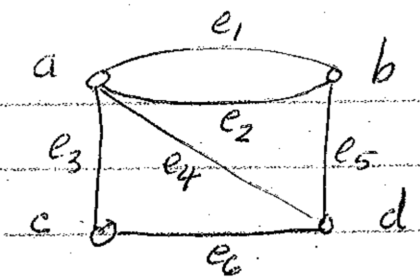
(c) $\langle a, b, c, d \rangle$ is the vertex sequence of a directed path.

(d) $\langle a, b, a \rangle$ & $\langle a, c, d, b, a \rangle$ are vertex sequences of directed cycles.

In graph-like objects we get similar concepts by replacing the directed edges by edges. Recall that an edge between two distinct vertices is really a pair of oppositely matched directed edges.

Def. A walk in a graph-like object G is a finite alternating sequence $\langle v_0, e_1, v_1, e_2, \ldots, v_{n-1}, e_n, v_n \rangle$ of vertices and edges. A trail is a walk in which no edge is repeated. A circuit is a closed trail and a path is a trail in which no vertex is repeated. A cycle is a circuit in which all vertices are distinct except that the first & last vertices are the same.

Def. The length of a walk is the number of edges in it

<u>Ex3</u> Let G =  . Then

(a) $\langle a, e_1, b, e_2, a, e_4, d, e_4, a, e_3, c\rangle$ is an open walk in G.

(b) $\langle c, e_3, a, e_1, b, e_2, a\rangle$ is an open trail in G.

(c) $\langle a, e_1, b, e_2, a\rangle$ is a cycle in G

(d) $\langle c, e_3, a, e_1, b\rangle$ is a path in G but $\langle c, e_3, a, e_1, b, e_2, a\rangle$ is <u>not</u> a path in G.

(e) $\langle a, e_1, b, e_2, a, e_4, d, e_5, b\rangle$ is a circuit in G.

<u>Prop. 1</u>. Let G be a graph-like object and $W = \langle v_0, e_1, v_1, \ldots, v_{n-1}, e_n, v_n\rangle$ be a walk from x to y in G. Then we can find a path P from x to y in G.

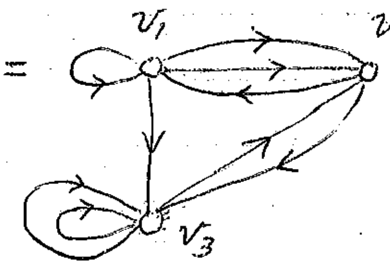
<u>Proof</u>: First observe that since W is a walk from x to y, then $v_0 = x$ and $v_n = y$. Now if if $x = y$, then the empty path (or path with no edges) $\langle v_0\rangle$ a directed path from x to y. So suppose $x \neq y$. Starting at v_0, keep going along the walk W until a vertex v_j is reached with $v_j = v_i$ for some $i<j$ for the first time or until you reach v_n. If you reach v_n, then you have your path P from x to y; otherwise, delete the portion of the sequence $\langle e_i, v_{i+1}, \ldots, v_{j-1}, e_j, v_j\rangle$ and keep going until a vertex is repeated for the first

or until you reach v_n. If we repeatedly do this we will end up with a path P from x to y.

Ex. 4 Let $W = \langle a, e_1, b, e_2, a, e_4, d, e_4, a, e_3, c \rangle$ be the walk from a to c in Ex. 3(a). Then the first repeated vertex is a, so we delete $\langle e_1, b, e_2, a \rangle$ to get $W_1 = \langle a, e_4, d, e_4, a, e_3, c \rangle$. Again a is the first repeated vertex, so we delete $\langle e_4, d, e_4, a \rangle$ to get $W_2 = \langle a, e_3, c \rangle$. W_2 is then the path P from a to c in G, that we sought.

