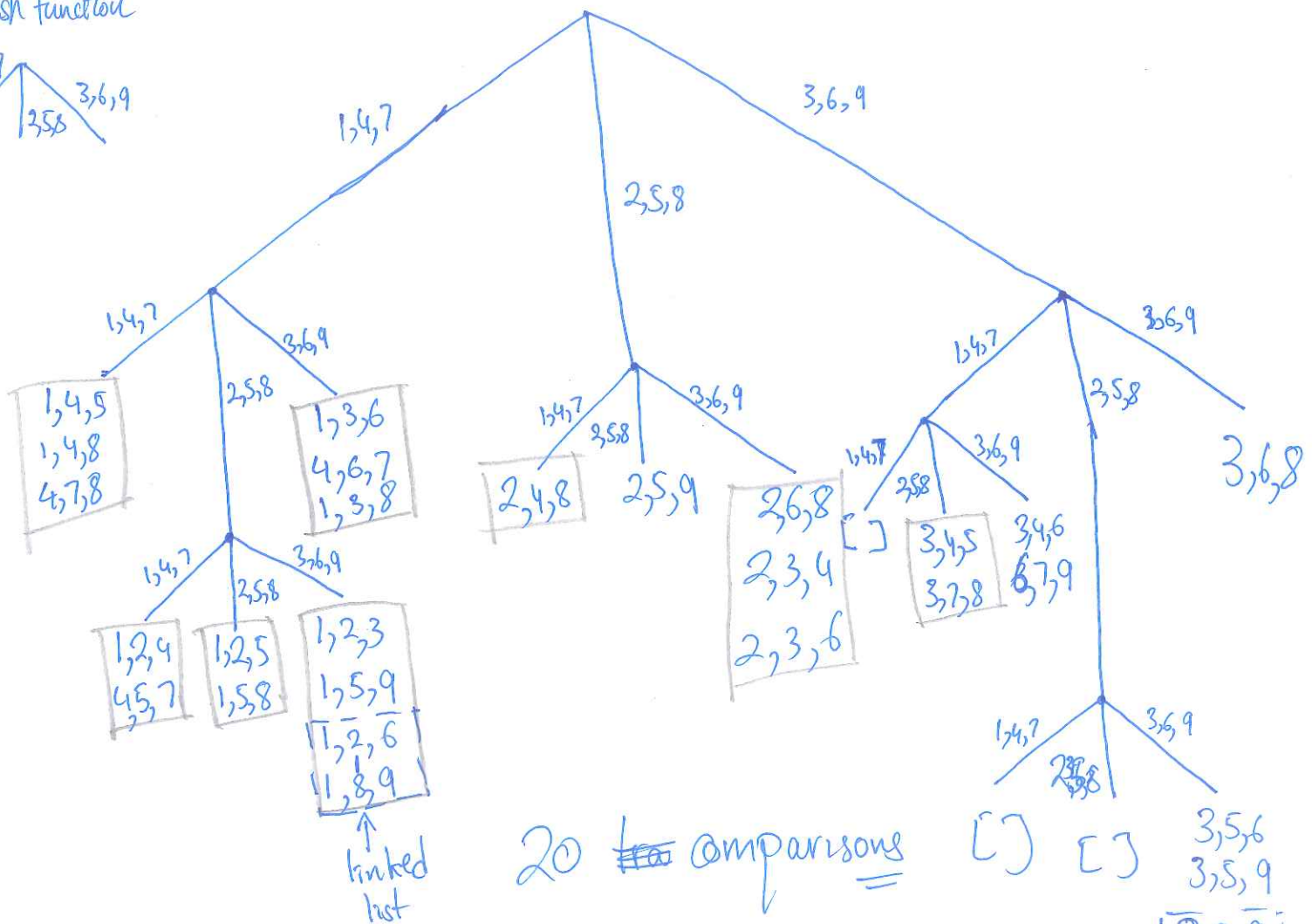Assignment 1

Q1

b) After sorting and deleting duplicates, there are 28 candidates

candidate_set = $\{1,2,3\}$, $\{1,4,5\}$, $\{1,2,4\}$, $\{1,2,5\}$, $\{1,5,9\}$, $\{1,3,6\}$, $\{1,4,8\}$
$\{1,2,6\}$, $\{1,8,9\}$, $\{1,5,8\}$, $\{1,3,8\}$

