

Solucionando el problema de la mochila en el Modelo Sticker

Ernesto Mancebo

Febrero 2020

Abstract

Este documento tiene como objetivo la resolución de los problema *suma de subconjuntos* y *problema de la mochila*, para éste último en sus versiones acotada y la no acotada, utilizando como apoyo para la resolución de ambos problemas las subrutinas *ordenado por cardinalidad* y *llenado en paralelo*.

1 Introducción



2 Modelo Sticker

El Modelo Sticker es un modelo de computación inspirado en cadenas de ADN propuesto por Sam Roweis [Sam Roweis et al. “A Sticker-Based Model for DNA Computation”](#). In: *Journal of computational biology : a journal of computational molecular cell biology* 5 (Feb. 1998), pp. 615–29. DOI: [10.1089/cmb.1998.5.615](#), donde tal modelo se basa en procedimientos de filtrado, con memoria de acceso aleatorio y una nueva manera de codificar la información. En ese mismo orden, se distingue de otros modelos computacionales por la manera en que codifica la información.

2.1 Concepto de Cadena de Memoria

El eje central de este modelo son las cadenas de memoria, tales cadenas consisten en hebras simples de ADN del tipo (n, k, m) , tal que $n \geq k.m$, siendo n la longitud de la cadena, k las cantidad de subcadenas o regiones, y m la longitud de cada subcadena. Las cadenas de memoria también se les conocen como moléculas y son representadas por sigma (σ).

En ese mismo orden de ideas, un sticker es un elemento asociado a una región en la cadena de memoria, una vez asociado se dice que la región está complementada, transformando la cadena en una hebra doble parcial. En ese sentido, se dice que en cada región donde haya un sticker complementándole, esta región está *activada*, mientras que la región donde se carezca de sticker se dice está *desactivada*. Tal abstracción binaria también se puede expresar de modo $0/1$.

2.2 Concepto de Tubo

Dentro del modelo sticker un tubo es un multiconjunto de complejos de memorias del mismo tipo. Las operaciones a principales dentro de este modelo son:

- $mezclar(T_1, T_2)$: Retorna la unión de los dos tubos devolviendo un solo tubo el cual contiene cada complejo de memoria proveniente de ambos tubos. También se nota como $(T_1 \cup T_2)$.
- $separar(T, i)$: Para un tubo T y un índice i tal que $(1 \leq i \leq n)$, retorna un $+(T, i)$ y $-(T, i)$, de modo que el primer tubo contiene todos los complejos de memoria cuya región i -ésima esté *activada*, mientras que el segundo contiene los complejos cuya región i -ésima *desactivada*.
- $encender(T, i)$: Para un tubo T y un índice i tal que $(1 \leq i \leq n)$, retorna un tubo T' con la región i -ésima *activada* en cada complejo de memoria.
- $apagar(T, i)$: Para un tubo T y un índice i tal que $(1 \leq i \leq n)$, retorna un tubo T' con la región i -ésima *desactivada* en cada complejo de memoria.
- $leer(T)$: Dado un tubo $T \neq \emptyset$, lee su contenido.

2.3 Concepto de Librería

Una librería dentro del modelo sticker consiste en complejos de memoria de modo (k, l) , teniendo k sub cadenas y las primeras l subcadenas están *activadas* o *desactivadas* en todas sus posibles combinaciones, mientras que las cadenas $k - l$ están *desactivadas*.

3 Subrutina Ordenado por Cardinalidad

Problema: Sea $A = \{1, \dots, p\}$, $B = \{b_1, \dots, b_s\} \subseteq A$, $F = \{D_1, \dots, D_t\} \subseteq P(A)$. Ordene los conjuntos de F de acuerdo a su cardinalidad en B , o en otras palabras, la cantidad de elementos de $B \cap D_i$.

El paper bajo estudio nos sugiere el implementar un programa molecular cuya entrada sea un tubo T_0 conteniendo complejos de memoria codificando cada conjunto de la familia F . Para esto, cada complejo de memoria en T_0 se representará mediante la función booleana $T_0 = \{\{\sigma : |\sigma| = p \wedge \exists j(\chi_{D_j} = \sigma)\}\}$; tal que χ_{D_j} es la función característica de D_j en A , siendo $(\chi_{D_j}(i) = 1$ si $i \in D_j$ de lo contrario $\chi_{D_j} = 0$ si $i \in A - D_j$).

El programa a diseñar consistirá en un bucle principal realizando s pasos, donde para el paso i -ésimo se tendrá $i + 1$ tubos (T_0, T_1, \dots, T_i) , tales tubos generados verificando la condición $\forall \sigma(\sigma \in T_j \rightarrow |\sigma \cap \{b_1, \dots, b_i\}| = j)$.

