

To prove the correctness, we need to prove the Apriori algorithm is sound and complete.

It's sound since each itemset in F_i has been checked for frequency.

We can prove the alg. is complete by induction:

1. We have all 1-itemsets are in C_1 and therefore all frequent 1-itemsets are in F_1 .
2. Assume all frequent n -itemsets are in F_n for $n \in N$
3. $n \rightarrow n + 1$
let X be a frequent $(n + 1)$ -itemsets
all n -itemsets of X are in F_n and frequent, so $X \in C_{n+1}$, $X \in F_{n+1}$

To prove the Irredundancy:

Let X be a frequent n -itemset

Then $X \notin \{F_1, \dots, F_{n-1}, F_{n+1}, \dots\}$, but $X \in F_n$,

Assume X appears twice in F_n

So X will appear twice in $C_n \Rightarrow$ some $(n-1)$ -subset appears twice in F_{n-1}

In this way, some 1-itemset will appear twice in F_1 and C_1

It is contradictory with $C_1 = I$ (which is a set each item only appear once).

According to above proof we can say the Apriori algorithm correctly and irredundantly generates all frequent itemsets.