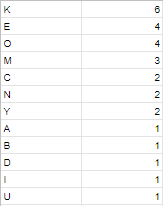
**1.) [40] Trace the FP-growth algorithm to find frequent itemsets. For this question, supmin = 60%.**

**1.a) Find the F-list**

**\*sup(min) = (6\*60)/100 =3.6 = I’ll use 4**



F-list = K – E – O *[– M – C – N – Y – A – B – D – I – U] \*pruned*

**1.b) Draw the FP-tree.**

Ordered frequent item sets (\*sup(min) = (6\*60)/100 =3.6 = 4):

T1: {K, E, O, M, N} => {K, E, O}

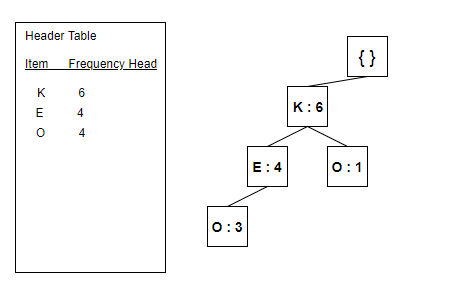
T2: {K, E, O, N, Y, D} => {K, E, O}

T3: {K, E, M, A} => {K, E}

T4: {K, M, C, Y, U} => {K}

T5: {K, E, O, C, I} => {K, E, O}

T6: {K, O, B} => {K, O}



**1.c & 1 .d)**

**Projected DB** *Conditional Pattern Base* *Frequent Item sets*

-----------------------------------------------------------------------------------------------------------

**O** {{K,E, : 3},{K : 1}} {K, 0: 4}, {E, O: 3}, {K, E, O: 3}

**E** {{K:4}} {K, E: 4}

**K** {nothing} {nothing}

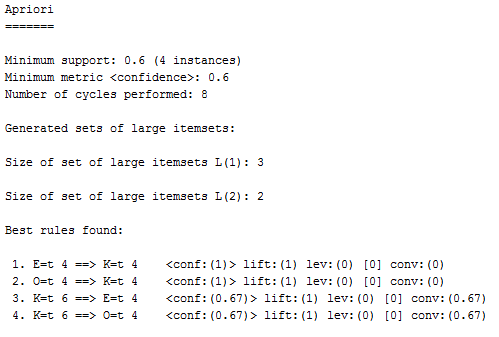
**2. [60] Convert the transaction data table into an hwk04.arff file, in which each item is a column with either a value t or ? (that is, with a type { t}). For example:**

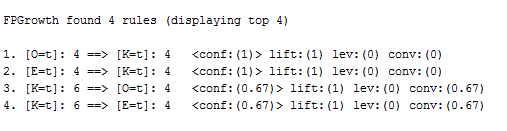
**2.1) Using support(min) = 60% and confidence = 60%, the input being :**





We get:





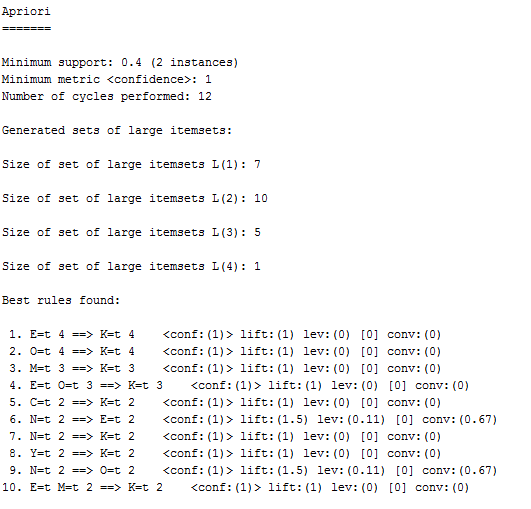
*\*Full report can be found in Apriori6060.txt and FPGrowth6060.txt*

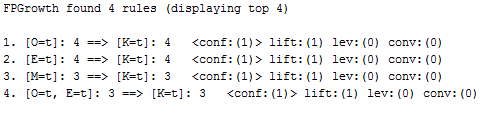
**2.2) Using support(min) = 40% and confidence = 100%, the input being :**





We get:



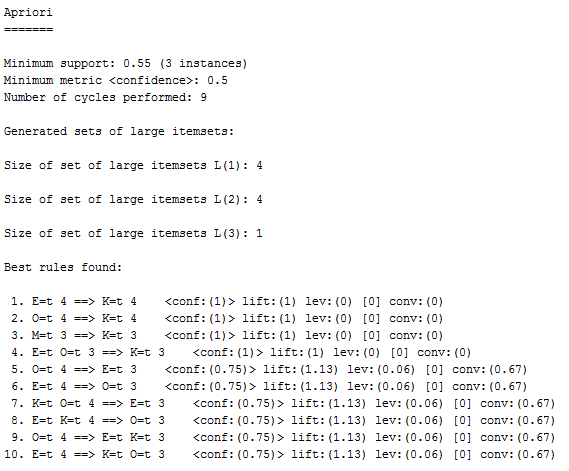


*\*Full report can be found in Apriori40100.txt and FPGrowth40100.txt*

**2.3) Using support(min) = 20% and confidence = 50%, the input being :**



For Apriori we get:



*\*Full report can be found in Apriori2050.txt*