3.1 Algoritmo

Las ideas previas nos indican el siguiente algoritmo:

```

1: procedure Cardinal_Sort( $T_0, B$ )
2:   for  $i = 1$  to  $s$  do
3:      $(T_0, T'_0) \leftarrow \text{separar}(T_0, b_i)$ 
4:     for  $j = 0$  to  $i - 1$  do
5:        $(T'_{j+1}, T''_j) \leftarrow \text{separar}(T_j, b_i)$ 
6:        $T_j \leftarrow \text{mezclar}(T'_j, T''_j)$ 
7:     end for
8:      $T_i \leftarrow T'_i$ 
9:   end for
10:  Return  $T_0, \dots, T_s$ 
11: end procedure

```

3.2 Traza Subrutina Ordenado por Cardinalidad

A modo de ilustrar el comportamiendo y los conceptos empleados en el algoritmo concebido, tenemos:

- $A : \{0, 1, 2, 3, 4, 5, 6\}$
- $B : \{1, 2, 4\}$
- $F : \{\{2, 6\}, \{3\}, \{4\}, \{2, 4\}\}$

Identificando los elementos que cumplen $B \cap D_j$ en F le resaltamos en rojo de modo que tenemos: $\{\{\textcolor{red}{2}, 6\}, \{3\}, \{\textcolor{red}{4}\}, \{\textcolor{red}{2}, \textcolor{red}{4}\}\}$. Codificando F para llevarlo a un tubo tendíamos:

$$\begin{bmatrix} 0010001 \\ 0001000 \\ 0000100 \\ 0010100 \end{bmatrix}$$

Este T_0 codificando F tras la ejecución de *Cardinal_Sort* tendremos:

$$\begin{aligned} T_0 &: [0001000] \\ T_1 &: \begin{bmatrix} 0000100 \\ 0010001 \end{bmatrix} \\ T_2 &: [0010100] \\ T_3 &: [] \\ T_4 &: [] \\ T_5 &: [] \\ T_6 &: [] \end{aligned}$$

Cumpliendo de esta manera la condición de $\forall \sigma (\sigma \in T_j \rightarrow |\sigma \in \{b_1, \dots, b_i\}| = j)$.

3.3 Otras Notaciones

Para usos posteriores dentro de este informe, podemos notar el uso de *Cardinal_Sort* de modo:

- $Cardinal_Sort(T_0)$ cuando $B = A$.
- $Cardinal_Sort(T_0, l, k)$ cuando $B = \{l, l + 1, \dots, k\}$.

4 Subrutina de Llenado

Problema: Sea $A = \{1, \dots, p\}$, $r \in \mathbb{N}$ y $f : A \rightarrow \mathbb{N}$ una función. Si $B \subseteq A$, notamos $f(B) = \sum_{i \in B} f(i)$, por conveniencia deifnimos que para $f(0) = 0$. Sea $q_f = f(A)$, $A_i = \{0, \dots, i\}$ ($0 \leq i \leq p$), y T_0 un multiconjunto de complejos de memoria de forma (n, k, m) con $k \geq p + r + q_f$.

Podemos decir que si $\sigma \in T_0$, podemos asumir que tal tubo está formado por las regiones:

- $(A\sigma) = \sigma(1) \cdots \sigma(p)$
- $(F\sigma) = \sigma(p+r+1) \cdots \sigma(p+r+q_f)$
- $(L\sigma) = \sigma(p+1) \cdots \sigma(p+r)$
- $(R\sigma) = \sigma(p+r+q_f+1) \cdots$

La subrutina opera sobre T_0 modificando sus elemento haciendo las moléculas del tubo de salida en $(F\sigma)$ el peso con respecto a f del subconjunto en A codificado en $(A\sigma)$, siendo esto:

$$\sum_{i=1}^p \sigma(i) f(i) \sum_{j=p+r+1}^{p+r+q_f} \sigma(j)$$

4.1 Algoritmo

```

1: procedure Parallel_Fill( $T_0, f, p, r$ )
2:   for  $i = 1$  to  $p$  do
3:      $(T_{i,0}^+, T_i^-) \leftarrow \text{separar}(T_{i-1}, i)$ 
4:     for  $j = 1$  to  $f(i)$  do
5:        $T_{i,j}^+ \leftarrow \text{encender}(T_{i,j}^+, p + r + f(A_{i-1}) + j)$ 
6:     end for
7:      $T_i \leftarrow \text{mezclar}(T_{i,f(i)}^+, T_i^-)$ 
8:   end for
9:   Return  $T_0$ 
10: end procedure

```

Para cada i ($1 \leq i \leq p$) se consiredan las regiones: $R_i = \{p + r + f(A_{i-1}) + 1, \dots, p + r + f(A_i)\}$.