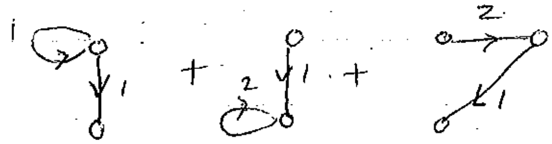
<u>Prop. 2</u>: Let G be a digraph-like object and $W = \langle v_0, \vec{e}_1, v_1, \cdots, v_{n-1}, \vec{e}_n, v_n \rangle$ be a directed walk from x to y in G. Then we can find a directed path P from x to y in G.

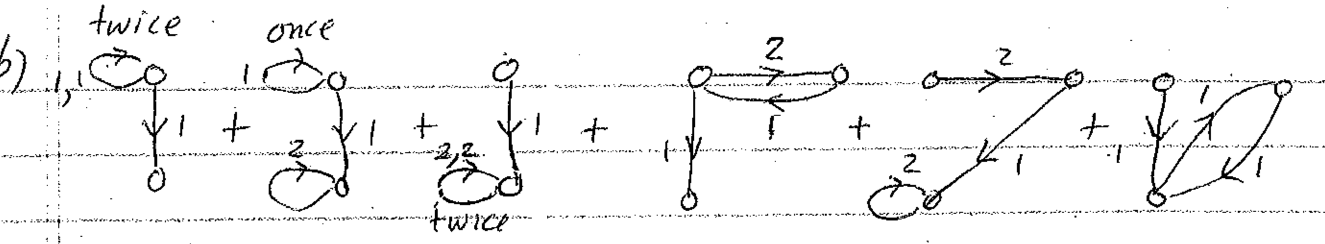
<u>Proof</u>: Do for H.W.

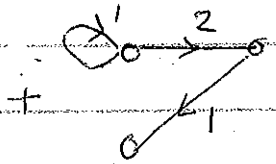
Ex 5 Let $G =$  be a digraph-like object.

(a) How many directed walks of length two are there from v_1 to v_3?

(b) How many directed walks of length three are there from v_1 to v_3?

(a)  Ans: $= 1(1) + 1(2) + 2(1) = \boxed{5}$

(b) 



Ans: $1(1)1 + 1(1)(2) + 1(2)2 + 2(1)1$
$+ 2(1)2 + 1(1)1 + 1(2)1 = \boxed{16}$

Surely, there must be an easier way – and there is!

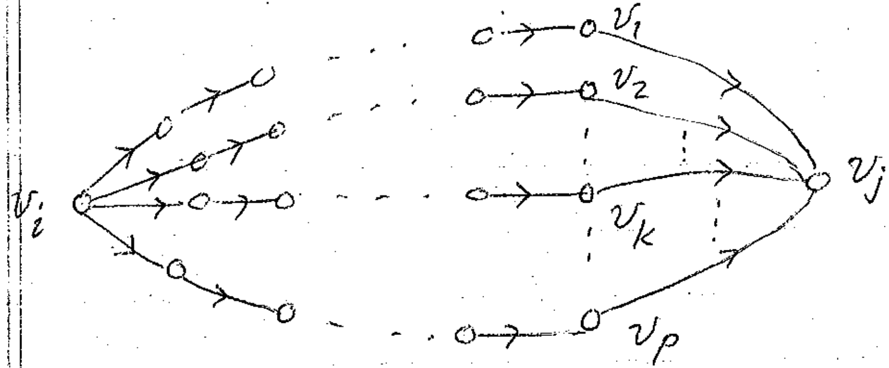
Recall that the adjacency matrix of the digraph-like object G on the vertices $v_1, \ldots, v_p$ was defined by $A[i,j]$ = no. of directed edges from v_i to v_j.
Recall also that the matrix A^n is defined recursively as follows. (a) $A^0 = I_p$ and
(b) $A^{n+1} = A^n . A$ for $n \geq 0$.

<u>Theorem 3</u> Let G be a digraph-like object with vertices $\{v_1, v_2, \ldots, v_p\}$. Then the number of directed walks of length n from v_i to v_j is $(A^n)[i,j]$.

<u>Proof</u>: We will prove the result by induction on n

<u>Basis</u>: For $n = 1$, we have
$A^1[i,j]$ = number of directed edges from v_i to v_j
= no. of directed walks of length 1 from v_i to v_j.
So the result is true for $n = 1$, for all i & j.

