Considere el siguiente esquema global y de fragmentación para una base de datos:

EMPLEADO=(ENUM, ENOMBRE, SALARIO, IRPF, DNUM)
DEPARTAMENTO=(DNUM, DNOMBRE, ÁREA, MGRNUM)
PROVEEDOR=(PNUM, PNOMBRE, CIUDAD)
SUMINISTRO=(PNUM, ANUM, DNUM, CANTIDAD)

$\text{EMPLEADO}_1 = \mathbf{SL}_{DNUM \leq 10}(\text{EMPLEADO})$
$\text{EMPLEADO}_2 = \mathbf{SL}_{10 < DNUM \leq 20}(\text{EMPLEADO})$
$\text{EMPLEADO}_3 = \mathbf{SL}_{DNUM > 20}(\text{EMPLEADO})$

$\text{DEPARTAMENTO}_1 = \mathbf{SL}_{DNUM \leq 10}(\text{DEPARTAMENTO})$
$\text{DEPARTAMENTO}_2 = \mathbf{SL}_{10 < DNUM \leq 20}(\text{DEPARTAMENTO})$
$\text{DEPARTAMENTO}_3 = \mathbf{SL}_{DNUM > 20}(\text{DEPARTAMENTO})$

$\text{PROVEEDOR}_1 = \mathbf{SL}_{CIUDAD = \text{“}San\ Francisco\text{”}}(\text{PROVEEDOR})$
$\text{PROVEEDOR}_2 = \mathbf{SL}_{CIUDAD = \text{“}Los\ Angeles\text{”}}(\text{PROVEEDOR})$
$\text{PROVEEDOR}_3 = \mathbf{SL}_{CIUDAD \neq \text{“}San\ Francisco\text{”}\ \wedge\ CIUDAD \neq \text{“}Los\ Angeles\text{”}}(\text{PROVEEDOR})$

$\text{SUMINISTRO}_1 = \text{SUMINISTRO}\ \mathbf{SJ}_{PNUM=PNUM}(\text{PROVEEDOR}_1)$
$\text{SUMINISTRO}_2 = \text{SUMINISTRO}\ \mathbf{SJ}_{PNUM=PNUM}(\text{PROVEEDOR}_2)$
$\text{SUMINISTRO}_3 = \text{SUMINISTRO}\ \mathbf{SJ}_{PNUM=PNUM}(\text{PROVEEDOR}_3)$

Considere la siguiente consulta:

“Listar la cantidad total de artículos suministrados a cada departamento (nombre) con números mayores de 17 y menores de 23, por proveedores de Los Ángeles, Sacramento o San José y que hayan facturado por cantidades totales superiores a 5000”

1. Traduzca dicha consulta a una expresión SQL y a una expresión del álgebra relacional (o del álgebra relacional extendida).

1.1. Expresión SQL.

```
SELECT SUM(CANTIDAD), DNOMBRE
FROM DEPARTAMENTO D, SUMINISTRO S, PROVEEDOR P
WHERE D.DNUM = S.DNUM AND S.PNUM = P.PNUM
AND S.ANUM BETWEEN 17 AND 23
AND (P.CIUDAD="Los Ángeles" OR P.CIUDAD="Sacramento" OR P.CIUDAD="San José")
GROUP BY DNOMBRE
HAVING SUM(CANTIDAD) > 5000;
```

1.2. Expresión algebraica.

Por simplicidad vamos a denominar:

j1: D.DNUM = S.DNUM
j2: S.PNUM = P.PNUM
c1: S.ANUM BETWEEN 17 AND 23
c2: P.CIUDAD = "Los Ángeles" OR P.CIUDAD="Sacramento" OR P.CIUDAD="San José"
c3: SUM(CANTIDAD) > 5000

