

## Ejemplo FP Growth

El siguiente ejemplo, busca ayudarlos a que puedan ver el patrón que sigue este algoritmo y les facilite la tarea de generalización para que puedan programarlo.

Tenemos las siguientes transacciones y un **soporte mínimo de 4**.

K S O	K T S
K S	K M
K O L	K M
B S O	K M
B S O	K T
T B S O	K T
T B S O	K T
T B O	K T
T B O	B S
T B O	B S
T B O L	B S
K T S	

Si contamos cada ítem, nos queda: K: 12 S: 11 O:10 T:12 L: 2 B:11 M:3

1. Ordenados quedan: K T B S O M L (En caso de empate, se optó por orden alfabético).
2. Se elimina el L y M cuyo soporte es menor a 4 de cada transacción.

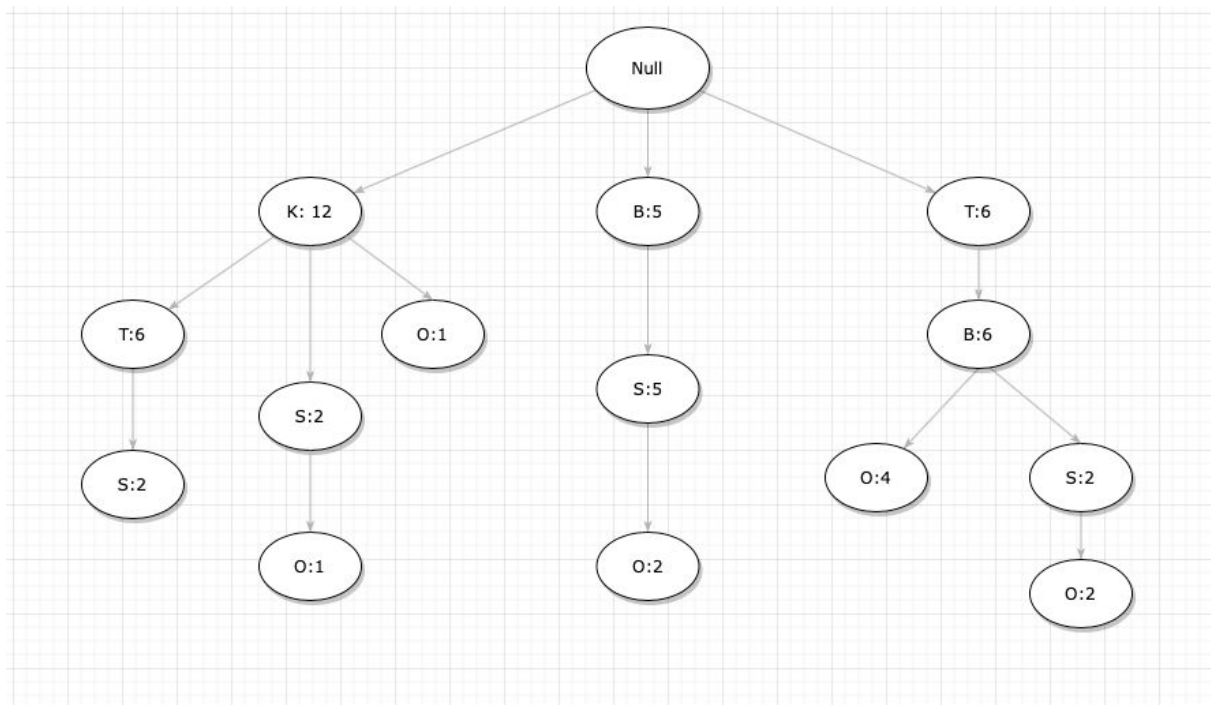
De este modo, se forma el siguiente árbol:

```

ROOT
├─ K:12
│  └─ T:6
│     └─ S:2
│        └─ S:2
│           └─ O:1
│              └─ O:1
│                 └─
│                    └─ B:5
│                       └─ S:5
│                          └─ O:2
│                             └─
│                                └─ T:6
│                                   └─ B:6
│                                      └─ O:4
│                                         └─ S:2
│                                            └─ O:2

```

Una versión más elegante de ver es:



Para este ejemplo, **haremos sólo el CPB del ítem “O”**, los demás son totalmente análogos. Por lo tanto, lo primero que hacemos es buscar **cada ruta al ítem O**.

