

### Exercise 1

Let  $\mathcal{D}$  a transaction dataset with the following horizontal representation  $\mathcal{H}_{\mathcal{D}}$  :

trans.	Items					
$t_1$		$B$	$C$	$D$		
$t_2$	$A$	$B$	$C$		$E$	
$t_3$	$A$	$B$	$C$	$D$		$F$
$t_4$				$D$	$E$	
$t_5$	$A$	$B$				
$t_6$	$A$		$C$		$E$	$F$
$t_7$	$A$	$B$			$E$	$F$
$t_8$				$D$		$F$
$t_9$			$C$		$E$	
$t_{10}$	$A$	$B$				$F$

- Execute the LCM algorithm on  $\mathcal{D}$  with a minsup  $\theta = 3$ .
- LCM is linear on what ? explain how.

### Exercise 2

Let us take the following query :

$$Q : frequent(P) \wedge closed(P) \wedge maxSize_{ub}(P)$$

with the two interpretations :

1. Mine all frequent closed itemsets which have the additional property of having a size less or equal to  $ub$ ;
  2. Mine all frequent itemsets having a size less or equal to  $ub$  which have the additional property of being closed.
- Give the set of solutions of  $Q$  with the two interpretations on  $\mathcal{D}$  with  $\theta = 3$ .
  - what is the correct semantic of such query ? explain ?

#### Algorithm 1: LCM

```

1 InOut :  $X$  : Closed Frequent Itemset ;
2 In :  $\theta$  : minsup
3 print( $X$ )
4 foreach  $i > tail(X)$  do
5   if  $frequent(X \cup \{i\}) \geq \theta$  then
6      $Y \leftarrow \bigcap_{t \in cover(X \cup \{i\})} t$ 
7     if  $Y = child(X)$  then  $LCM(X, \theta)$ 

```