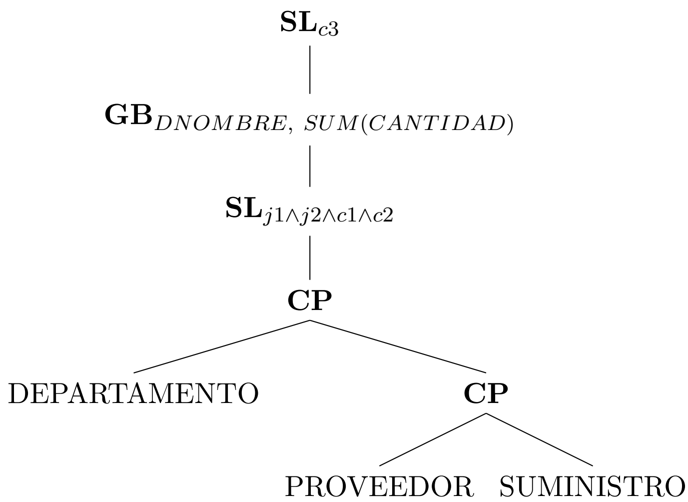
Y por tanto la expresión algebraica asociada a la consulta será:

$\mathbf{SL}_{c3}(\mathbf{GB}_{DNOMBRE,\ SUM(CANTIDAD)}(\mathbf{SL}_{j1 \wedge j2 \wedge c1 \wedge c2}(D\ \mathbf{CP}\ (P\ \mathbf{CP}\ S))))$

2. Transforme esta consulta global a una consulta reducida sobre fragmentos, usando la heurística de transformación y los criterios 1 a 4 para simplificarla.

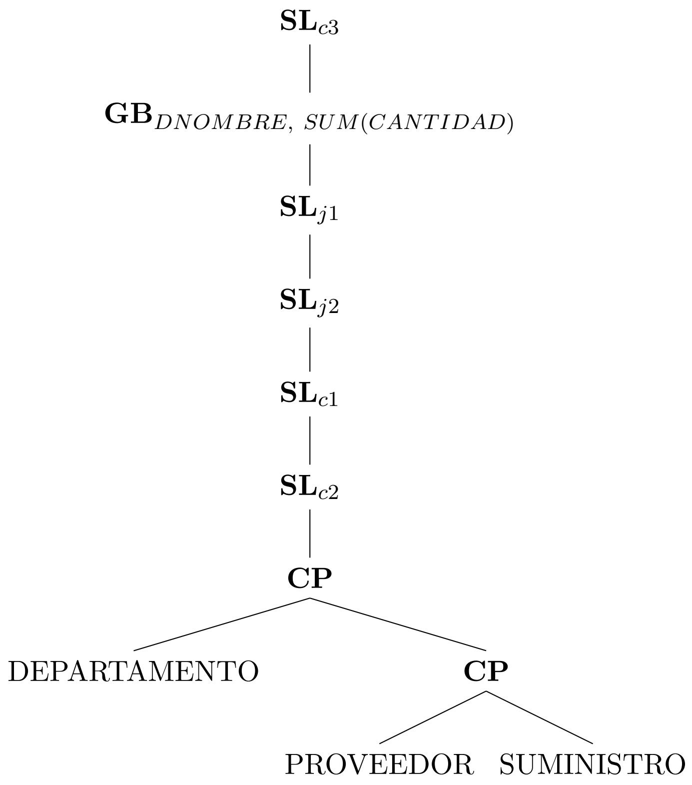
2.1. Etapa de descomposición de consultas.