<u>Ind. Step</u>: Suppose the result is true for directed walks of length n, for all i & j. Then
$(A^n)[i,j]$ = no. of directed walks of length n from v_i to v_j for all $1 \leq i, j \leq p$. Now



No. of directed walks of length $n+1$ from v_i to v_j

$$= \sum_{k=1}^{p} \begin{pmatrix}\text{no. of directed walks} \\ \text{of length } n \text{ from } v_i \text{ to } v_k\end{pmatrix} \cdot \begin{pmatrix}\text{no. of directed walks} \\ \text{of length } 1 \text{ from } v_k \text{ to } v_j\end{pmatrix}$$

$$= \sum_{k=1}^{p} (A^n)[i,k] \cdot A^1[k,j] = (A^n \cdot A^1)[i,j] = (A^{n+1})[i,j].$$

So if the result is true for n, it will be true for $n+1$. Hence the result is true for all n by the Principle of Mathematical Induction.

Note the result is also true for $n=0$ because

$$A^0[i,j] = (I_p)[i,j] = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j \end{cases} = \begin{pmatrix}\text{no. of walks of} \\ \text{length 0 from } v_i \text{ to } v_j\end{pmatrix}$$

Ex. 5 (again) (a) Number of directed walks of length 2 from v_1 to v_3 $= (A^2)[1,3] = \boxed{5}$

$$\underbrace{\begin{bmatrix} 1 & 2 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 1 & 2 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} 3 & 3 & \boxed{5} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}}_{A^2}$$

(b) No. of directed walks of length 3 from v_1 to v_3 $= (A^3)[1,3] = \boxed{16}$

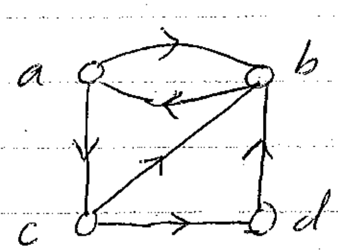
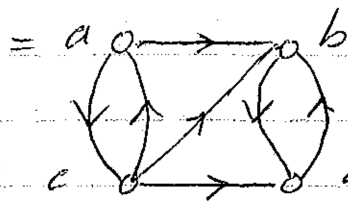
$$\underbrace{\begin{bmatrix} 3 & 3 & 5 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}}_{A^2} \underbrace{\begin{bmatrix} \cdot & \cdot & 1 \\ \cdot & \cdot & 1 \\ \cdot & \cdot & 2 \end{bmatrix}}_{A} = \begin{bmatrix} \cdot & \cdot & 16 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$$

Theorem 3': Let G be a graph-like object with vertices $\{v_1, v_2, \dots, v_n\}$. Then the number of walks of length n from v_i to v_j is $(A^n)[i,j]$.

§2 Connected digraphs and graphs

Def A digraph G is connected (some textbooks say strongly connected) if there is a directed path from any vertex of G to any other vertex of G.

Def Let G be a digraph. A semi-path in G is an alternating sequence $v_0, e_1, v_1, \dots, v_{n-1}, e_n, v_n$ of vertices and directed edges such that the endpoints of e_i are v_{i-1} and v_i. (We do not require that e_i be a directed edge from v_{i-1} to v_i.) The digraph G is said to be semi-connected (some textbooks say weakly-connected) if there is a semi-path from any vertex of G to any other vertex of G.

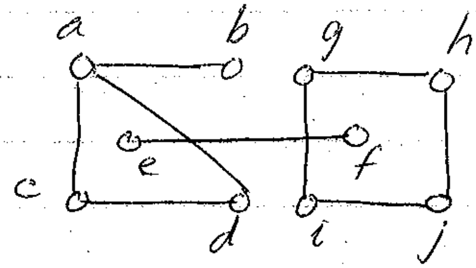
Ex. 1 Let $G_1 =$  & $G_2 =$ 

Then G_1 is connected but G_2 is not connected because there is no directed path from b to a. G_2 is, however, semi-connected.

