Question 1

Build inverted index

101	$I\text{-}S_0$	201	$I\text{-}S_1$	301	$I\text{-}S_2$
cat pat mat sat cat cat		pat mat sat pat mat cat		sat mat cat pat pat mat	

M.D

cat	pat	mat	sat	cat	pat	mat	sat	cat	pat	mat	sat	cat	pat	pat
2	1	1	1	1	2	2	1	1	1	2	1	1	1	1

output

((cat, 101), 2)	((pat, 201), 2)	((sat, 301), 1)
((pat, 101), 1) M_0	((mat, 201), 2) M_1	((mat, 301), 2) M_2
((mat, 101), 1)	((sat, 201), 1)	((cat, 301), 1)
((sat, 101), 1)	((cat, 201), 1)	((pat, 301), 1)
((cat, 101), 1)		((pat, 301), 1)

S.S

Assume after partitioning

cat, mat, sat ↓ — R_0 Input

cat, pat ↓ — R_1 Input

R_0 Input	R_1 Input
→ ((cat, 101), [2])	→ ((cat, 101), [1])
((cat, 301), [1])	((cat, 201), [1])
→ ((mat, 101), [1])	→ ((pat, 101), [1])
((mat, 201), [2])	((pat, 201), [2])
((mat, 301), [2])	((pat, 301), [1])
→ ((sat, 101), [1])	
((sat, 201), [1])	
((sat, 301), [1])	

R_0 output	R_1 output
(cat, [(101,2), (301,1)])	(cat, [(101,1), (201,1)])
(mat, [(101,1), (201,2), (301,2)])	(pat, [(101,1), (201,2), (301,1)])
(sat, [(101,1), (201,1), (301,1)])	

```
class Mapper
    method initialize ()
        tprev ← ϕ ,    P ← new PostingsList ()

    method reduce ( pair (t, n) , integer [f] )
        if (t != tprev && tprev ≠ ϕ)
            Emit (term tprev, postings P)
            P.reset ()
        P.add (new pair (n, f))
        t ← tprev

    method close ()
        Emit (term tprev , postings P)
```

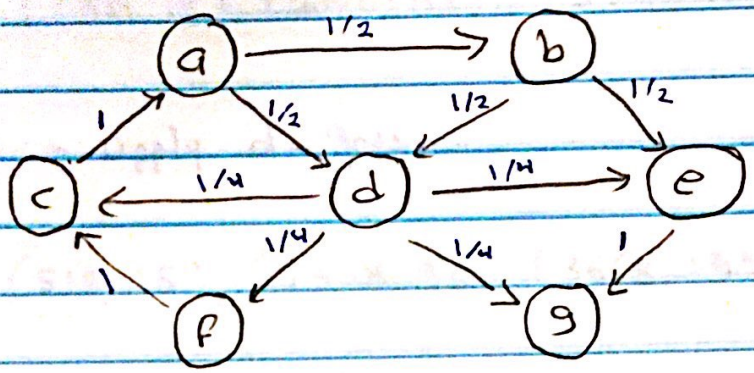
Inverted indexing pseudo code

```
Class Mapper
    method initialize ( )
        H ← new AssociativeArray ( )
        id ← 0

    method map (docid n, doc d)
        id ← n
        for all term t in record r do:
            H{t} ← H{t} + 1

    method close ( )
        for all term t in H do:
            Emit ((t, id), H{t})
                                  └> term frequency
```

Question 2



step	a	b	c	d	e	f	g
0	1/7	1/7	1/7	1/7	1/7	1/7	1/7
1	0.1428	0.0714	0.1785	0.1428	0.1071	0.0357	0.1785
2	0.1785	0.0714	0.0714	0.1071	0.0714	0.0357	0.1428
3	0.0714	0.0893	0.0625	0.125	0.0625	0.0268	0.0982

$\frac{1}{7} + \frac{1}{4} \times \frac{1}{7} = \frac{4}{28} + \frac{1}{28}$

Room 105 and this room

c = 0.25 * 0.1428 + 1 * 0.0357 =

d = 0.5 * a + 0.5 * b = 0.5

e = 0.5 * b + 0.25 * d = 0.0357 + 0.0357 =

f = 0.25 * d =

g = 1 * e + 0.25 * d = 0.1071 + 0.0357

Question 3

[(512, 15), (2080, 93), (5748, 195), (7080, 255)]

Step 0 → apply d-gaps

[(512, 15), (1568, 93), (3668, 195), (1332, 255)]

→ 512 15 1568 93 3668 195 1332 255

Step 1

512 = 100|00000000
15 = 1111
1568 = 1100|0100000
93 = 1011101
3668 = 111001|101100
195 = 11000011
1332 = 1010|0110100
255 = 11111111

00000100	10000000	10001111	
00001100	10100000	11011101	
00011100	11010100	00000001	11000011
00001010	11111111		