Concurrent Programming Languages

Channel-based Concurrency Module
Lab 3: Project

26 October 2021

MIEI - Integrated Masters in Comp. Science and Informatics
Specialization Block

Bernardo Toninho
(with António Ravara and Carla Ferreira)

NOVALINCS

Course Infrastructure

- We are going to use GitHub Classroom for the project
- The project handout is available here:

https://classroom.github.com/a/-ZTUtS0d

- Sign up, teams of 2 people as before.
- Deadline on 8/11/2021 23:59. Have fun.

Protein Synthesis

Protein synthesis in cells occurs in two steps:

- DNA Transcription (from DNA to mRNA template)

- mRNA Translation (from mRNA to an amino acid chain)

In this project you will implement these two operations, concurrently.

Transcription

Going from DNA to mRNA:

- DNA is a molecule composed of two chains of nucleotides.
- Each nucleotide is composed of one of four nucleobases: **C**ytosine, **G**uanine, **A**denine and **T**hymine.
- DNA chain ~ a string of **C**s, **G**s, **A**s and **T**s.
- DNA is transcribed into an mRNA molecule out of so-called complementary nucleotides.
- The RNA transcript carries the same information as the DNA chain, but where the base Thymine (T) is replaced by the RNA-only base, **U**racil (U).

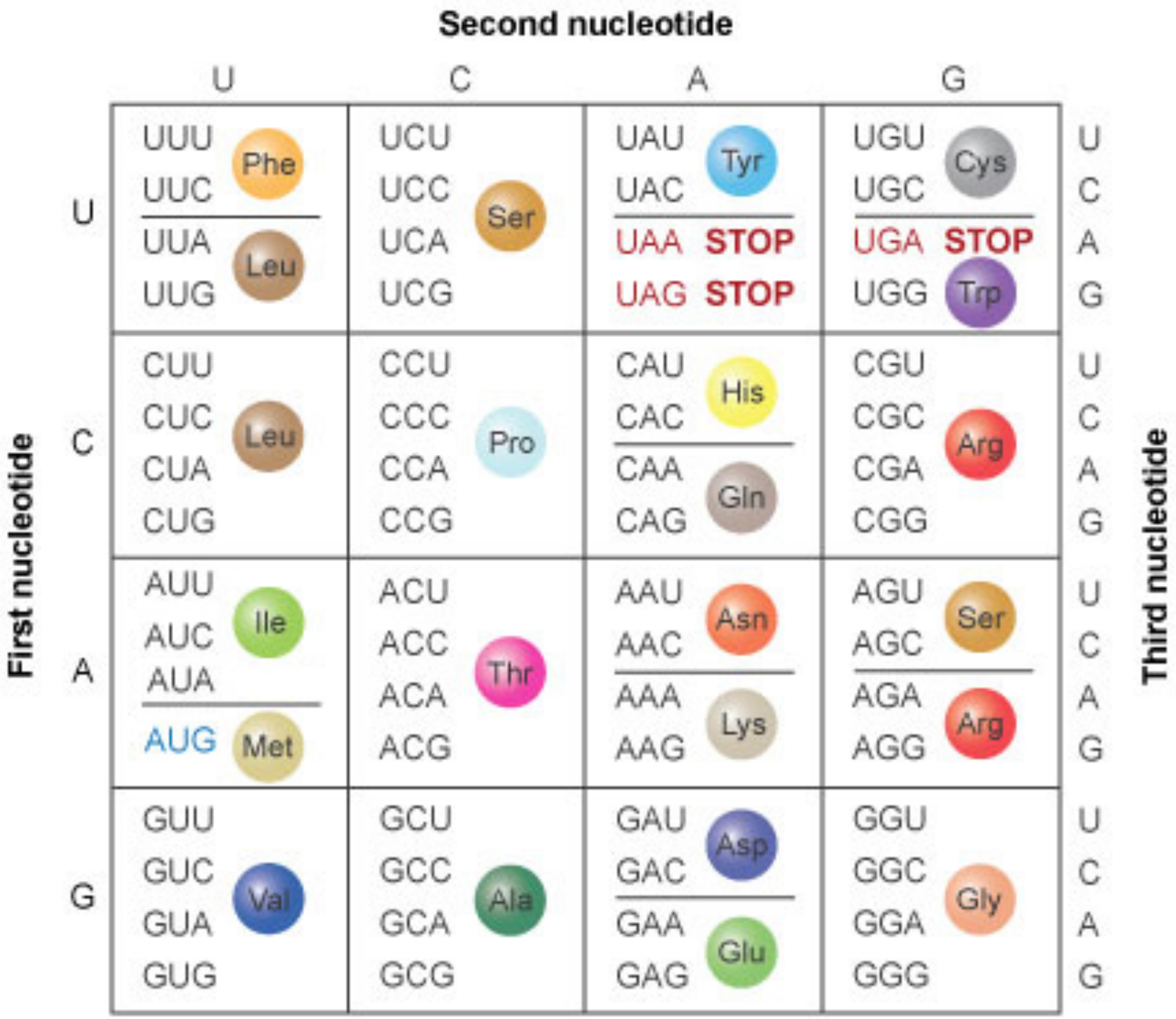
Transcription

DNA Chain: GGGCCG**T**C**TT**C**TT**CG**TT**AAA

mRNA Transcription: GGGCCG**U**C**UU**C**UU**CG**UU**AAA

Translation

- Amino acids are characterized by **3** mRNA nucleotide bases (a **codon**)



Second nucleotide

First nucleotide	U	C	A	G	Third nucleotide
U	UUU UUC Phe UUA UUG Leu	UCU UCC UCA UCG Ser	UAU UAC Tyr UAA STOP UAG STOP	UGU UGC Cys UGA STOP UGG Trp	U C A G
C	CUU CUC CUA CUG Leu	CCU CCC CCA CCG Pro	CAU CAC His CAA CAG Gln	CGU CGC CGA CGG Arg	U C A G
A	AUU AUC AUA Ile AUG Met	ACU ACC ACA ACG Thr	AAU AAC Asn AAA AAG Lys	AGU AGC Ser AGA AGG Arg	U C A G
G	GUU GUC GUA GUG Val	GCU GCC GCA GCG Ala	GAU GAC Asp GAA GAG Glu	GGU GGC GGA GGG Gly	U C A G

- GUU translates to Valine (Val)

- AUG translates to Methionine (Met)

Translation

- Amino acids are characterized by **3** mRNA nucleotide bases (a **codon**)

- For the resulting amino acid chain to form a valid protein, it must adhere to the following (simplified) rules:

  - its first amino acid must be Methionine (Met)

  - it must be ended by one of the three termination codons (UAA, UAG or UGA).

Assignment

- Implement a concurrent protein synthesis analyzer!
- Your analyzer must be able to answer two kinds of queries:
    - Given a DNA sequence, what is its transcription into mRNA.
    - Given a DNA sequence, what is its corresponding amino acid chain (and does that chain form a valid protein).

System Spec

