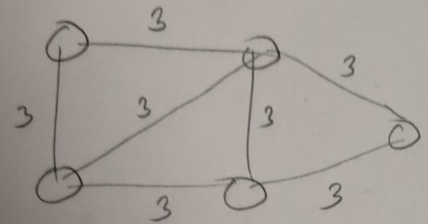
Mark Gomeng

1) a- Queue structure. Since visits all immediate neighbors first then moves to next level and so on.

b-



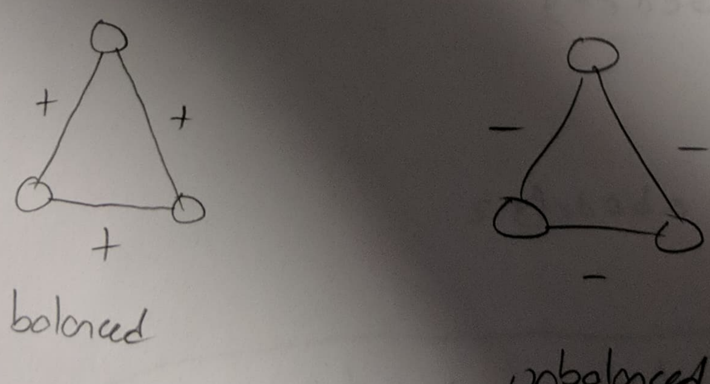
c- Common properties are small average path length (< 6 degrees of separation), the power law distribution and a high clustering coefficient

d- The diameter of a graph is the length of the longest shortest path between any two nodes in the graph.

e- with n as # of nodes

min edges spanning tree = n - 1 so m - (n - 1) edges must be removed

f-



← enemy of my enemy is my friend

2)

a) a → c = 3 		a, c visited

a → b = 4

b) a → b = 4 		a, b, c, visited

a → c → b = 5

a, c, d = 6

a, c, e = 9

c) a, c, b, d = 10

a, b, d = 9 		a, b, c, d

a, c, d = 6

a, c, e = 9

d) a, c, e, = 9

a, c, d, e = 7 		a b c d e

a, c, d, f = 11

e) a, c, d, f = 11 		a b c d e f

a, c, d, e, g = 12

f) a, c, d, f, g = 13

a, c, d, e, g = 12 		a b c d e f g

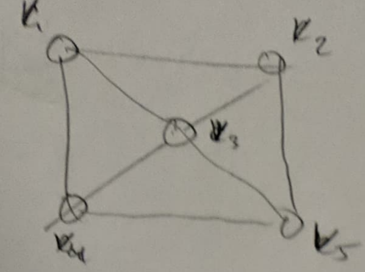
g) a, c, d, f, z = 18

a, c, d, e, g, z = 16 		a b c d e f g z

Thus shortest path between a and z

is a → c → d → e → g → z which costs 16.

3) a)



$C_b(v_1) = 2 \times (\underbrace{\tfrac{1}{2}}_{v_4,\ v_2}) = 1$

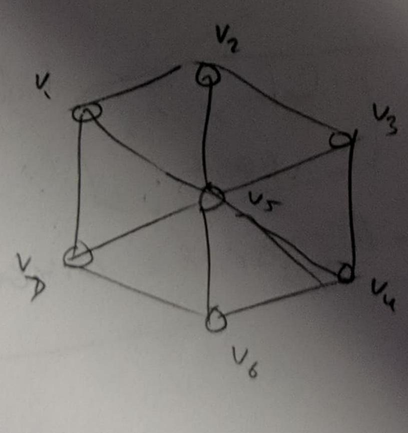
$C_b(v_2) = 2 \times (\underbrace{\tfrac{1}{2}}_{v_5,\ v_1}) = 1$

$C_b(v_3) = 2 \times (\underbrace{\tfrac{1}{2}}_{v_4\ v_2} + \underbrace{\tfrac{1}{2}}_{v_5\ v_1}) = 2$

$C_b(v_4) = 2 \times (\underbrace{\tfrac{1}{2}}_{v_1\ v_5}) = 1$

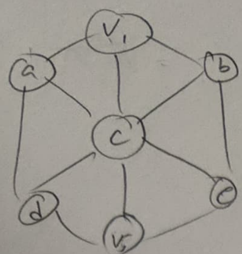
$C_b(v_5) = 2 \times (\underbrace{\tfrac{1}{2}}_{v_4\ v_2}) = 1$

b) for non central vertex in a wheel graph when $n \geq 6$, is $\underline{\text{one}}$ because the two other non-central nodes are the only one using the node which is just 1/2 since central is also used.



$C_b(v_1) = 2 \times (\underbrace{\tfrac{1}{2}}_{v_7\ v_2} + 0) = 1$

4)



$N(v_i) = a, b, c$

$N(v_j) = d, c, e$

$N(v_i) \cap N(v_j) = \{c\}$

$N(v_i) \cup N(v_j) = \{a, b, c, d, e\}$

$\sigma_{Jaccard}(v_i, v_j) = \dfrac{|\{c\}|}{|\{a, b, c, d, e\}|} = \dfrac{1}{5}$

5) $d_{in}(A) = 2$

$d_{in}(B) = 0$

$d_{in}(C) = 3$

$d_{in}(E) = 0$

$d_{in}(D) = 2$

X	A	B	C	D	E
$P(v_x)$	2/7	0	3/7	2/7	0

6) a) status = 20 Senior, 50 Junior, [20, 50]