$\{2,3,4\}$, $\{2,5,9\}$, $\{2,3,6\}$, $\{2,4,8\}$, $\{2,6,8\}$

$\{3,4,5\}$, $\{3,5,6\}$, $\{3,8,9\}$, $\{3,5,9\}$, $\{3,4,6\}$, $\{3,6,8\}$, $\{3,7,8\}$

$\{4,5,7\}$, $\{4,7,8\}$, $\{4,6,7\}$,
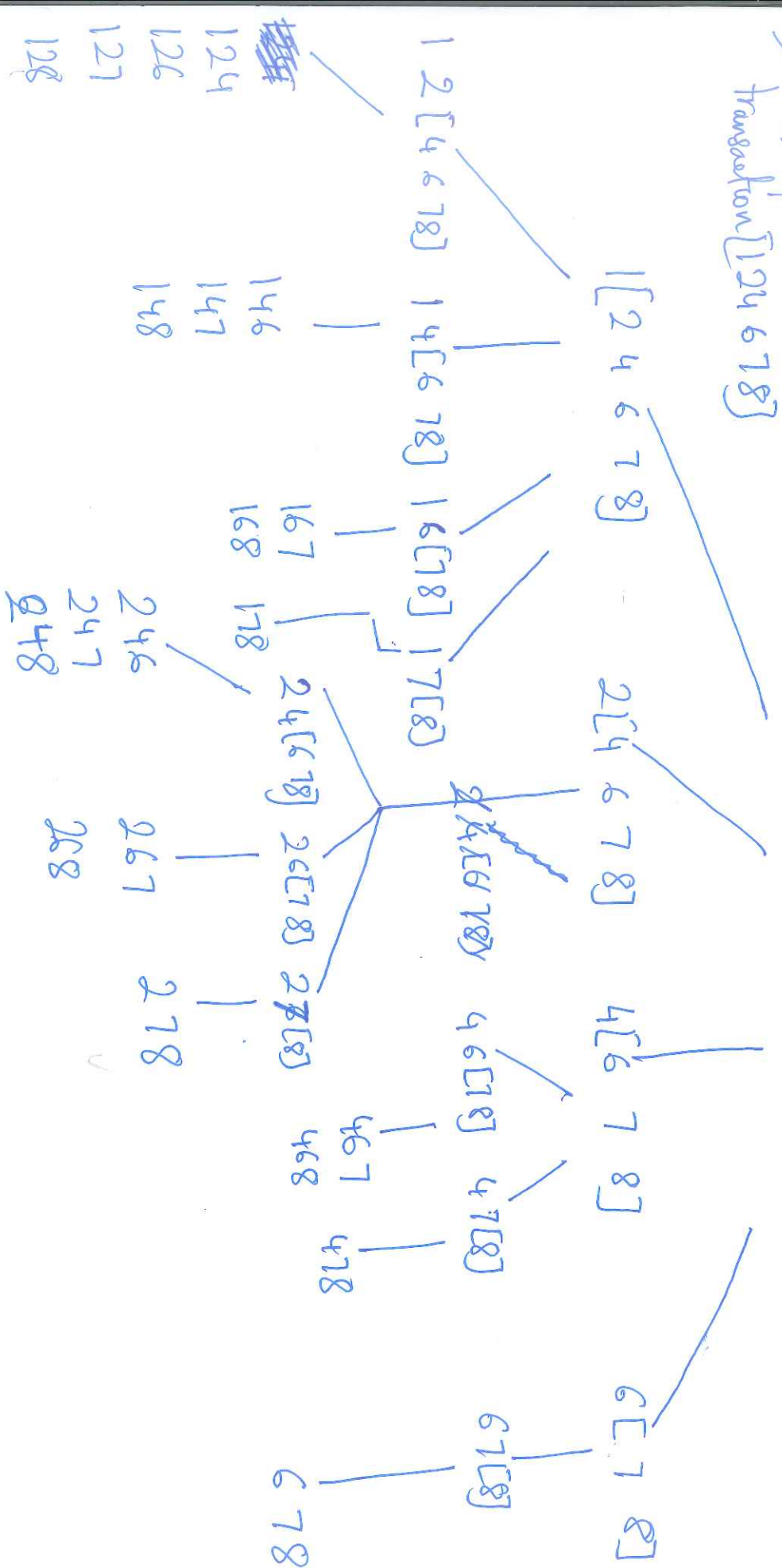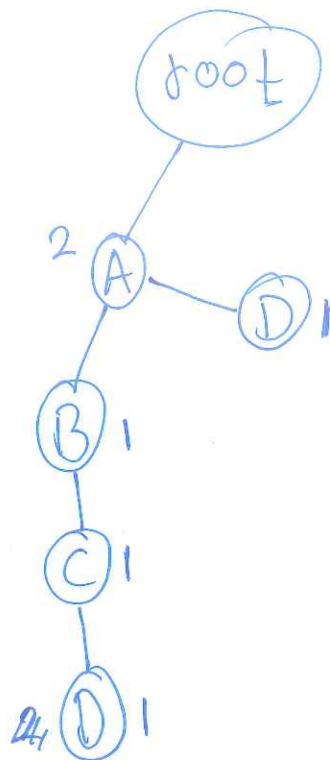
$\{6,8,9\}$, $\{6,7,9\}$

