Example 2 - UAI 655 (less advanced)

## Exam1 – oral part – summer 2018

|  |  |
| --- | --- |
| Name of student |  |
| Date |  |

# Create the Jupyter (ipynb) program in the /mnt/exam/exam1/”your\_name” and name it “example2-yourname.ipynb”. Put your name on the top of your program as a comment.

# Put the following tasks into your program.

# Create protein fasta files - start codons and stop codons

# Small theory

# About stop and start codons see more [https://en.wikipedia.org/wiki/Genetic\_code#RNA\_codon\_table](https://en.wikipedia.org/wiki/Genetic_code" \l "RNA_codon_table) <https://en.wikipedia.org/wiki/DNA_codon_table>

# <https://www.ncbi.nlm.nih.gov/orffinder/>

# 

# 

# **From UniProt database download**

* Put both id and
* sequence to variables for later comparison

# Requirements

# **Create Jupyter program which**

# Read nucleotide fasta files

# */mnt/hepatitis\_b\_genome.fasta*

# */mnt/ebov\_2014.fasta*

# */mnt/ebov\_2007.fasta*

# Take into clipboard the codontable */mnt/codontable.txt* and put it into your program

* From nucleotide fasta file create (print on the screen or write in a disk) protein fasta file. The header take as it is. Replace the sequence.
  + Find start\_codon in sequence start\_codon = ["ATG"]
  + Find one of stop\_codon stop\_codon = ["TAG","TAA","TGA"]
* All between the start\_codon and the first stop\_codon replace according to the codontable. The stop\_codons are “-” in the codontable, omit them.
* Continue till the end of the nucleotide fasta file.
* Consider that the fastafile can be multiple (more headers and more sequencies