$H(\text{status}) = -\left(\frac{2}{7}\right)\log\left(\frac{2}{7}\right) - \left(\frac{5}{7}\right)\log\left(\frac{5}{7}\right) = 0{,}86$

b) E_{salary} low $= [0, 20)$, $E = 0$

med $= [0, 20]$, $E = 0$

high $= [20, 10]$, $E = 0{,}918$

$\text{Gain} = 0{,}86 - \left(\frac{20}{70}\right)(0) - \left(\frac{20}{70}\right)0 - \left(\frac{30}{70}\right)0{,}918 = 0{,}466$

Salary as root

highest gain

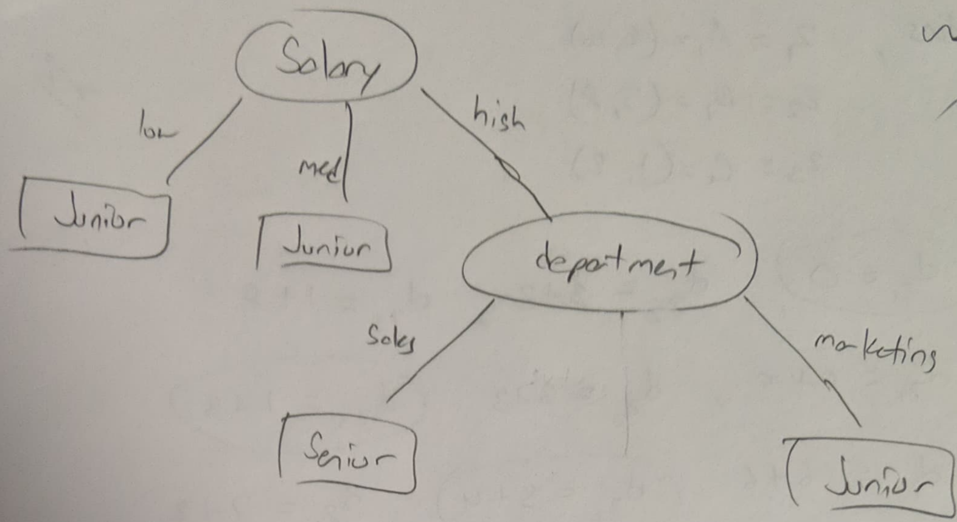
E_{dept} sales $= [15, 20]$, $E = 0{,}98$

systems $= [0, 20]$, $E = 0$

marketing $= [5, 10]$, $E = 0{,}92$

$\text{Gain} = 0{,}86 - \left(\frac{35}{70}\right)0{,}98 - \left(\frac{20}{70}\right)0 - \left(\frac{15}{70}\right)0{,}92 = 0{,}17$

6) c)



7) dept = Sales , salary = high

$P(\text{status} = \text{Junior} \mid \text{sales, high}) = P(\text{Sales} \mid \text{Junior}) \times P(\text{high} \mid \text{Junior}) \times \dfrac{P(\text{Junior})}{P(\text{sales, high})}$

$= \dfrac{20}{50} \times \dfrac{10}{50} \times \dfrac{50/70}{P(\text{sales, high})} = \dfrac{20}{50} \times \dfrac{10}{50} \times \dfrac{50}{70} \times \dfrac{1}{P(\text{sales, high})}$

$= 0{,}057 \; P(\text{sales, high})$

$P(\text{status} = \text{Senior} \mid \text{sales, high}) = P(\text{sales} \mid \text{Senior}) \times \ldots$

$= \dfrac{15}{20} \times \dfrac{20}{20} \times \dfrac{(20/70)}{P(\text{sales, high})}$

$= 0{,}21 \; P(\text{sales, high})$

Status = Senior since P(...) is greater

8) first clusters, $z_1 = A_1 = (2,10)$

$z_2 = B_1 = (5,8)$

$z_3 = C_1 = (1,2)$

$A_1 = (2,10)$, $d_{z_1} = 0$, $d_{z_2} = 3+2$, $d_{z_3} = 1+8$

$A_2 = (2,5)$, $d_{z_1} = 0+5$, $d_{z_2} = 3+3$ $d_{z_3} = 1+3$

$A_3 = (8,4)$, $d_{z_1} = 6+6$ $d_{z_2} = 3+4$ $d_{z_3} = 7+2$

$B_1 = (5,8)$ $d_{z_1} = 3+2$ $d_{z_2} = 0$ $d_{z_3} = 4+6$

$B_2 = (7,5)$ $d_{z_1} = 5+5$ $d_{z_2} = 2+3$ $d_{z_3} = 6+3$

$B_3 = (6,4)$ $d_{z_1} = 4+6$ $d_{z_2} = 1+4$ $d_{z_3} = 5+2$

$C_1 = (1,2)$ $d_{z_1} = 1+8$ $d_{z_2} = 4+6$ $d_{z_3} = 0$

$C_2 = (4,9)$ $d_{z_1} = 2+1$ $d_{z_2} = 1+1$ $d_{z_3} = 3+7$

after one round

z_1 cluster has $\{A_1\}$

z_2 cluster has $\{A_3, B_1, B_2, B_3, C_2\}$

z_3 cluster has $\{A_2, C_1\}$

z_1 cluster center $= A_1 = (2,10)$

z_2 cluster center $= (30/5, 30/5) = (6,6)$

z_3 cluster center $= (3/2, 7/2) = (1.5, 3.5)$

9) a) False

b) True

c) True

d) False

e) True

f) False

g) False

h) True

i) False

j) True