Árbol inicial:



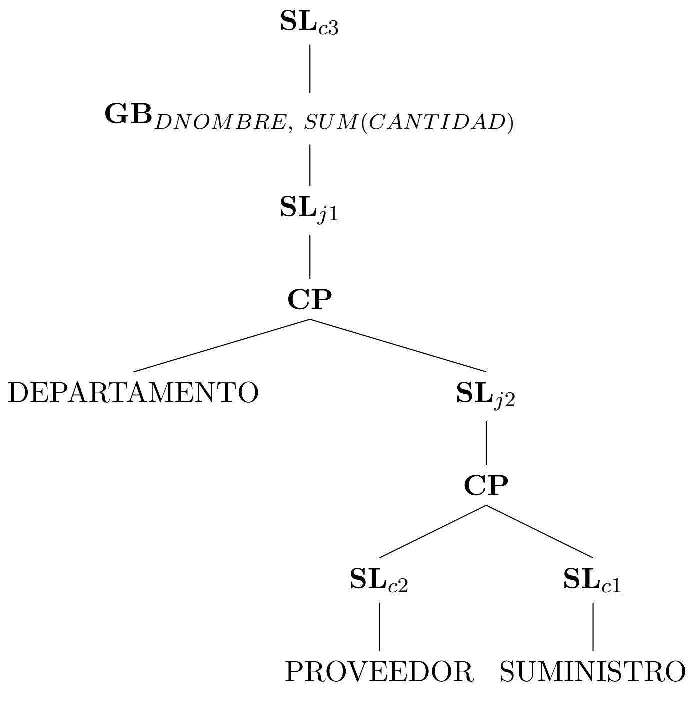
Paso 1: Empleando la idempotencia de operadores unarios, descomponemos las operaciones de selección con predicados conjuntivos.

Obtenemos así el siguiente árbol:



Paso 2: Empleando la conmutatividad de la selección con operadores binarios, desplazamos cada selección tan abajo en el árbol como lo permitan los atributos del predicado.

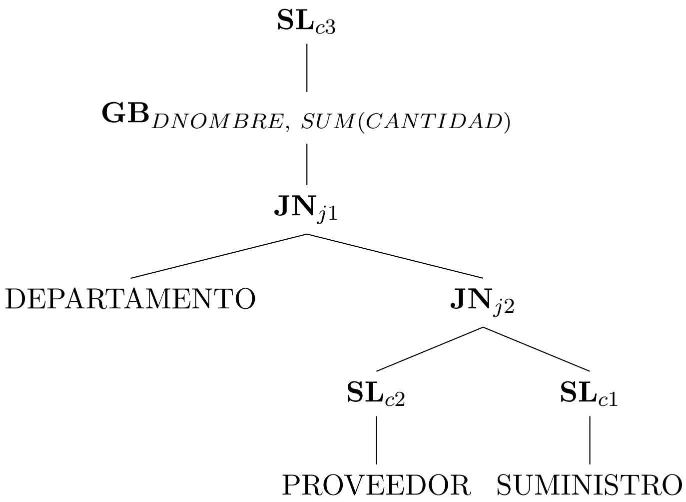
Obtenemos así el siguiente árbol:



No es necesario readaptar los nodos hoja puesto que no hay selecciones más restrictivas que se ejecuten primero, por ello nos saltamos el **Paso 3**.

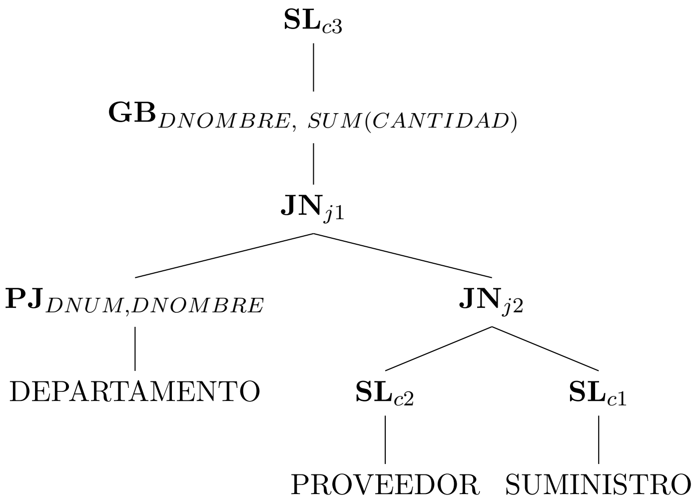
Paso 4: Combinamos una operación de producto cartesiano con una selección subsiguiente, cuyo predicado represente una condición de reunión (join).

