1. Genomic texts

- The cell, atom of the living world
- At the heart of the cell: the DNA macromolecule
- DNA codes for genetic information
- What is an algorithm?
- **Counting nucleotides**
- GC and AT contents of DNA sequence
- DNA walk
- Compressing the DNA walk
- Predicting the origin of DNA replication?
- Overlapping sliding window

Counting nucleotides

The input of our algorithm

AGCTTTTCATTCTGACTGCAACGGGCAATATGTCTCTGTGTGGATTAAAAAAAGAGTGTCTGATAGCAGC*

```
nbA,nbC,nbG,nbT, TotalNb, index: integer
sequence: character string [1:*]

nbA,nbC,nbG,nbT, TotalNb ← 0
index ← 1
```

declaration of variables

initialization; assignments of values

```
nbA,nbC,nbG,nbT, TotalNb, index: integer
sequence: character string [1:*]

nbA,nbC,nbG,nbT, TotalNb ← 0
index ← 1

repeat
  case sequence [index] of
    "A": nbA ← nbA + 1
    "C": nbC ← nbC + 1
    "G": nbG ← nbG + 1
    "T": nbT ← nbT + 1
  endcase
  TotalNb ← TotalNb + 1
  index ← index + 1
until sequence [index] = "#"
```

control instructions

incrementation of counters

condition

```
nbA,nbC,nbG,nbT, TotalNb, index: integer
sequence: character string [1:*]

nbA,nbC,nbG,nbT, TotalNb ← 0
index ← 1

repeat
  case sequence [index] of
    “A”: nbA ← nbA + 1
    “C”: nbC ← nbC + 1
    “G”: nbG ← nbG + 1
    “T”: nbT ← nbT + 1
  endcase
  TotalNb ← TotalNb + 1
  index ← index + 1
until sequence [index] = “#”
display “Length of the sequence:” TotalNb
display “%A=” (nbA/TotalNb)*100, “  %C=“, (nbC/TotalNb)*100, “  %G=”, (nbG/
TotalNb)*100, “  %nbT=”, (nbT/TotalNb)*100
```