4.2 Traza Subrutina de Llenado

Con tal de ilustrar la subrutina bajo estudio, utilicemos como entrada:

- $A : \{1, 2, 3, 4\}$
- $B : \{2, 3, 4\}$
- $T_0 : \{\{2\}, \{4\}, \{3\}\}$

Una vez codificado el tubo T_0 , tenemos:

$$\begin{bmatrix} 01000000000000 \\ 00010000000000 \\ 00100000000000 \end{bmatrix}$$

Nótese que las primeras cuatro regiones resaltadas en rojo corresponden al tamaño p , reservado para codificar el valor de cada B . Una vez ilustrado el tubo T_0 , una vez procesado por *Parallel_Fill*, tendremos:

$$\begin{bmatrix} 0, 1, 0, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0 \\ 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1 \\ 0, 0, 1, 0, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0 \end{bmatrix}$$

5 Problema de Suma de Subconjuntos

Problema: Sea $A = \{1, \dots, p\}$ y $\omega \rightarrow \mathbb{N}$ una función peso. Sea $k \in \mathbb{N}$, tal que $k \leq \omega(A) = q_\omega$, determinar si existe un subconjunto $B \subseteq A$, tal que la suma de los pesos en B sea exactamente k .

Definido una vez el problema, el paper bajo estudio presenta un diseño de un programa dentro del modelo sticker tal que resuelve el problema de la suma de subconjuntos. El tubo de entrada T_0 conteniendo una librería de forma $p + q_\omega, q$.

La primera fase, la fase de llenado, cada molécula σ del tubo de entrada *será llenado a fin de conseguir en los últimos componentes q el peso de los subconjuntos*

que codifica. Los tubos obtenidos luego serán ordenados de acuerdo a su cardinalidad. Por último, el tubo k -ésimo es leído.

El tubo obtenido del proceso descrito codifica los subconjuntos de A para el peso k , en caso de existir.

5.1 Algoritmo

La descripción del planteamiento explicado sugiere un programa molecular como el que sigue:

```

1: procedure Subset_Sum( $p, w, k$ )
2:    $q_w \leftarrow \sum_{i=1}^p w(i)$ 
3:    $T_0 \leftarrow \text{Libería}(p + q_w, p)$ 
4:    $T_1 \leftarrow \text{Parallel\_Fill}(T_0, w, p, 0)$ 
5:    $T_{salida} \leftarrow \text{Cardinal\_Sort}(T_1, p + 1, p + q_w)[k]$ 
6:   leer( $T_{salida}$ )
7: end procedure

```

5.2 Traza

Tomando como referencia el tubo T_0 del apartado anterior:

$$\begin{bmatrix} 0100,1110000000 \\ 0001,1111111111 \\ 0010,1111110000 \end{bmatrix}$$

Para un $k = 10$, una vez el T_0 sea procesado por *Subset_Sum* tendríamos como salida: 00011111111111, pues si contamos los bits en la región azul nos da un total de 10.

5.3 Diseño

6 Problema de la Mochila Acotado

Problema: sea $A = \{1, \dots, p\}$ un conjunto finito no vacío, $\omega : A \rightarrow \mathbb{N}$ una función peso y $\rho : A \rightarrow \mathbb{N}$ una función valor. Sea $k, k' \in \mathbb{N}$ tal que $k \leq \omega(A) = q_\omega$ y $k' \leq \rho(A) = q_\rho$; determine si existe un subconjunto $B \subseteq A$ tal que $\omega(B) \leq k$ y $\rho(B) \geq k'$.

Esto nos lleva plantear el diseño de un programa molecular el cual resuelva el problema propuesto empezando con una librería $(p + q_\omega + q_\rho, p)$.

En la fase inicial, la fase de llenado, se procede como en el problema previo: cada complejo de memoria en el tubo inicial es llenado apropiadamente tal que codifique el *peso* asociado al subconjunto, seguido de la ordenación de los mismos de acuerdo a su cardinalidad de $\omega(A\sigma)$, empezando por obtener algunos tubos T_0, \dots, T_{q_ω} tales que $\forall \sigma (\sigma \in T_j \Rightarrow |\omega(A\sigma)| = j)$.