Obtenemos así el siguiente árbol:



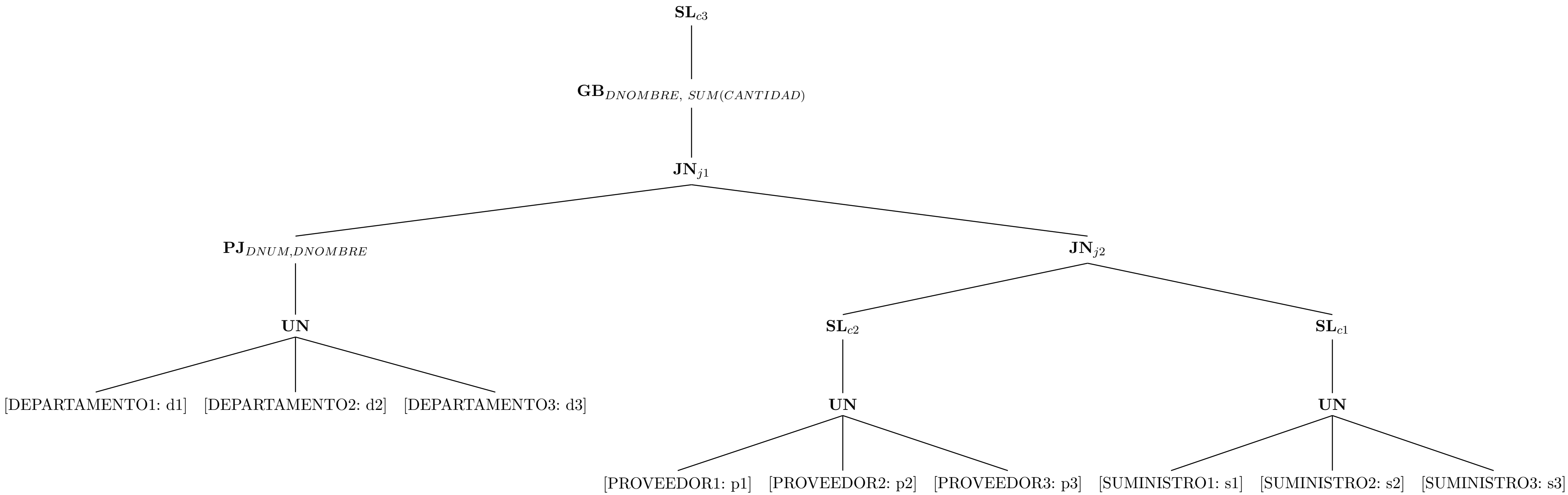
Paso 5: Empleando la idempotencia de operadores unarios y la conmutatividad de la proyección con operadores binarios, descomponemos las listas de atributos de proyección y las desplazamos lo más bajo posible en el árbol.

Obtenemos así el Árbol final:



1.2. Etapa de localización de datos distribuidos.

Árbol canónico:



d1: DNUM $\leq$ 10
d2: 10 $<$ DNUM $\leq$ 20
d3: DNUM $>$ 20
p1: CIUDAD = "San Francisco"
p2: CIUDAD = "Los Ángeles"
p3: CIUDAD $\neq$ "San Francisco" $\wedge$ CIUDAD $\neq$ "Los Ángeles"
s1: S.PNUM=P.PNUM $\wedge$ P.CIUDAD = "San Francisco"
s2: S.PNUM=P.PNUM $\wedge$ P.CIUDAD = "Los Ángeles"
s3: S.PNUM=P.PNUM $\wedge$ P.CIUDAD $\neq$ "San Francisco" $\wedge$ P.CIUDAD $\neq$ "Los Ángeles"

Aplicación del criterio 1: Evaluación de las cualificaciones.