K O: 1  
 K S O:1  
 B S O:2  
 T B O:4  
 T B S O:2

Ahora vamos a contar cada ítem {K:2 S:5 B:8 T:6 O:10}

Dado que **K tiene un soporte menor a 4**, se elimina. De este modo, el **real CPB** que nos interesa es:

S:1  
 B S:2  
 T B:4  
 T B S:2

Una forma de continuar es hacer borrón y cuenta nueva, es decir, tomar el CPB como nuevas transacciones de **un nuevo árbol** (ordenados por su soporte), es decir:

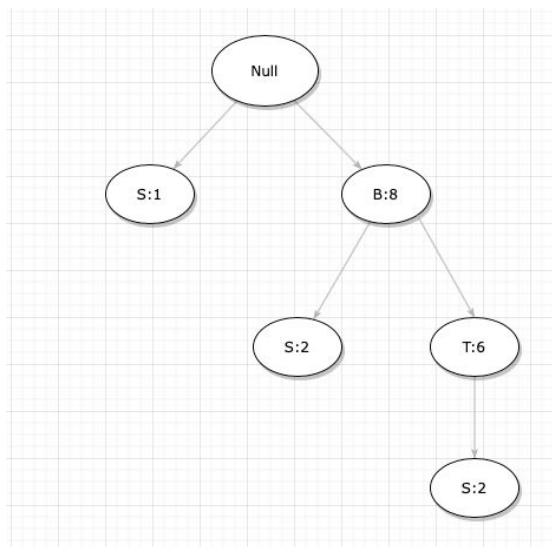
S	B T
B S	B T
B S	B T S
B T	B T S
B T	

Volvemos a generar un árbol. Sólo que ahora **debemos considerar que este árbol fue formado dado que se usó el ítem O**. Es un árbol condicionado al ítem O, es decir, todos los *itemset* frecuentes que se encuentren en el nuevo árbol, **si o si** tienen el ítem O.

**Hint:** Notar que en esta parte se aprecia un proceso recursivo: (1) Formamos un árbol con las transacciones, (2) encontramos el CPB y (3) formamos un árbol sólo con las transacciones de ese CPB, sólo que este nuevo árbol está condicionado al ítem O.

```
ROOT
├ S:1
│
├ B:8
│   ├── S:2
│   ├── T:6
│   └ S:2
```

Una versión elegante de este árbol es:



Ahora vamos a hacer un nuevo CPB de este árbol **para cada ítem**.

#### 1. CPB de B

No olvidar que en verdad es CPB de B dado el CPB de O

Aquí ya tenemos un **itemset frecuente {B, O: 8}** El 8 sale porque la B está en total 8 veces en el nuevo árbol.

CPB: Null

Dado que el CPB el **null**, no podemos seguir. Paramos en esta parte.

## 2. CPB de T

No olvidar que en verdad es CPB de T dado el CPB de O

Aquí ya tenemos un **itemset frecuente {T, O: 6}**. El 6 sale porque el T está en total 6 veces en el nuevo árbol.

CPB: {B:6}

Nuevamente hacemos borrón y cuenta nueva y generamos un nuevo árbol con sólo las transacciones de B:6, es decir

B  
B  
B  
B  
B  
B

Lo cual da un árbol de la forma

ROOT  
|\_B:6

### 2.1 Hacemos el CPB de B

No olvidar que en verdad es CPB de B dado el CPB de T dado el CPB de O.

Aquí ya tenemos un **itemset frecuente {B, T, O : 6}**. El 6 sale porque la B está en total 6 veces.

CPB: Null

Dado que el CPB es **null**, no podemos seguir. Paramos en esta parte.

### 3. CPB de S

No olvidar que en verdad es CPB de S dado el CPB de O

Aquí ya tenemos un **itemset frecuente {S, O: 5}**. El 5 sale porque la S está en total 5 veces.

CPB {B:2} y {B T: 2}

Con una suma de { B:4 T:2 }

Dado que hay sólo 2 T que es menor al mínimo soporte (4), se elimina. Así que el real CPB es

B:2

B:2

Nuevamente hacemos borrón y cuenta nueva y generamos un nuevo árbol con sólo las transacciones de B:2 y B:2, es decir

B

B

B

B

Lo cual da un árbol de la forma

Root

|\_ B:4

#### 3.1 CPB de B.

No olvidar que en verdad es CPB de B dado el CPB de S dado el CPB de O.

Aquí ya tenemos un **itemset frecuente {B, S, O : 4}**. El 4 sale porque la B está en total 4 veces en el nuevo árbol.

CPB: Null

Dado que el CPB es **null**, no podemos seguir. Paramos en esta parte.

Con todo lo anterior, los itemset frecuentes **que contienen O** son:

- {B, O: 8}
- {T, O: 6}.
- {B, T, O : 6}.
- {S, O: 5}
- {B, S, O : 4}

Luego se hace lo mismo, a partir del árbol inicial pero con otro ítem distinto a O.