INVESTIGACIÓN OPERATIVA

Segundo Cuatrimestre — 2012

Trabajo práctico

Fecha de entrega: 7 de diciembre

Genética Alienígena

GigaFarma, una de las compañías farmacéuticas más grandes del mundo, se encuentra actualmente experimentando con ADN alienígena. Su objetivo es producir una cadena de ADN alienígena que le reporte el mayor beneficio posible a la hora de comercializarla.

Una cadena de ADN alienígena puede verse como una secuencia no vacía de *genes*, conectados entre sí por uniones. A su vez, cada gen es una secuencia no vacía de bases. GigaFarma logró catalogar los genes que aparecen en las cadenas de ADN alienígena, ya que no toda secuencia de bases es un gen válido. Cada gen tiene un valor que depende de su funcionalidad, y cada cadena de ADN alienígena tiene un valor de mercado que es la suma de los valores de cada uno de sus genes. Para facilitar las cosas, vamos a representar a las diferentes bases con letras minúsculas, y a una unión con un guión medio (“-”). En el siguiente ejemplo podemos ver a la izquierda una posible lista de genes y sus valores; a la derecha aparecen algunas cadenas de ADN alienígena que pueden formarse a partir de ellos, con sus valores de mercado calculados.

hola	(5)	va-como-va	(9=2+5+2)
como	(5)	les	(3)
les	(3)	como-les	(8=5+3)
va	(2)	como-como-les	(13=5+5+3)
		como-como-como-les	(18=5+5+5+3)

Una cadena de ADN producible es una cadena de ADN que GigaFarma puede producir. En la práctica es una secuencia no vacía de *porciones* que la empresa puede sintetizar, sin conexiones adicionales de una a otra. Cada porción es una secuencia de bases y uniones, conteniendo al menos una unión, pero sin uniones iniciales, finales ni consecutivas. Cada porción tiene un costo que depende de la dificultad para producirla, y cada cadena de ADN producible tiene un costo de producción que es la suma de los costos de cada una de sus porciones. En el ejemplo que aparece a continuación podemos ver a la izquierda una posible lista de porciones y sus costos; a la derecha aparecen algunas cadenas de ADN producible que pueden formarse a partir de ellas, con sus costos de producción asociados.

como-co	(3)	como-co	(3)
mo-co	(8)	mo-les	(4)
mo-les	(4)	hasta-esta-no-sirven	(300=200+100)
como-como-les	(12)	como-como-les	(7=3+4 o 12)
ta-no-sirven	(100)	como-como-como-les	(15=3+8+4)
hasta-es	(200)		

Observar que puede existir más de una manera de formar algunas cadenas de ADN producible. Tal es el caso de “como-como-les” en el ejemplo, que puede obtenerse a partir de las porciones “como-co” y “mo-les” con costo de producción 7, o usando simplemente la porción “como-como-les” con costo de producción 12. Cuando existe más de una manera de sintetizar cierta cadena de ADN producible, GigaFarma siempre lo hace de alguna manera que tenga costo de producción mínimo.

El conjunto de cadenas de ADN alienígena es infinito, lo mismo que el conjunto de cadenas de ADN producible. Pero GigaFarma no está interesada directamente en ninguno de esos dos conjuntos, sino en su intersección. Si

Este problema fue usado originalmente para el TAP2012 (Torneo argentino de programación). Se puede encontrar una copia de los enunciados en <http://www.dc.uba.ar/events/icpc/download/problems/taip2012-problems.pdf>.

revisamos los ejemplos anteriores, podemos notar que “como-les” es alienígena pero no producible, “mo-les” es producible pero no alienígena, y “como-como-les” es alienígena y producible. Por cada cadena de ADN alienígena y producible, la compañía puede obtener una ganancia neta que es igual a su valor de mercado menos su costo de producción. Si esa ganancia neta no es positiva, la cadena nunca será producida. Dado que hay tanto material genético dando vueltas, GigaFarma pagaría cualquier cosa por saber cuál es la máxima ganancia neta positiva que puede obtener por alguna cadena de ADN alienígena y producible.

Entrada

Cada caso de prueba se describe utilizando varias líneas. La primera línea contiene dos enteros G y P que representan respectivamente la cantidad de genes y de porciones ($1 \leq G, P \leq 100$). Cada una de las G líneas siguientes describe un gen distinto mediante una cadena S y un entero V ; la cadena S tiene entre 1 y 10 caracteres, y esta formada únicamente por letras minúsculas que representan las bases del gen; el entero V indica el valor del gen ($1 \leq V \leq 1000$). Cada una de las P líneas siguientes describe una porción distinta mediante una cadena T y un entero C ; la cadena T tiene entre 1 y 30 caracteres, y esta formada únicamente por letras minúsculas y guiones medios (“-”) que representan respectivamente las bases y las uniones; esta cadena contiene al menos un guion, pero no tiene guiones iniciales, finales ni consecutivos; el entero C indica el costo de la porción ($1 \leq C \leq 100$). En cada caso de prueba los genes son todos diferentes, y las porciones son todas diferentes. El final de la entrada se indica con una línea que contiene dos veces el número -1 .

Salida

Para cada caso de prueba, imprimir en la salida una línea conteniendo un entero que representa la máxima ganancia neta positiva que GigaFarma puede obtener por una cadena de ADN alienígena y producible. Si ninguna ganancia neta es positiva imprimir el valor 0. Si la ganancia neta puede ser arbitrariamente grande imprimir un asterisco (“*”).

Ejemplo

Entrada de ejemplo	Salida para la entrada de ejemplo
4 6	6
hola 5	0
como 5	*
les 3	0
va 2	
como-co 3	
mo-co 8	
mo-les 4	
como-como-les 12	
ta-no-sirven 100	
hasta-es 200	
2 3	
xyz 1000	
zyxyz 1000	
xyz-zyx 1	
zyx-xyz 1	
xyz-xyz-zyx-xyz 1	
2 1	
abc 1	
abcabc 1000	
abc-abc 999	
1 1	
ser 10	
no-ser 5	
-1 -1	

Programación lineal

En todos los casos, entregar un informe detallado, con resultados y un análisis cualitativo y cuantitativo de los mismos.

2.1 El master

Realizar la selección del Master chileno utilizando los datos del último año del paper (2009). Considerar la misma cantidad de seleccionados y las mismas restricciones que en el paper. Implementar al menos 3 métodos distintos de selección (no necesariamente los mismos del paper) y un algoritmo que combine los resultados (no necesariamente el del paper).

2.2 El fixture de fútbol

Generar una propuesta de fixture para el Torneo Inicial 2012 del fútbol argentino. Considerar todas las restricciones usuales e incorporar restricciones adicionales que les parezcan razonables. Utilizar el modelo de patrones para simplificar la resolución del segundo modelo. Comparar los resultados con el fixture actual.

2.3 El fixture de voley

Generar una propuesta estilo TTP para armar la liga de Voley de Argentina 2012-2013. Considerar al menos 2 sets distintos de parejas (utilizar el modelo generador de parejas del paper, o alguno similar). Entregar una solución espejada y una solución no espejada para cada conjunto de parejas. Proponer además alguna función objetivo adicional, además de la clásica del TTP. Hacer un análisis comparativo de los distintos resultados (les quedaran al menos 8 fixtures distintos) y comparar también con el fixture actual.