$\mathbf{SL}_{\text{P.CIUDAD=“Los Ángeles” } \vee \text{ P.CIUDAD=“Sacramento” } \vee \text{ P.CIUDAD=“San José”}}$ [Pi: pi]

- **q1**: CIUDAD=“San Francisco” $\wedge$ (P.CIUDAD=“Los Ángeles” $\vee$ P.CIUDAD=“Sacramento” $\vee$ P.CIUDAD=“San José”)

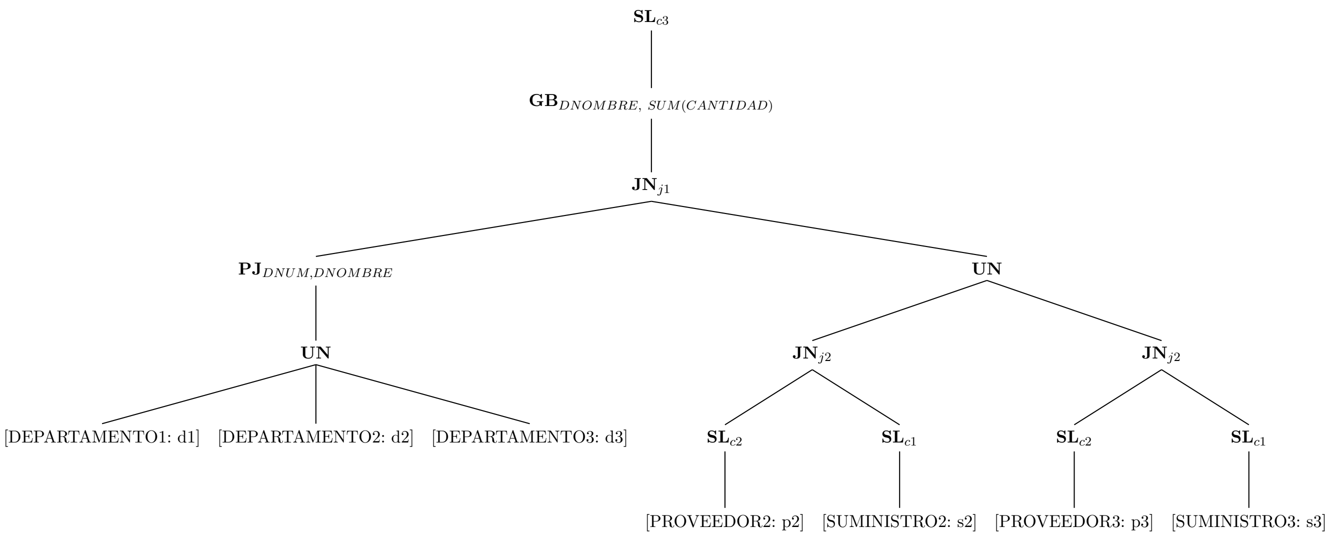
  Contradicción.

- **q2**: CIUDAD = “Los Ángeles” $\wedge$ (P.CIUDAD=“Los Ángeles” $\vee$ P.CIUDAD=“Sacramento” $\vee$ P.CIUDAD=“San José”)
- **q3**: CIUDAD $\neq$ “San Francisco” $\wedge$ CIUDAD $\neq$ “Los Ángeles” $\wedge$ (P.CIUDAD=“Los Ángeles” $\vee$ P.CIUDAD=“Sacramento” $\vee$ P.CIUDAD=“San José”)

$\mathbf{SL}_{\text{ANUM BETWEEN 17 AND 23}}$ [Si: si]

En este caso, ninguna cualificación resultante es una contradicción.

Aplicación del criterio 2 y 3: Para el segundo subimos uniones por encima de los productos naturales y para el tercero evaluamos las cualificaciones de los productos naturales.



Aplicación de nuevo del criterio 2 y del criterio 3: evaluación de cualificaciones de los productos naturales.

