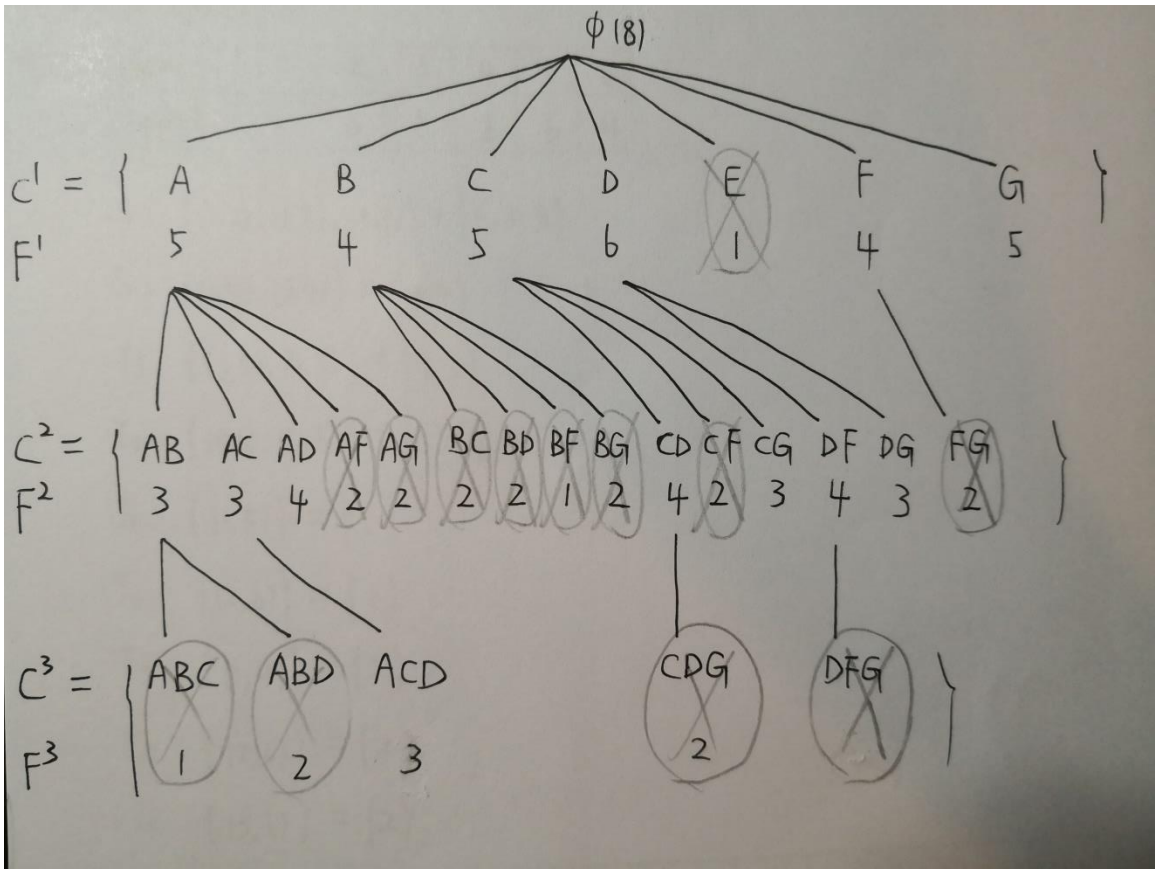
Name: Chaoyang Zhu UIN: 930001840

Q1.

(a)



So the frequent itemsets are A(5), B(4), C(5), D(6), F(4), G(5), AB(3), AC(3), AD(4), CD(4), CG(3), DF(4), DG(3) and ACD(3).

Q4.

sup(ABE)=2

sup(AB)=3 sup(AE)=2 sup(BE)=4 sup(A)=4 sup(B)=5 sup(E)=4

conf(AB->E)= $\frac{2}{3}$ interest(AB->E)= $\frac{2}{3}-\frac{4}{6}=0$

conf(AE->B)= $\frac{2}{2}=1$ interest(AE->B)= $1-\frac{5}{6}=\frac{1}{6}$

conf(BE->A)= $\frac{2}{4}=\frac{1}{2}$ interest(BE->A)= $\frac{1}{2}-\frac{4}{6}=-\frac{1}{6}$

conf(A->BE)= $\frac{2}{4}=\frac{1}{2}$ interest(A->BE)= $\frac{1}{2}-\frac{4}{6}=-\frac{1}{6}$

conf(B->AE)= $\frac{2}{5}$ interest(B->AE)= $\frac{2}{5}-\frac{2}{6}=\frac{1}{15}$

conf(E->AB)= $\frac{2}{4}=\frac{1}{2}$ interest(E->AB)= $\frac{1}{2}-\frac{3}{6}=0$

Exercise 6.3.1

(a)

Item	1	2	3	4	5	6
Support	4	6	8	8	6	4

T1: {(1, 2), (1, 3), (2, 3)} = {2, 3, 3}

T2: {(2, 4), (3, 4)} = {4, 4}

T3: {(3, 5), (4, 5)} = {4, 3}

T4: {(4, 6), (5, 6)} = {3, 2}

T5: {(1, 5)} = {1}

T6: {(2, 6)} = {1}

T7: {(1, 4)} = {2}

T8: {(2, 5)} = {2}

T9: {(3, 6)} = {2}

Item pair	(1,2)	(1,3)	(2,3)	(2,4)	(3,4)	(3,5)	(4,5)	(4,6)	(5,6)	(1,5)	(2,6)	(1,4)	(2,5)	(3,6)
Support	2	3	3	4	4	4	3	3	2	1	1	2	2	2

(b)

Item pair	(1,2)	(1,3)	(2,3)	(2,4)	(3,4)	(3,5)	(4,5)	(4,6)	(5,6)	(1,5)	(2,6)	(1,4)	(2,5)	(3,6)
Bucket	2	3	6	8	1	4	9	2	8	5	1	4	10	7

(c)

Bucket	Content	Count
0		0
1	(3, 4), (2, 6)	5
2	(1, 2), (4, 6)	5
3	(1, 3)	3
4	(3, 5), (1, 4)	6
5	(1, 5)	1
6	(2, 3)	3
7	(3, 6)	2
8	(2, 4), (5, 6)	6
9	(4, 5)	3
10	(2, 5)	2

Since the support threshold is 4, bucket 1, 2, 4, 8 are frequent.

(d)

According to the table above, the pairs counted on the second pass of the PCY Algorithm are (3, 4), (2, 6), (1, 2), (4, 6), (3, 5), (1, 4), (2, 4), (5, 6).