Hash Function

1,4,7   3,6,9
    3,5,8

1,4,5
1,4,8
4,7,8

1,3,6
4,6,7
1,3,8

1,2,4
4,5,7

1,2,5
1,5,8

1,2,3
1,5,9
1,2,6
1,8,9

2,4,8

2,5,9

2,6,8
2,3,4
2,3,6

[ ]

3,4,5
3,7,8

3,4,6
6,7,9

3,6,8

[ ]   [ ]   3,5,6
            3,5,9
            3,8,9
            6,8,9

linked list

20 comparisons

linked list

b) All possible subsets for
transaction [1 2 4 6 7 8]

1 2 4 6 7 8

1[2 4 6 7 8]   2[4 6 7 8]   4[6 7 8]   6[7 8]

1 2[4 6 7 8]   1 4[6 7 8]   1 6[7 8]   1 7[8]   2 4[6 7 8]   2 6[7 8]   2 7[8]   4 6[7 8]   4 7[8]   6 7 8

124   146   167   178   246   267   278   467   478   678
126   147   168   247   268
127   148   248
128

2)b)

root

2 A —— D 1

B 1

C 1

2↓ D 1

Support threshold = 2
Sup(A)=2   Sup(C)=1
Sup(B)=1   Sup(D)=2

As ~~B~~ sup(B) < minsup and sup(c) <minsup
the resulting conditional fp tree
for D is:

root

2 A —— D 1

D 1

~~{(A:1)~~
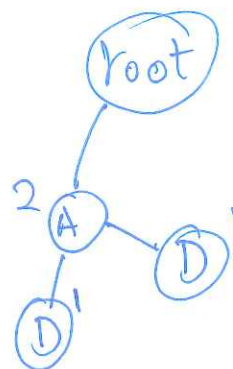
result:
{(A:1, D:1), (A:1, D:1)}

Hash Tree Output for Q1 (A)

[[[[1, 4, 5], [1, 4, 8], [4, 7, 8]], [[[1, 2, 4], [4, 5, 7]], [[1, 2, 5], [1, 5, 8]], [[1, 2, 3], [1, 5, 9], [1, 2, 6], [1, 8, 9]]], [[1, 3, 6], [4, 6, 7], [1, 3, 8]]], [[[2, 4, 8]], [[2, 5, 9]], [[2, 3, 4], [2, 3, 6], [2, 6, 8]]], [[[], [[3, 4, 5], [3, 7, 8]], [[3, 4, 6], [6, 7, 9]]], [[], [], [[3, 5, 6], [3, 5, 9], [3, 8, 9], [6, 8, 9]]], [[3, 6, 8]]]]