[Di: di] $\mathbf{JN}_{j1}$ [Pj $\mathbf{JN}_{j2}$ Sj: pj $\wedge$ sj $\wedge$ $\mathbf{JN}_{j2}$] para i=1,2,3, j=2,3.

- **q122**: DNUM $\leq$ 10 $\wedge$ P.CIUDAD=“Los Ángeles” $\wedge$ S.PNUM=P.PNUM $\wedge$ P.CIUDAD = “Los Ángeles”
- **q123**: DNUM $\leq$ 10 $\wedge$ P.CIUDAD=“Los Ángeles” $\wedge$ S.PNUM=P.PNUM $\wedge$ P.CIUDAD $\neq$ “San Francisco” $\wedge$ P.CIUDAD $\neq$ “Los Ángeles”
- **q132**: DNUM $\leq$ 10 $\wedge$ P.CIUDAD $\neq$ “San Francisco” $\wedge$ P.CIUDAD $\neq$ “Los Ángeles” $\wedge$ S.PNUM=P.PNUM $\wedge$ P.CIUDAD = “Los Ángeles”

- **q133**: DNUM $\leq$ 10 $\wedge$ P.CIUDAD $\neq$ “San Francisco” $\wedge$ P.CIUDAD $\neq$ “Los Ángeles” $\wedge$ S.PNUM=P.PNUM $\wedge$ P.CIUDAD $\neq$ “San Francisco” $\wedge$ P.CIUDAD $\neq$ “Los Ángeles”
- **q222**: 10 $<$ DNUM $\leq$ 20 $\wedge$ P.CIUDAD=“Los Ángeles” $\wedge$ S.PNUM=P.PNUM $\wedge$ P.CIUDAD = “Los Ángeles”
- **q223**: 10 $<$ DNUM $\leq$ 20 $\wedge$ P.CIUDAD=“Los Ángeles” $\wedge$ S.PNUM=P.PNUM $\wedge$ P.CIUDAD $\neq$ “San Francisco” $\wedge$ P.CIUDAD $\neq$ “Los Ángeles”
- **q232**: 10 $<$ DNUM $\leq$ 20 $\wedge$ P.CIUDAD $\neq$ “San Francisco” $\wedge$ P.CIUDAD $\neq$ “Los Ángeles” $\wedge$ S.PNUM=P.PNUM $\wedge$ P.CIUDAD = “Los Ángeles”
- **q233**: 10 $<$ DNUM $\leq$ 20 $\wedge$ P.CIUDAD $\neq$ “San Francisco” $\wedge$ P.CIUDAD $\neq$ “Los Ángeles” $\wedge$ S.PNUM=P.PNUM $\wedge$ P.CIUDAD $\neq$ “San Francisco” $\wedge$ P.CIUDAD $\neq$ “Los Ángeles”
- **q322**: DNUM $>$ 20 $\wedge$ P.CIUDAD=“Los Ángeles” $\wedge$ S.PNUM=P.PNUM $\wedge$ P.CIUDAD = “Los Ángeles”
- **q323**: DNUM $>$ 20 $\wedge$ P.CIUDAD=“Los Ángeles” $\wedge$ S.PNUM=P.PNUM $\wedge$ P.CIUDAD $\neq$ “San Francisco” $\wedge$ P.CIUDAD $\neq$ “Los Ángeles”
- **q332**: DNUM $>$ 20 $\wedge$ P.CIUDAD $\neq$ “San Francisco” $\wedge$ P.CIUDAD $\neq$ “Los Ángeles” $\wedge$ S.PNUM=P.PNUM $\wedge$ P.CIUDAD = “Los Ángeles”
- **q333**: DNUM $>$ 20 $\wedge$ P.CIUDAD $\neq$ “San Francisco” $\wedge$ P.CIUDAD $\neq$ “Los Ángeles” $\wedge$ S.PNUM=P.PNUM $\wedge$ P.CIUDAD $\neq$ “San Francisco” $\wedge$ P.CIUDAD $\neq$ “Los Ángeles”

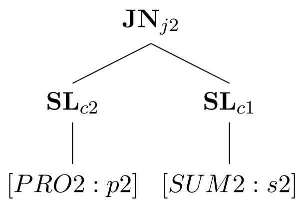
No hay ninguna cualificación resultante contradictoria.