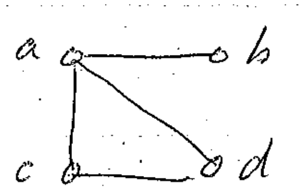
Def. A graph G is connected if there is a path between any two vertices of G. (Note that when a graph is viewed as a digraph, it will also be connected.)

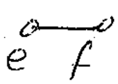
Def. Let G be a graph. A subgraph H of G is said to be a <u>maximal connected subgraph</u> of G if H is connected and there is no connected subgraph H' of G which properly contains H. The maximal connected subgraphs of G are called the <u>connected components</u> of G.

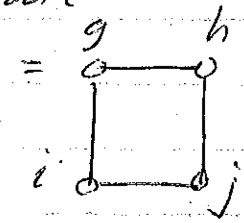
Ex 2 Let G =



Then the connected components of G are

G_1 = 

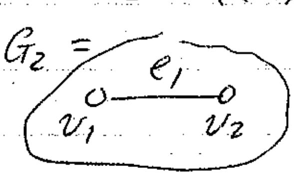
, G_2 = 

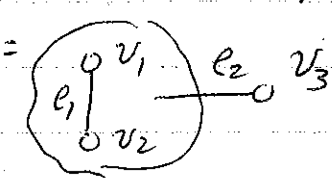
, & G_3 = 

<u>Prop. 4</u> If G is a connected graph with p vertices, then G has at least $p-1$ edges.

<u>Proof:</u> Choose any vertex of G and call it v_1. Put $G_1 = \langle \{v_1\}, \emptyset \rangle$. Then G_1 is a subgraph of G with 0 edges. Since G is connected, there must be an edge, e_1 say from a vertex in $V(G_1)$ to a new vertex, v_2 say, in $V(G) - V(G_1)$. Let $G_2 = \langle \{v_1, v_2\}, \{e_1\} \rangle$. Similarly,

G_1 = 

G_2 = 

G_3 = 

since G is connected, there must be an edge,

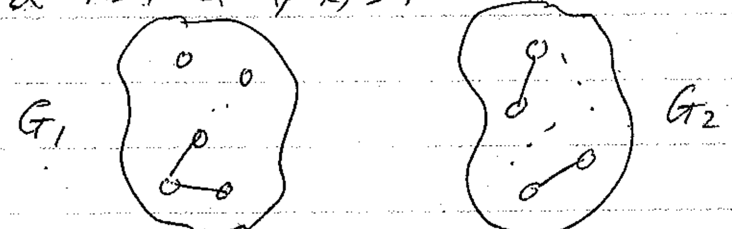
e_2 say, from G_2 to a new vertex, v_3 say, in $V(G)-V(G_2)$. Let $G_3 = \langle\{v_1, v_2, v_3\}, \{e_1, e_2\}\rangle$. If we keep doing this we will end up with a subgraph

$$G_p = \langle\{v_1, \dots, v_p\}, \{e_1, \dots, e_{p-1}\}\rangle$$

of G. Since G_p has $p-1$ edges, it follows that G will have at least $p-1$ edges.

<u>Prop. 5</u> (a) If G is a disconnected graph with p vertices then $|E(G)| \leq (p-1)(p-2)/2$.
(b) If G has p vertices & $|E(G)| > (p-1)(p-2)/2$, then G is connected.

<u>Proof</u>: (a) Suppose G is disconnected. Then we can split G into two parts G_1 & G_2 such that there is no edge from G_1 to G_2. Let $k = |V(G_1)|$. Then $|V(G_2)| = p-k$ & $k \geq 1$ & $(p-k) \geq 1$.



