

Lab Assignment Week 09

CSC/DSCI 1301 – Principles of CS/DS I

Week of March 4th, 2024

Introduction

Welcome to the ninth programming lab of CSC/DSCI 1301! Today we will be covering the following topics:

- Using Default Argument Values
- Using Arbitrary Arguments
- String Methods

Lab policy reminders:

- Attendance is mandatory.
- Labs must be completed **individually**.
- TAs are here to help you. Ask them for help!
- Lab assignments are due at the end of each lab.

Deliverables:

1. Python files for your DNA translation program.
 - a. dna_translation.py
 - b. helper.py
2. Screenshots of your program output

If you have any questions, please do not hesitate to ask your TA!

Program: dna_translation.py

Computers are extremely efficient at storing and manipulating large data. For this Lab, you will need to write a program that transcribes a DNA sequence into mRNA, then translates that mRNA sequence into the sequence of amino acids that make up the protein. You will need to write two functions, `transcription()` and `translation()`, to accomplish this task. Your program should also print out the sequence after each step.

DNA to mRNA

The `transcription()` function takes a DNA sequence as its input, transcribes it into the complementary mRNA sequence and returns the mRNA sequence as its output. You will need to program your `transcription()` function to perform two tasks:

1. Replace the Nucleotide Thymine (T) with Uracil (U).
2. Create the Base Pair Complement of the sequence.

The rules for creating the Base Pair Complement are shown in the table below.

RNA Base Paring Rules Table

Nucleotides	Codons
Adenine to Uracil	A-U
Uracil to Adenine	U-A
Cytosine to Guanine	C-G
Guanine to Cytosine	G-C

Hints

- The built-in String method `replace()` will replace the Thymine (T) with Uracil (U) DNA string.
 - The `replace()` method returns the string after replacement.
 - You can read more about this method [here](#).

mRNA to Protein

Protein translation comprises two steps. The mRNA is broken up into nucleotide triplets. The nucleotides that comprise the triplet determine which amino acid make up that part of the protein.

The translation() function will take in the mRNA sequence as input, translate it into a sequence of amino acids, and return the amino acid sequence as its output. Your translation() function will need to do the following:

1. Separate the mRNA sequence into nucleotide triplets.
 - a. **Note: There is no built-in method that does this. You will have to write a helper function that separates a string into triplets. See Chunk Function Section.**
2. Look up the amino acid that corresponds to the nucleotide triplet.

Chunk Function

There is no native function that splits a string into triplets. You will need to write this helper function yourself. You can find the start of the function definition in the helper.py file on iCollege. This function will take a string and chunk size as its input. And return a list of strings of the inputted chunk size. Your chunk function may assume that the length of the input string will always be a multiple of the chunk size.

Example Input

Chunk Size = 3

A	A	C	U	G	A	U	G	G
---	---	---	---	---	---	---	---	---

Example Output

A	A	C
U	G	A
U	G	G

Hint

- Start with an empty string and use the built-in String method append() to append each of the chunks to the output list you will return.

Nucleotide to Amino Acid Chart

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

Hint

- Store your Nucleotide to Amino Acid Chart as a Dictionary. This dictionary is already provided for you in the helper.py file.
 - Nucleotide triplet as the key.
 - Amino Acid Abbreviation as the Value.

Example Output

DNA Sequence

TACGCAGAAAAAATCAGCGGGGTTGTTGGTCATTAGTCTGAATT

mRNA Sequence

AUGCGUCUUUUUUUAGUCGCCCCAACAACCAAGUAAUCAGACUUAA

Protein Sequence

Met Arg Leu Phe Leu Val Ala Pro Thr Thr Ser Asn Gln Thr STOP

Skills Covered

- Creating User-defined Functions
- String Methods

Deliverables

For this program you will need to provide the python file containing your code as well as a screenshot of the output of your program. Please name your files as follows:

- Python Files
 - lastname_firstname_filename.py
 - For example: **hawamdeh_faris_dna_translation.py**
- Screenshots
 - lastname_firstname_filename.png
 - For example: **hawamdeh_faris_dna_translation.png**