Assignment4

Ex2

The result for chess ,mushroom ,T1014D100k in A-Priori are saved in four different text.file

The support is set to 1% of itemset being tested,. For example, if the size of sample is 10000, so the support is 100:

$$Support\ For\ Samples = Support\ \times Size\ of\ samples$$

When the A-Priori is applied, with the size of samples expanding, the more frequent items can be found

The SON Algorithm is parallel computing, splitting the whole itemset to different part, find the frequent in the sub-parts and sum up the frequent items from different part. This is the perfect and non-error algorithm to find the frequent items. Comparing with the sample-randomizer algorithm, the SON algorithm finds out all the frequent items

A- priori is applied in Chess.dat, mushroom.dat

EX3

Part1. Greedy algorithm is make the best choice or matching for the current state without thinking the overall optimization

For example , in the map-navigation design , the greedy algorithm is finding the location which is the closet to the current location without considering the shortest route in general from the start to end location

Balance algorithm is the greedy algorithm considering for the future optimized .For example, the client has limited budget, it can choose the advertiser with the largest remaining budget . If the choice for the largest remaining budget is more than 1 , it can select the advertiser with greatest benefit

Competitive Ratio is, what is the difference betweenworst of all possible inputs?

Worst / optimal

Part2. We have these arrangements

ABC: 6
A xx
B yy
C zz

ACB: 4
A xx
C None
B yy

BAC: 4
B xx
A None
C yy

BCA: 4
B xx
C yy
A None

2:
And the optimal is 6
So, the half is 3

∴ the sequence may like this xxxzyy

ABC: 5
A xx
B x
C zy

ACB: 6
A xx
C xz
B yy

BAC: 5
B xx
A x
C zy