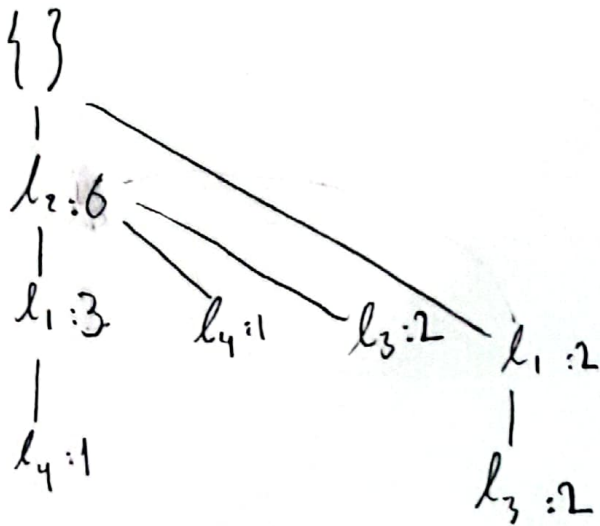


TID	ordered frequent itemset
T100	$l_2 - l_1$
T200	$l_2 - l_4$
T300	$l_2 - l_3$
T400	$l_2 - l_4 - l_1$
T500	$l_1 - l_3$
T600	$l_2 - l_3$
T700	$l_1 - l_3$
T800	$l_2 - l_1$

: F12-tree (A)

$l_1: 5, l_2: 6, l_3: 4, l_4: 2$  \*

F-list =  $l_2 - l_1 - l_3 - l_4$



: F12-Tree \*

\* Conditional Pattern base :

item	Conditional Pattern base	Con. frequent pattern tree	Frequent Patterns
$l_2$			
$l_1$	$l_2: 3$	$l_2: 3$	$\{l_2, l_1\}: 3$
$l_3$	$l_2: 2, l_1: 2$	$l_2: 2, l_1: 2$	$\{l_2, l_3\}: 2, \{l_1, l_3\}: 2$
$l_4$	$l_2: 1, l_2, l_1: 1$	$l_2: 2$	$\{l_2, l_4\}: 2$

Frequent item sets:  $\{l_2, l_1\} - \{l_2, l_3\} - \{l_1, l_3\} - \{l_1, l_2, l_3\} - \{l_2, l_4\}$

(3)

$K=1$ ,  $\text{min-sup}=3$

Item	TID
$l_1$	$T_{100} - T_{400} - T_{500} - T_{700} - T_{800} \checkmark$
$l_2$	$T_{100} - T_{200} - T_{300} - T_{400} - T_{600} - T_{800} \checkmark$
$l_3$	$T_{300} - T_{500} - T_{600} - T_{700} \checkmark$
$l_4$	$T_{200} - T_{400} \times$

$K=1$

Items	TID
$l_1 l_2$	$T_{100}, T_{400}, T_{800} \checkmark$
$l_1 l_3$	$T_{500} - T_{700} \times$
$l_2 l_3$	$T_{300} - T_{600} \times$

$K=2$

frequent Itemsets =  $\{l_1, l_2, l_3, l_1 l_2\}$

© الگوریتم  $FP_{Growth}$  یک الگوریتم DFS است عکس الگوریتم  $ECLAT$  دیگر روش دارد  
 BFS دارد.  $FP$  حافظه کمتر نسبت به  $ECLAT$  نیاز دارد و همچنین ممکن است  
 طولانی تر از  $FP$  به جواب برسد.



$$\text{Support}(\overset{A}{\text{hot dogs}}, \overset{B}{\text{hamb}}) = \frac{2000}{5000} = 40\% \quad -2$$

$$\text{Conf}(\overset{A}{\text{hot dogs}} \rightarrow \overset{B}{\text{hamb}}) = \frac{\text{sup}(\text{hot, hamb})}{\text{sup}(\text{hamb})} = \frac{2000}{3000} = \frac{2}{3} = 66\%$$

(A) این قانون صحیح است، زیرا  $\text{sup}$  بیشتر از 25٪، و  $\text{Conf}$  بیشتر از 50٪ است.

$$\text{Lift}(A, B) = \frac{C(A \rightarrow B)}{S(B)} = \frac{S(A \cup B)}{S(A) \times S(B)} = \frac{0.4}{0.6 \times 0.8} = \frac{100}{12} \quad (B)$$

$$S(A) = 0.6 \quad S(B) = 0.8 \quad \approx 8$$

Expected: چون  $\text{Lift}(A, B) > 1$  در نتیجه همبستگی مثبت بین این دو به نظر می آید.

	hot day	hot
HAM	$\frac{2500 \times 3000}{5000} = 1500$	$\frac{2000 \times 2500}{25000} = 2000$
$\neg \text{HAM}$	$\frac{3000 \times 2500}{25000} = 3000$	$\frac{2000 \times 2500}{25000} = 2000$

$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}$$

$$\chi^2 = \frac{(2000 - 1500)^2}{1500} + \frac{(1000 - 500)^2}{1000} + \frac{(1500 - 1500)^2}{1500} + \frac{(1500 - 1000)^2}{1000}$$

$$= \frac{250000}{1500} + \frac{250000}{1000} + 0 + \frac{250000}{1000} = 500 + 166.7$$

$$= 666.7$$

$$\chi^2 < E(\text{hot, HAM}) \Rightarrow$$

$$666.7 < 1500$$

→ negatively Correlated!