So $|E(G_1)| \leq k(k-1)/2$ & $|E(G_2)| \leq (p-k)[(p-k)-1]/2$.
Hence $|E(G)| \leq \frac{1}{2}\{k(k-1) + (p-k)(p-k-1)\}$. So

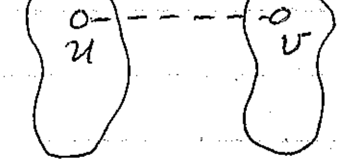
$$\begin{aligned}
&(p-1)(p-2)/2 - |E(G)| \\
&\geq \tfrac{1}{2}\{(p-1)(p-2) - k(k-1) - (p-k)(p-k-1)\} \\
&= \tfrac{1}{2}\{p^2 - 3p + 2 - k^2 + k - p^2 + pk + pk - k^2 + p - k\} \\
&= \tfrac{1}{2}\{2pk - 2k^2 - 2p + 2\} = (pk - k^2 - p + 1) \\
&= \underbrace{(k-1)}_{\geq 0}\underbrace{(p-k-1)}_{\geq 0} \geq 0.
\end{aligned}$$

Hence $|E(G)| \leq (p-1)(p-2)/2$. Part (b) follows immediately from part (a).

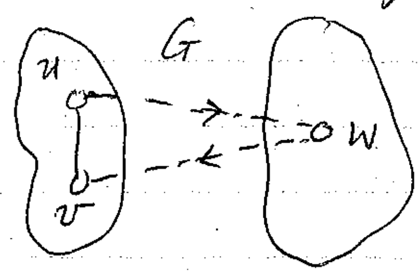
<u>Prop 6</u>: If G is a disconnected graph, then G^c will be a connected graph.

<u>Proof</u>: Let u & v be any two vertices of G^c. We must show that there is a path from u to v in G^c. There are two cases.

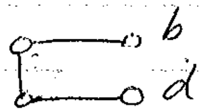
<u>Case (i)</u>: $uv \notin E(G)$.
In this case uv must be an edge in G^c. So we instantly get a path from u to v in G^c.

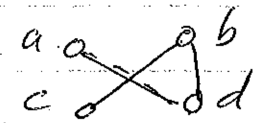
$G =$ 

<u>Case (ii)</u>: $uv \in E(G)$.
In this case u & v will be in the same connected component of G because there is an edge from u to v in G. Since G is a disconnected graph, G must have at least one other connected component. Let w be any vertex in this other connected component. Then $uw \notin E(G)$ & $wv \notin E(G)$. So $uw \in E(G^c)$ & $wv \in E(G^c)$. Hence we again get a path $\langle u, w, v \rangle$ from u to v in G^c.

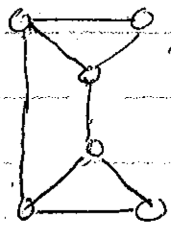
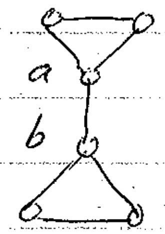


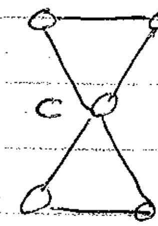
So in either case we got a path from u to v in G^c. Hence G^c is a connected graph.

<u>Ex. 3</u>. Let $G =$ 

. Then $G^c =$ 

Observe that both G & G^c are connected. Note also that $G^c \cong G$. Graphs such as G are called self-complimentary.

<u>Ex.5(a)</u> Let $G_1 =$ , $G_2 =$ 

, & $G =$ 

Which of the graphs above is the most connected?
Which of the graphs above is the least connected?

<u>Def.</u> We define the <u>vertex-connectivity</u>, $k_V(G)$, of a graph G by
$k_V(G) =$ minimum no. of vertices whose removal will disconnect G or reduce it to K_1.
We define the <u>edge-connectivity</u>, $k_E(G)$, of G by
$k_E(G) =$ minimum no. of edges whose removal will disconnect G or reduce it to K_1.

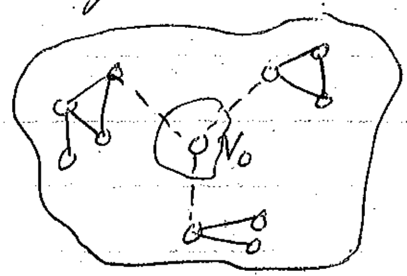
<u>Ex.5(b)</u> In Ex.5(a) $k_V(G_1) = 2$, $k_V(G_2) = 1$, $k_V(G_3) = 1$
$k_E(G_1) = 2$, $k_E(G_2) = 1$, $k_E(G_3) = 2$.
So in a certain sense G_1 is the most connected and G_2 is the least connected.