Con los tubos $T_0 \cup \dots \cup T_k$ se lleva a cabo una segunda operación de llenado, esta con respecto a los *valores*. Posteriormente, las moléculas σ del tubo resultante son ordenadas con respecto a su cardinalidad en $\rho(A\sigma)$, obteniendo nuevamente una serie de tubos de forma T_0, \dots, T_{q_ρ} , tales que $\forall \sigma (\sigma \in T_j \Rightarrow |\rho(A\sigma)| = j)$.

Finalmente, se obtiene la solución al problema con tubos $T'_k \cup \dots T_{q_\rho}$ y estos son leídos.

6.1 Algoritmo

Las directrices del apartado anterior guían a definir un programa molecular de la siguiente forma

```
1: procedure Bounded_Knapsack( $p, w, \rho, k, k'$ )
2:    $q_w \leftarrow \sum_{i=1}^p w(i); q_\rho \leftarrow \sum_{i=1}^p \rho(i); T_0 \leftarrow \text{Libería}(p + q_w + q_\rho, p)$ 
3:    $T_0 \leftarrow \text{Parallel\_Fill}(T_0, w, p, 0)$ 
4:    $T_0 \leftarrow \text{Cardinal\_Sort}(T_0, p + 1, p + q_w)$ 
5:    $T_1 \leftarrow \emptyset$ 
6:   for  $i = 1$  to  $k$  do
7:      $T_1 \leftarrow \text{mezclar}(T_1, \text{Cardinal\_Sort}(T_0, p + 1, p + q_w)[i])$ 
8:   end for
9:    $T_0 \leftarrow \text{Parallel\_Fill}(T_1, \rho, p, q_w)$ 
10:   $\text{Cardinal\_Sort}(T_0, p + q_w + 1, p + q_w, q_\rho)$ 
11:   $T_1 \leftarrow \emptyset$ 
12:  for  $i = k'$  to  $q_\rho$  do
13:     $T_1 \leftarrow \text{merge}(T_1, \text{Cardinal\_Sort}(T_0, p + q_w + 1, p + q_w + q_\rho)[i])$ 
14:  end for
15:  leer( $T_1$ )
16: end procedure
```

6.2 Diseño

7 Problema de la Mochila No Acotado

Problema: bajo las mismas condiciones del problema anterior, determinar si un subconjunto $B \subseteq A$ tal que $\rho(B) = \max\{\rho(C) : C \subseteq A \wedge \omega(C) \leq k\}$.

Una solución molecular es obtenida a partir del problema la mochila acotado, mas un cambio en la fase final donde se ordenan los tubos de salida con respecto a su valor y se elige sólomente el tubo con mayor valor.

7.1 Algoritmo

La nueva propuesta lleva a idear un programa molecular similar al previo mas se ha de notar las variaciones en la salida del programa al retornar sólo el tubo con mayor valor.

```
1: procedure Unbounded_Knapsack( $p, w, \rho, k, k'$ )
2:    $q_w \leftarrow \sum_{i=1}^p w(i); q_\rho \leftarrow \sum_{i=1}^p \rho(i); T_0 \leftarrow \text{Libería}(p + q_w + q_\rho, p)$ 
3:    $T_0 \leftarrow \text{Parallel\_Fill}(T_0, w, p, 0)$ 
4:    $T_0 \leftarrow \text{Cardinal\_Sort}(T_0, p + 1, p + q_w)$ 
5:    $T_1 \leftarrow \emptyset$ 
6:   for  $i = 1$  to  $k$  do
7:      $T_1 \leftarrow \text{mezclar}(T_1, \text{Cardinal\_Sort}(T_0, p + 1, p + q_w)[i])$ 
8:   end for
9:    $T_0 \leftarrow \text{Parallel\_Fill}(T_1, \rho, p, q_w)$ 
10:   $i = q_\rho; t = 0$ 
11:  while  $i \geq 1 \wedge t == 0$  do
12:     $T' \leftarrow \text{Cardinal\_Sort}(T_0, p + q_w + 1, p + q_w + q_\rho)[i]$ 
13:    if  $T' \neq \emptyset$  then
14:       $\text{leer}(T')$ 
15:       $t = 1$ 
16:    else
17:       $i = i - 1$ 
18:    end if
19:  end while
20: end procedure
```

8 Conclusión

9 Bibliografía

References

Roweis, Sam et al. “A Sticker-Based Model for DNA Computation”. In: *Journal of computational biology : a journal of computational molecular cell biology* 5 (Feb. 1998), pp. 615–29. DOI: 10.1089/cmb.1998.5.615.