

ϵ : seuil support minimal

I : ens des items

D : dataset

```
main :  
    foreach e : I do  
        backtracking( $\emptyset$ , e, D,  $\epsilon$ )
```

```
backtracking(P, e, D,  $\epsilon$ ) :  
    if supportD(P  $\cup$  {e})  $\geq \epsilon$   
        output (P  $\cup$  {e}) // on a trouvé un itemset fréquent  
        Q  $\leftarrow$  P  $\cup$  {e}  
        foreach i  $\in$  I tq i  $\supset$  e and i  $\not\subset$  P do  
            backtracking(Q, i, D,  $\epsilon$ )
```

Itemsets fréquents maximax

Soit P un itemset fréquent

Il est **maximal** s'il n'existe pas Q (itemset fréquent) tq P \supset Q

Itemset fréquents fermés (1999 découvert par un français)

Soit P un itemset fréquent

Il est **fermé** s'il n'existe pas Q (items fréquent)

tq :

- P \supset Q
- et
- support(P) = support(Q)

ens de tous les itemsets fréquents

\supseteq

ens de tous les itemsets fréquents fermés

\supseteq

ens de tous les itemsets maximaux

A partir des "fermés" on peut retrouver tous les itemsets fréquents

→ compression **sans perte**

algo LCM :

```
main :  
    Pfermé $\emptyset$   $\leftarrow \bigcap_{t \in D} t$   
    output(Pfermé $\emptyset$ )  
    foreach e  $\in$  I tq e  $\not\subset$  Pfermé $\emptyset$  do  
        expand(Pfermé $\emptyset$ , e, D,  $\epsilon$ )
```

D_P : dataset "réduit" par le pattern P

$D_P = \{t \in D \mid t \supseteq P\}$

```
expand(P, e, D_P, ε)
  if supportD_P(P ∪ {e}) ≥ ε
     $P_{ext} \leftarrow \bigcap_{t \in DP[\{e\}]} t$  //  $D_P[\{e\}] = \{t \in D_P \mid e \in t\}$ 
    if maxitem( $P_{ext}$ ) = e
       $Q \leftarrow P \cup P_{ext}$ 
      output(Q)
       $D_Q = \{t \setminus Q \mid t \in D_P[\{e\}]\}$ 
      foreach  $i \in I$  tq  $i < e$  and  $i \notin Q$  do
        expand(Q, i,  $D_Q$ , ε)
```