<u>Def.</u> Let G be a connected graph. A <u>cut-vertex</u> of G is any vertex v such that $G - \{v\}$ is a disconnected graph. A <u>bridge</u> (or cut-edge) is any edge e such that $G - \{e\}$ is a disconnected graph.

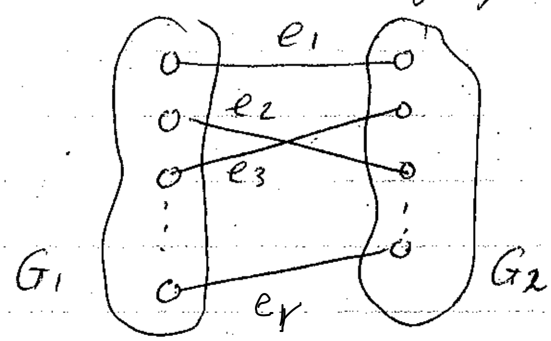
<u>Ex.5(c)</u> In the graph G_2 in Ex.5(a), the edge $e = ab$ is a bridge. In G_2, both a & b are cut-vertices; and in G_3, c is also a cut-vertex.

Prop. 7: In any connected graph, $k_V(G) \leq k_E(G) \leq \delta(G)$.

Proof: $k_E(G) \leq \delta(G)$. Let v_0 be any vertex in G with $\deg(v_0) = \delta(G)$, the minimum degree in G. If we remove the $\delta(G)$ edges that are incident to v_0, we will get a disconnected graph. Hence $k_E(G) \leq \delta(G)$.



$k_V(G) \leq k_E(G)$: Let $r = k_E(G)$. Then we can find r edges $e_1, e_2, \ldots, e_r$ whose removal will change G into a disconnected graph with two parts G_1 & G_2.



Now by judiciously removing one endpoint from each edge e_i (so as to leave at least one vertex in both G_1 & G_2) we will change G into a disconnected graph — or we will change G into K_1. Hence $k_V(G) \leq k_E(G)$.

Qu: Let G be a connected graph. When is it possible to remove exactly one directed edge from each of the pairs of oppositely oriented directed edges in G and still end up with a connected digraph?

§3 Weighted digraphs & graphs and distance

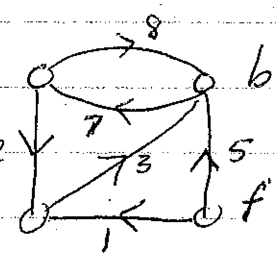
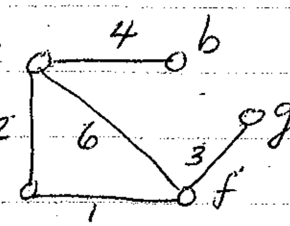
Def. A weighted digraph is an ordered pair $\langle G, w \rangle$ where G is a digraph and $w : E(G) \to \mathbb{R}^{\#}$ is a function called the weight-function, and $\mathbb{R}^{\#}$ = the set of non-negative real numbers.

Weighted graphs are defined in the same way except that when we view a graph as a digraph, we have to insist that each pair of oppositely oriented directed edges must have the same weight.

Def. The distance from a vertex u to a vertex v in a weighted digraph G is the function $d : V(G) \times V(G) \to [0, \infty]$ defined by

$$d(u,v) = \begin{cases} \text{length of the shortest directed path in } G \text{ from } u \text{ to } v, \\ \infty, \quad \text{if there is no directed path in } G \text{ from } u \text{ to } v. \end{cases}$$

The length of a directed path is the sum of the weights of the directed edges in the path.