- Two software components, a **client** and a **server**. Each component must compile to a separate binary.

- **Client**:

  - Sends user-specified requests to the server, receives and prints the result to standard output.

  - Client and the server communicate using a standard TCP socket.

  - If a client loses contact with the server, it should print a “Disconnected from server” message to standard output and terminate.

  - Message manipulation code provided (package **`message`**)

System Spec

- Two software components, a **client** and a **server**. Each component must compile to a separate binary.
- **Server**:
  - Receives requests from **one or more** clients, performs the corresponding transcription or translation (and validation) request and answers back.
  - The server **must** be able to handle client requests concurrently.
  - The partition of the request into jobs and their aggregation into a final answer is relatively straightforward, but there are some design choices to think about:
    - Order of the DNA and the transcribed mRNA matters, and so the aggregation must preserve the original order of the DNA sequence.
    - For **transcription**, the operation is at the single nucleotide level.
    - For **translation**, the operation is at the codon level (but requires transcription).
    - Think on how to divide the request such that there is the **most potential for parallelism** across the combined transcription-translation steps.

Requirements

- The server must use multiple workers to perform the client jobs, but of a bounded number. The focus is on concurrency, but aim to *maximize* the potential for parallelism.

- Choosing a good way of organizing (and creating) workers, suited to the specific context of the problem is part of the challenge.

- Choosing a good way of aggregating the work is also part of the challenge, considering the order of the sequence must be preserved.

- **Correctness is mandatory.**

- The server should try to **minimize mean response time** for client requests.

Grading

- You must turn in your code and a brief report (add a PDF to the repository). Standard plagiarism rules apply (see Univ. and Dept. CoC).

- Your code **must not** use locks or mutexes.

- Your project will receive a better grade according to the following (unordered) criteria:

    - **Correctness**: Does the server answer correctly? (Critical!)

    - **Worker creation:** When and how are workers created?

    - **Worker management:** How are workers organized amongst themselves and in the overall system?

    - **Data management**

    - **Fairness and responsiveness**

    - **Overall code quality**

    - **Use of tests to validate your code**