

Instituto Superior Manuel Teixeira Gomes

Licenciatura de Engenharia Informática, 2º Semestre

Linguagens de Programação I

Prof. Doutor António Anjos

Programa proposto sobre Genética

Trabalho elaborado por:

Carlos Manuel Pacheco Soares, Nº 21103408

Luciano Higinio Teixeira Neves, Nº 21100341

Portimão, 20 de Janeiro de 2012

Índice

Introdução	2
Livrarias utilizadas.....	2
Estrutura base do programa	3
Versão 1	3
Versão 2	3
Dificuldades	5
Conclusão	7

Introdução:

No âmbito da disciplina de Linguagens de Programação I, irei tentar explicar o funcionamento do programa que criamos, proposto pelo professor, utilizando o compilador gcc.

O trabalho proposto foi, em suma:

“1. Pretende-se um programa que, a partir de um ficheiro, que contenha a sequência de uma das duas cadeias de DNA que permita a criação de:

1. A hélice complementar;

2. A sequência RNA;

3. A conversão em proteínas.

2. Após apresentação de um menu e seleção da opção por parte do utilizador, o programa deverá solicitar o nome do ficheiro de entrada (DNA) e o nome do ficheiro de saída que irá conter o resultado relativo à opção escolhida. Os resultados não são para apresentar no ecrã.”

Para proporcionar um maior rigor no nosso programa decidimos criar duas versões diferentes: uma que irá fazer exatamente o que foi proposto para dar resultados idênticos aos dados pelo professor, que será a versão 1 (v1) e outro utilizando conceitos de Biologia que consideramos mais corretos, que será a versão 2 (v2).

Livrarias utilizadas:

Para a realização do programa em causa utilizei as três livrarias permitidas, as quais são:

```
#include <stdio.h>
```

```
#include <stdlib.h>
```

```
#include <string.h>
```

Estas livrarias são livrarias padrão do C.

Estrutura base do programa:

Para separar o que foi proposto e as melhorias introduzidas (principalmente no que diz respeito à disciplina de Biologia) dividimos o trabalho em duas versões, criando dois programas diferentes para cada uma delas (versão 1 e versão 2).

Versão 1:

A versão 1 consiste, basicamente, em criar um programa que funcionasse da forma como foi proposto nos apêndices. Criando um menu dentro de um ciclo, só terminando o mesmo quando selecionada a opção para sair. O menu na primeira opção cria um ficheiro com a cadeia complementar à de DNA, a qual é retirada de um ficheiro escolhido pelo utilizador. Na segunda opção cria-se um ficheiro com uma cadeia de RNA a partir da de DNA utilizada e, por fim, na terceira opção cria-se uma proteína baseada na cadeia de RNA utilizada, ignorando os conceitos de codões start/stop e omitindo os codões que não reconhece.

Versão 2:

Na versão 2 existem algumas melhorias, focando-nos de uma forma mais aprofundada relativamente aos conceitos de Biologia. A maneira como é criado o menu é idêntica e a primeira opção do menu é igual. A segunda opção varia bastante, pois uma cadeia de DNA tem um início e um fim, representando-o normalmente da seguinte forma:

5' - ATGCGATGAATC - 3'

Isto tem a ver com as ligações covalentes das moléculas (não queremos entrar muito em pormenor). Assim, como este tipo de cadeia é considerada a cadeia base de cada gene, iremos também considerar a cadeia inserida no ficheiro escolhido pelo utilizador como uma cadeia base.

Então a nossa cadeia de DNA complementar será a seguinte (aproveitando o exemplo anterior):

3' - TACGCTACTTAC- 5'

Porque quando uma cadeia começa por 5' e acaba em 3' então qualquer cadeia que é complementar a esta começa por 3' e acaba em 5'.

Mas cada cadeia de mRNA, isto é, o RNA mensageiro (o tipo de RNA que nos interessa para o nosso programa) começa em 5' e acaba 3', então não o podemos formar a partir do DNA, mas temos de o formar pelo DNA complementar. O qual irá resultar no seguinte:

5' - AUGCGAUGAAUC - 3'

Obtemos o mesmo resultado quando utilizamos logo a cadeia de DNA principal e substituímos os 'T's por 'U's.

No entanto, esta não é a única modificação que introduzimos, na terceira opção do menu também existem inúmeras melhorias no programa.

Indo pelos conceitos da Biologia, uma proteína é obtida a partir do processo de tradução de uma cadeia de mRNA, do qual já vimos anteriormente como a obtemos. Esta cadeia irá se ligar a uma ribossoma, o qual vai permitir a ligação de cada anticodão de tRNA a cada codão da cadeia de mRN. Cada anticodão ou codão é constituído por três aminoácidos, mais especificamente, três nucleótidos. Como a cada anticodão está ligada uma molécula de aminoácidos, pode-se criar uma relação direta entre os codões e os aminoácidos, isto é, o chamado código genético. Mas para não fazer este processo todo podemos para simplificar proceder da seguinte maneira: se conseguimos obter a cadeia de mRNA substituindo os 'T's de uma cadeia principal de DNA por 'U's, então, utilizando a matriz fornecido conseguimos, substituindo no código genético, (neste caso na matriz) os 'U's por 'T's. Isto facilita imenso o processo, no entanto, é preciso dizer que isto só é possível ao desprezar os conceitos de intrões e exões.

Mas cada polipéptido (ou proteína), que é uma sequência de aminoácidos, tem um início e um fim, para tal ser possível é necessário utilizar os codões de iniciação e os de finalização. O codão que representa o aminoácido Metionina (Met/M) é quase sempre o codão de iniciação e três outros codões, os quais tivemos de introduzir na matriz (tem escrito entre parenteses “stop”), de finalização.

Deste modo, na terceira opção do menu do programa iremos procurar na sequência de DNA do ficheiro selecionado o primeiro codão de iniciação, fazendo a correspondência de cada codão a partir desse codão inclusivo, com os aminoácidos que irão formar a proteína até se encontrar um codão de finalização. Cada tradução irá formar somente uma proteína.

Dificuldades:

De um modo geral, não considerámos ter muitas dificuldades na criação do programa, surgindo grande parte delas ao criar a terceira opção do programa. Isto porque sentimos certas dificuldades em colocar o algoritmo a funcionar corretamente. Por fim, podemos dizer que uma das maiores dificuldades foi ao criar este relatório, uma vez que nunca tivemos uma explicação minimamente pormenorizada de como o fazer.

Conclusão:

Enriquecendo os nossos conhecimentos acerca da linguagem de programação C e, ao mesmo tempo, adquirindo uma certa experiência.

Como ultrapassamos facilmente as dificuldades com que nos deparamos, consideramos que este projecto proposto foi uma mais-valia.

Por fim, consideramos também que, uma vez que o manuseamento de ficheiros é uma função importante e muito útil na programação, este trabalho contribuiu muito para nos auxiliar em futuros programas.