Ex. 6 Let $G_1 =$  & $G_2 =$ 

In G_1, $d(a,a) = 0$, $d(a,b) = 5$, $d(b,a) = 7$, $d(a,f) = \infty$.
In G_2, $d(b,b) = 0$, $d(a,f) = 3$, $d(b,g) = 10$ and $d(g,b) = 10$.

Note that in any weighted graph, $d(u,v) = d(v,u)$.

Def. Let $\langle G, w \rangle$ be a weighted graph. We define the <u>eccentricity</u> of a vertex u in G by

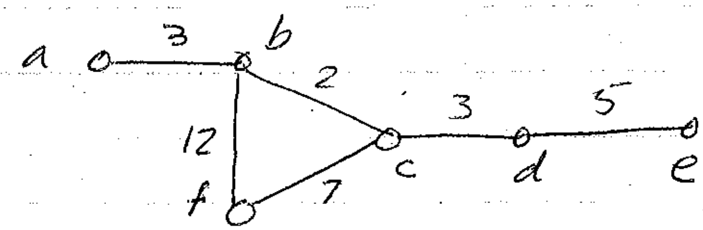
$$ecc(u) = \max \{d(u,v) : v \in V(G)\}.$$

We define the <u>diameter</u> & <u>radius</u> of G by

$$diam(G) = \max \{ecc(u) : u \in V(G)\}$$
$$rad(G) = \min \{ecc(u) : u \in V(G)\}.$$

A <u>center</u> of G is any vertex u_0 such that $ecc(u_0) = rad(G)$. A pair of <u>peripheral</u> vertices in G is any pair of vertices $\{u_0, v_0\}$ such that $d(u_0, v_0) = diam(G)$.

Ex. 7 Let $G =$



$d(u,v)$	a	b	c	d	e	f	$ecc(u) = \max\{d(u,v) : v \in V(G)\}$
a	0	3	5	8	13	12	13
b	3	0	2	5	10	9	10
c	5	2	0	3	8	7	8
d	8	5	3	0	5	10	10
e	13	10	8	5	0	15	15
f	12	9	7	10	15	0	15

$diam(G) = \max\{13, 10, 8, 10, 15, 15\} = 15$

$rad(G) = \min\{13, 10, 8, 10, 15, 15\} = 8$

The vertex c is the only center.

The only pair of peripheral vertices are $\{e, f\}$.

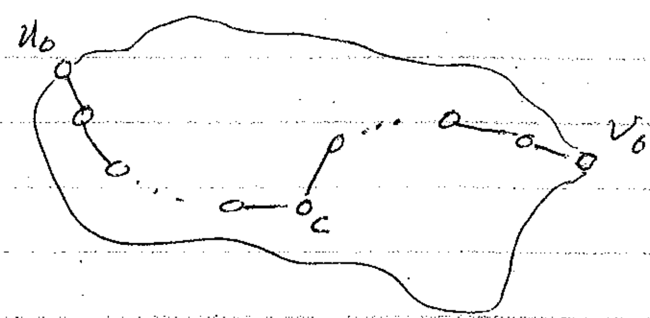
<u>Prop. 8</u>: Let $\langle G, w\rangle$ be a weighted graph. Then $rad(G) \leq diam(G) \leq 2 \cdot rad(G)$.

<u>Proof</u>: From the definitions of $rad(G)$ & $diam(G)$ we have

$$rad(G) = \min \{ecc(u) : u \in V(G)\}$$
$$\leq \max \{ecc(u) : u \in V(G)\} = diam(G).$$

Now let $diam(G) = D$. Then we can find a pair of peripheral vertices $\{u_0, v_0\}$ in G such that $d(u_0, v_0) = diam(G)$. Let c be a center of G. Then

$$D = d(u_0, v_0)$$
$$\leq d(u_0, c) + d(c, v_0)$$
$$\leq ecc(c) + ecc(u)$$
$$= 2\, ecc(c) = 2 \cdot rad(G).$$



<u>Remark</u>: We have used the triangular inequality in the proof of Prop. 5. This says that if u, v and w are any three vertices in a weighted graph $\langle G, w\rangle$, then $d(u,w) \leq d(u,v) + d(v,w)$.

