LA SUBSECUENCIA COMÚN MÁXIMA

En muchos problemas se quiere conocer cuál es la composición de la solución óptima, es decir, los elementos que la conforman. En este caso, se deben conservar los valores de las soluciones parciales y la forma como se llega a ellas.

En el siguiente ejemplo se crea una tabla y a partir de ella se reconstruye la solución, tratándose del cálculo de la subsecuencia común máxima.

Ejemplo 1

Dada una secuencia X = $\{x_1\ x_{2\ldots}\ x_m\}$, se dice que Z = $\{z_1\ z_{2\ldots}\ z_k\}$, es una subsecuencia de X (siendo k ≤ m) si existe una secuencia creciente $\{i_1\ i_{2\ldots}\ i_k\}$ de índices de X tales que para todo j = 1,2,..., k se tiene $X_{ij} = Z_j$

Dadas dos secuencias, X y Y, se dice que Z es una subsecuencia común de X y Y si es subsecuencia de X y subsecuencia de Y. Se debe determinar la subsecuencia de longitud máxima común a dos secuencias.

Solución:
A la longitud de la secuencia común máxima de las secuencias X_i y Y_j se llamará L (i,j), siendo X_i el i – ésimo estipulo de X (es decir, X_i = $\{x_1\ x_{2\ldots}\ x_i\}$) y Y_j el j – ésimo estipulo de Y, (es decir, Y_j = $\{y_1\ y_{2\ldots}\ y_j\}$).

Se puede plantear la solución como una sucesión de decisiones, aplicando el principio de óptimo en el que en cada marcha se establece si una representación forma parte o no de la secuencia común máxima. Comenzando por los últimos caracteres de las dos secuencias X y Y, la solución viene dada por la siguiente relación en recurrencia:

$$L\,(i,j) = \begin{Bmatrix} 0 & si\ i\ = 0\ o\ j = 0 & \\ L\,(i\ -1, j-\ 1) + 1 & si\ i \neq 0, & j \neq 0\ y\ x_i =\ y_j \\ Max\,\{L\,(i, j-1), L\,(i-1, j)\} & si\ i \neq 0, & j \neq 0\ y\ x_i \neq\ y_j \end{Bmatrix}$$

La salida recursiva es de orden exponencial y, por tanto, se construirá una tabla con los valores L (i,j) para evitar la repetición de cálculos ayudados de la programación dinámica. Para la construcción de la tabla suponga que X y Y son las sucesiones de valores:

X = {1 0 0 1 0 1 0 1}
Y = {0 1 0 1 1 0 1 1 0}

		0	1	2	3	4	5	6	7	8
			1	0	0	1	0	1	0	1
0		0	0	0	0	0	0	0	0	0
1	0	0	0	1	1	1	1	1	1	1
2	1	0	1	1	1	2	2	2	2	2
3	0	0	1	2	2	2	3	3	3	3
4	1	0	1	2	2	3	3	4	4	4
5	1	0	1	2	2	3	3	4	4	5
6	0	0	1	2	3	3	4	4	5	5
7	1	0	1	2	3	4	4	5	5	6
8	1	0	1	2	3	4	4	5	5	6
9	0	0	1	2	3	4	5	5	6	6

Tabla 4.3. Cálculo de la subsecuencia común máxima.

La tabla se va construyendo por filas y rellenando de izquierda a derecha. En cada L[i,j] hay dos datos: uno para la construcción de lo que viene después de la secuencia óptima y otro necesario que corresponde a la longitud de cada subsecuencia.

La salida a lo que viene después de la secuencia normal máxima de las secuencias X y Y, se encuentra en el extremo inferior derecho (L[9,8]) y, por tanto, su longitud es 6. Si se quiere obtener lo que viene después de la secuencia, se debe recorrer la tabla (zona sombreada) a partir de esta posición y seguir la información que indica cómo obtener las longitudes óptimas a partir de su procedencia (izquierda, diagonal o superior).