



Introducción a la Biología Molecular Computacional IC-8022

Profesor:

Francisco Torres Rojas

Alineamiento con espacio lineal y cuadrático

Estudiantes:

Carlos Mauricio Girón Alas - 2014113159

Amanda Solano Astorga - 2013100025

31 de mayo del 2017

Índice

1. Introducción	1
2. Casos de Ejemplo:	2
2.1. Caso simple, hileras de tamaño 7	2
2.2. Caso con mayor consumo de recursos, hileras de 1000	3
2.3. Caso con grande, hileras de 5000	4
2.4. Caso en el que no corre ningún algoritmo	5
3. Conclusión	6

1. Introducción

El siguiente documento trata sobre los resultados del desarrollo de un pequeño sistema que implementa alineamiento global en sus versiones de espacio cuadrático y espacio lineal. El objetivo del sistema es presentar a manera de ejemplos aproximados a la vida real, las diferencias entre las versiones del algoritmo. Los resultados de varios casos de pruebas serán analizados en la conclusión.

2. Casos de Ejemplo:

2.1. Caso simple, hileras de tamaño 7

ecvsel

Ingreso de hileras

Ingrese el tamaño de la hilera V

Ingrese el tamaño de la hilera W

Hileras Generadas

Hilera V

Hilera W

Ingrese los pesos a usar:

Match - +

Mismatch - +

GAP Penalty - +

ecvsel

Datos del Needleman-Wunsch Cuadrático

Scoring

Tiempo

Complejidad Espacial

Alineamiento

G	C	T	-	G	T	G	C
A	T	T	C	G	C	G	-

Datos del Needleman-Wunsch Lineal

Scoring

Tiempo

Complejidad Espacial

Alineamiento

G	C	T	G	T	G	C	-
-	A	T	T	C	G	C	G

2.2. Caso con mayor consumo de recursos, hileras de 1000

ecvsel

Ingreso de hileras

Ingrese el tamaño de la hilera V

Ingrese el tamaño de la hilera W

Hileras Generadas

Hilera V
AGCGTAAGCGCCAGTCAGTTCGGAACATGCCGCACGTCCGTACTGGGGGAATAACTCTAA/

Hilera W
GGACCCCACTTCGACGCGCCAAAAGTCCTAGTAGGCTTCGCGCCGATTGGGGAGGGATCCCTI

Ingrese los pesos a usar:

Match - +

Mismatch - +

GAP Penalty - +

ecvsel

Datos del Needleman-Wunsch Cuadrático

Scoring

Tiempo

Complejidad Espacial

Alineamiento

_	_	A	G	C	G	T	A	_	A	G	C	G	_	C	C	A	G	T
G	G	A	C	C	C	C	A	C	T	T	C	G	A	C	G	C	G	C

Datos del Needleman-Wunsch Lineal

Scoring

Tiempo

Complejidad Espacial

Alineamiento

A	G	_	C	G	T	A	A	G	_	_	C	G	C	C	_	A	G	T
G	G	A	C	C	C	C	A	C	T	T	C	G	A	C	G	C	G	C

2.3. Caso con grande, hileras de 5000

ecvsel

Ingreso de hileras

Ingrese el tamaño de la hilera V

Ingrese el tamaño de la hilera W

Hileras Generadas

Hilera V

GCTGTGCATTGCGGGCACAAGAGTCCCGGTCCCTGTAGCTTTGATCAGCTCGAATCCATTAA

Hilera W

GCGGCATAGAGATGACCATCCCTTCAGTCCTTAAACACCGCTAGAACGAGTGCAGGGATTTC

Ingrese los pesos a usar:

Match - +

Mismatch - +

GAP Penalty - +

ecvsel

Datos del Needleman-Wunsch Cuadrático

Scoring

Tiempo

Complejidad Espacial

Alineamiento

G	C	T	G	T	G	C	A	T	T	C	G	C	G	G	C	A	C	A
G	C	-	G	-	G	C	A	-	T	A	G	-	A	G	-	A	T	G

Datos del Needleman-Wunsch Lineal

Scoring

Tiempo

Complejidad Espacial

Alineamiento

G	C	T	G	T	G	C	A	T	T	C	G	C	G	G	C	A	C	A
G	C	-	G	-	G	C	A	-	T	A	G	-	A	G	-	A	T	G

2.4. Caso en el que no corre ningún algoritmo

The screenshot shows the 'ecvsel' application window. Under the 'Ingreso de hileras' section, the 'Ingrese el tamaño de la hilera V' field contains '100000000000' and the 'Ingrese el tamaño de la hilera W' field contains '100000000000'. A 'Generar Hileras' button is visible. Below, the 'Hileras Generadas' section has two empty text boxes labeled 'Hilera V' and 'Hilera W'. The 'Ingrese los pesos a usar:' section has three rows: 'Match' with value '1', 'Mismatch' with value '-1', and 'GAP Penalty' with value '-2'. Each row has minus and plus buttons. At the bottom is an 'Iniciar algoritmos' button.

This screenshot shows the same 'ecvsel' application window, but with an error dialog box overlaid. The dialog box has a title bar 'ecvsel' and contains the text 'Las hileras no se pueden generar' with an 'Aceptar' button. The background interface is partially obscured by the dialog box.

3. Conclusión

Como se puede notar en los casos mostrados, la tendencia es a que la cantidad de espacio utilizado por el algoritmo de espacio cuadrático sea una orden de magnitud mayor que la cantidad de espacio utilizado por el algoritmo de espacio lineal.

Sucede algo similar pero de manera inversa con la tendencia del tiempo. Como es de esperar, debido a las alteraciones al algoritmo original y debido a las complicaciones producto de buscar el uso de menos memoria, el tiempo que tarda el algoritmo lineal es considerablemente mayor al tiempo que tarda el cuadrático.

En nuestro caso, ya que la máquina utilizada para las pruebas poseía 16GB de memoria RAM, el único aproximado (debido a lo que tardaba el sistema en generar respuesta o error) en el que se logró hacer que el algoritmo cuadrático dejara de ser factible, fue con hileras de tamaño: 1 000 000 000. El algoritmo lineal dejó de ser factible con hileras de tamaño 100 000 000 000.