We can see that this result is true because

$d(u,w)$ = length of the shortest path from u to w
$\leq$ length of the shortest path from u to w extracted from the walk that goes along the shortest path from u to v & the shortest path from v to w.
$\leq$ (length of the shortest path from u to v) + (length of the shortest path from v to w)
$= d(u,v) + d(v,w)$.

Qu: How can we find the distance from the vertex u to the vertex v in a weighted digraph $\langle G, w\rangle$?

Algorithm 1 (Dijkstra's Distance Algorithm)

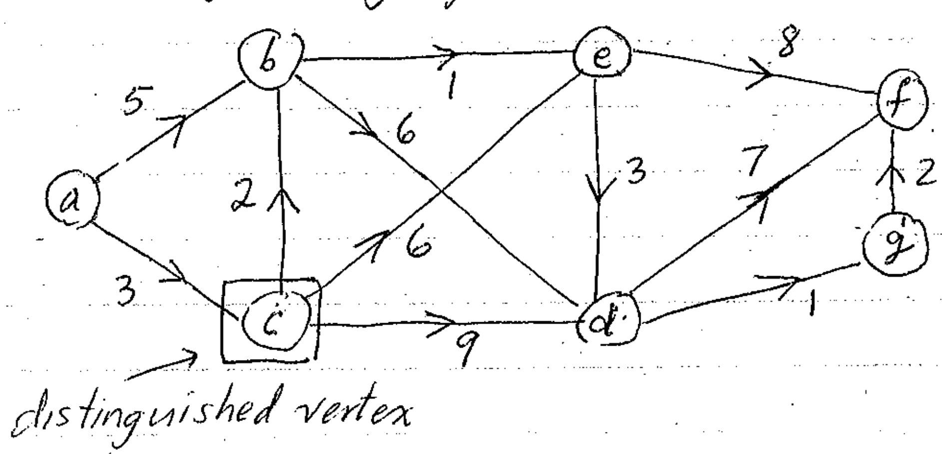
INPUT: A weighted digraph G with distinguished vertex x.

OUTPUT: The distances $d(x,v)$ from x to each vertex in G.

METHOD: Label each vertex v with the length of the shortest path, with $\leq i$ directed edges, from x to v.

1. Let $i \leftarrow 0$, $T \leftarrow V$, $L(x) \leftarrow 0$ & $L(v) \leftarrow \infty$ for each $v \in V(G) - \{x\}$
2. If $T = \emptyset$, STOP; else choose a vertex $v_0 \in T$ with minimum label $L(v_0)$.
3. For every directed edge $\vec{e} = \overrightarrow{v_0 u}$ from v_0 to another vertex $u \in T$, let
   $$L(u) \leftarrow \min\{L(u),\ L(v_0) + w(\vec{e})\}.$$
4. Let $T \leftarrow T - \{v_0\}$ and $i \leftarrow i+1$. Then go to step 2.

Ex. 8 Find the distances from c to each of the vertices in the weighted graph below.



v

$L(a)$	$L(b)$	$L(c)$	$L(d)$	$L(e)$	$L(f)$	$L(g)$	T	i	v_0
∞	∞	$\underline{0}$	∞	∞	∞	∞	$\{a,b,c,d,e,f,g\}$	0	c
∞	$\underline{2}$	'	9	6	∞	∞	$\{a,b,d,e,f,g\}$	1	b
∞	'	'	8	$\underline{3}$	∞	∞	$\{a,d,e,f,g\}$	2	e
∞	'	'	$\underline{6}$	'	11	∞	$\{a,d,f,g\}$	3	d
∞	'	'	'	'	11	$\underline{7}$	$\{a,f,g\}$	4	g
∞	'	'	'	'	$\underline{9}$	'	$\{a,f\}$	5	f
$\underline{\infty}$	'	'	'	'	'	'	$\{a\}$	6	a
'	'	'	'	'	'	'	$\emptyset$	7	
$d(c,v) = \infty$	2	0	6	3	9	7			