Para el resto de árboles utilizaremos la siguiente notación para poder representar al completo los gráficos resultantes.

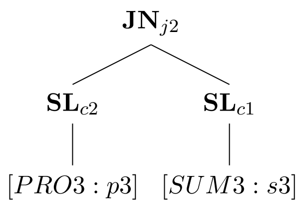
Para abreviar utilizaremos DEPARTAMENTO=DEP, PROVEEDOR=PRO y SUMINISTRO=SUM.

Utilizaremos los siguientes subárboles que se usan varias veces para evitar ramificar demasiado el árbol.

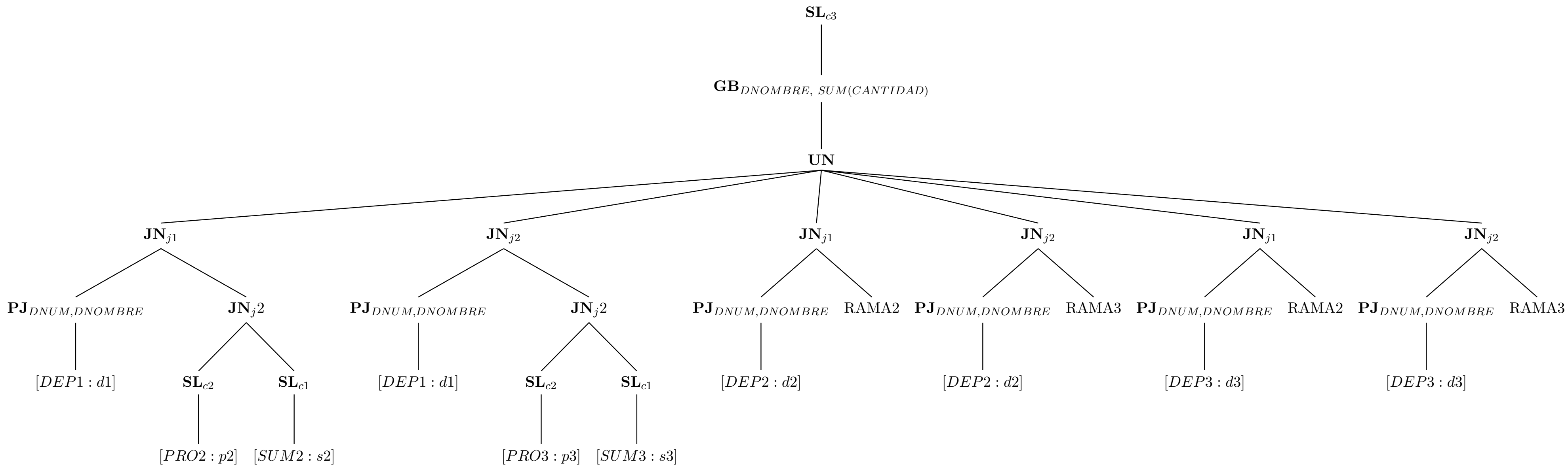
Denotaremos como RAMA2 a:



Denotaremos como RAMA3 a:



Aplicando la idempotencia del operador proyección obtenemos:



Realmente se realiza la misma ramificación que la obtenida para el DEPARTAMENTO1 para los otros dos departamentos.

Se cumplen todas las condiciones para aplicar el **Criterio 4** (subir las uniones por encima del **GB**) a la vez que bajamos $\mathbf{SL}_{c3}$ por fragmento.

Obtenemos el Árbol final: