

BioBIKE Syntax Problem Set 3 - Loops

1. Write a loop that says hello 10 times.
2. Write a loop that prints the name of every organism, each on a separate line.
(Key function: *DISPLAY-LINE*)
(Key constant: **all-organisms**)
3. Write a loop that prints the genome size of every organism, next to the name of the organism.
(Key function: *LENGTH-OF*)
(Key constant: **tab**)
4. Write a loop that calculates the total number of nucleotides known by BioBIKE.
5. Write a loop that calculates the probability of encountering a given nucleotide sequence (e.g. "CCGATCGG") in a genome with $[A] = [C] = [G] = [T]$, i.e. all frequencies 25%. Start off as follows:

```
(ASSIGN sequence = "CCGATCGG") ; or any other nucleotide sequence
(FOR-EACH letter IN sequence
  INITIALIZE product = fill-in
  AS product = fill-in
  FINALLY (RETURN fill-in ) )
```

6. Write a loop that calculates the probability of encountering a nucleotide sequence in a genome with $[A] = 0.3$. The following template may be helpful:

```
(FOR-EACH letter IN "fill-in-sequence"
  WITH fill-in
  AS A% = 0.3
  AS C% = fill-in
  AS G% = fill-in
  AS T% = fill-in
  (IF-TRUE (EQUAL letter "A")
    THEN fill-in)
  (IF-TRUE (EQUAL letter "C")
    THEN fill-in)
  . . .
  FINALLY (RETURN fill-in))
```

7. Write a loop that calculates the average length of a protein in SS120.
8. Write a loop that calculates and displays the name of each organism followed by the average length of a protein in that organism (this might take some 10's of seconds to execute)

9. Explain and fix the following:

- 9a. FOR-EACH i FROM 1 TO 10
 (DISPLAY-LINE "HELLO")
- 9b. (FOR-EACH i FROM 1 TO 10)
 (DISPLAY-LINE "HELLO"))
- 9c. (FOR EACH i FROM 1 TO 6
 SUM i)
- 9d. (FOR-EACH (FROM 100 TO 1 BY -1)
 SUM n)
- 9e. (FOR-EACH gene IN (GENES-OF ss120)
 AS gene size = (LENGTH-OF gene)
 DO collect gene size)
- 9f. (FOR-EACH (GENES-OF ss120)
 COLLECT length)
- 9g. (FOR-EACH gene IN GENES-OF ss120
 (COLLECT (LENGTH-OF gene)))
- 9h. (FOR-EACH gene IN (GENES-OF ss120)
 COLLECT (LENGTH-OF gene)
 SUM (LENGTH-OF gene))
- 9i. (FOR-EACH n from 1 to 100
 INITIALIZE sum = 0
 (+ n sum))
- 9j. (FOR-EACH n from 1 to 100
 INITIALIZE sum = 0
 (INCREMENT sum BY n))
- 9k. (FOR-EACH gene IN ss120
 DISPLAY LENGTH-OF gene)
- 9l. (FOR EACH gene IN (GENES-OF ss120)
 DO (LENGTH-OF gene)
 COLLECT length)
- 9m. (FOR-EACH gene IN (GENES-OF ss120)
 INITIALIZE atg-count = 0
 INITIALIZE non-atg-count = 0
 AS start-codon = (SEQUENCE-OF gene FROM 1 TO 3)
 DO (IF-TRUE (SAME start-codon "atg")
 THEN (INCREMENT atg-count)
 ELSE (INCREMENT non-atg-count)
 FINALLY (COLLECT (LIST atg-count non-atg-count))))
- 9n. (FOR-EACH gene IN (GENES-OF ss120)
 INITIALIZE length = (LENGTH-OF gene)
 COLLECT length)

; This one is a bit subtle