FACULDADE ENGENHARIA

Departamento de
Informática

Bioinformática
(15861, 13422)

Algorithms! Algorithms Everywhere...

- Objectives for today's lesson:
    - Understand the concept of algorithms.
    - Explore the importance of algorithms in biotechnology.
    - Introduce real-world applications of algorithms in the field.
    - From an algorithm to a program.

- An *algorithm* is a sequence of instructions that one must perform in order to solve a well-formulated *problem*.

- Analogy: Cooking recipe — Just as each step in a recipe must be followed to achieve the desired dish, an algorithm is a recipe for solving problems.

Pumpkin soup

🍴 2 servings   🕐 15 minutes

INGREDIENTS

100 ml milk
50 g butter
3 eggs
1 tbs cocoa
2 tsp baking soda
a pinch of salt
3 eggs

DIRECTIONS

1. Nunc nulla velit, feugiat vitae ex quis, lobortis porta leo.
2. Donec dictum lectus in ex accumsan sodales. Pellentesque habitant morbi tristique.
3. Nunc nulla velit, feugiat vitae ex quis, lobortis porta leo. Donec dictum lectus in ex. lentesque habitant morbi tristique. Nunc nulla velit, feugiat vitae ex quis, lobortis porta leo. Donec dictum lectus in ex.
4. Habitant morbi tristique.Nunc nulla velit, feugiat vitae ex quis, lobortis porta leo. Donec dictum lectu,
5. Donec dictum lectus in ex accumsan sodales. Pellentesque habitant morbi tristique.
6. Nunc nulla velit, feugiat vitae ex quis, lobortis porta leo. Donec dictum lectus in ex. lobortis porta leo.

NOTES

Nunc nulla velit, feugiat vitae ex quis, lobortis porta leo. Donec dictum lectus in ex accumsan sodales. Pellentesque habitant morbi tristique.

- An *algorithm* is a sequence of instructions that one must perform in order to solve a well-formulated *problem*.

- We may specify problems in terms of their *inputs* and their *outputs*, and the algorithm will be the method of translating the *inputs* into the *outputs*.

- **Do you know?** Nature uses algorithm-like procedures to solve biological problems, for example, in the process of DNA replication (Biological Algorithms).

- Algorithms are everywhere, often working behind the scenes in ways we don't even realize!

- Algorithms are everywhere, often working behind the scenes in ways we don't even realize!

  - <u>GPS Navigation</u>: Just as a GPS guides us to the best route, algorithms steer us toward the most efficient solutions.

  - <u>Search Engines</u>: Algorithms sift through millions of web pages to find the information we seek, much like searching for the perfect recipe in a cookbook.

  - <u>*Social Media Feeds*</u>: Algorithms determine what shows up in our news feeds based on our likes and interactions, similar to how a chef tailors a dish to a customer's taste preferences.

Do you can identify other *algorithms* in everyday life? Provide examples, please!!!

- Algorithms simplify complex <u>decisions</u> and <u>daily tasks </u> by breaking them down into manageable steps, streamlining processes and predicting outcomes.
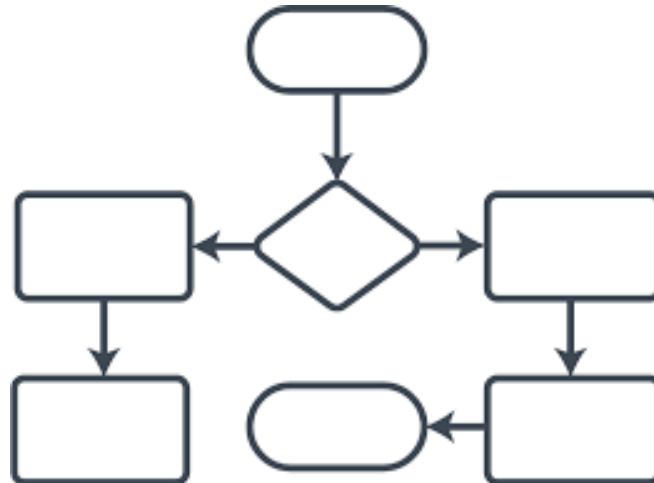
- Enhance efficiency by optimizing tasks.
- Improve accuracy in data analysis and predictions.
- Save time by automating repetitive processes.

- Handling vast amounts of biological data.
- Making precise predictions and decisions.
- Enhancing the speed and accuracy of research outcomes.

- DNA Sequencing and Alignment
  - DNA sequencing is the process of determining the exact order of the four bases — adenine (A), guanine (G), cytosine (C), and thymine (T) — in a DNA molecule.

- # DNA Sequencing and Alignment
  - ## Algorithms are used to align sequences to find similarities and differences, which is essential for genetic analysis, evolutionary studies, and identifying genetic disorders.

Unaligned sequences

Human	ACATTATGGACAGGTAAGTAAAAAACATATT
Chimpanzee	ACATTATGGACAGGTAAGTAAAAAACATATT
Macaque	ATATACATTACGGACAGGTAAGTAAAAACAT

Aligned sequences

Human	ACA	TTATGGACAGGTAAGTAAAAAACATATT
Chimpanzee	ACA	TTATGGACAGGTAAGTAAAAAACATATT
Macaque	ATATACATTACGGACAGGTAAGTAAAAACAT

- # DNA Sequencing and Alignment
  - ## The Smith-Waterman algorithm, a form of dynamic programming, is commonly used for local DNA sequence alignment.

 - ## *Finding the most similar pieces in a puzzle.*

Initialize the scoring matrix

	T	G	T	T	A	C	G	G
0	0	0	0	0	0	0	0	0
G	0							
G	0							
T	0							
T	0							
G	0							
A	0							
C	0							
T	0							
A	0							

Substitution matrix: $S(a_i, b_j) = \begin{cases} +3, & a_i = b_j \\ -3, & a_i \neq b_j \end{cases}$

Gap penalty: $W_k = kW_1$
$W_1 = 2$

- Protein Structure Prediction
  - Protein folding is the physical process by which a protein chain acquires its native 3D structure, a conformation that is biologically functional.

Symmetric fragment assembly

+

Symmetric rigid body moves

Symmetric all-atom refinement
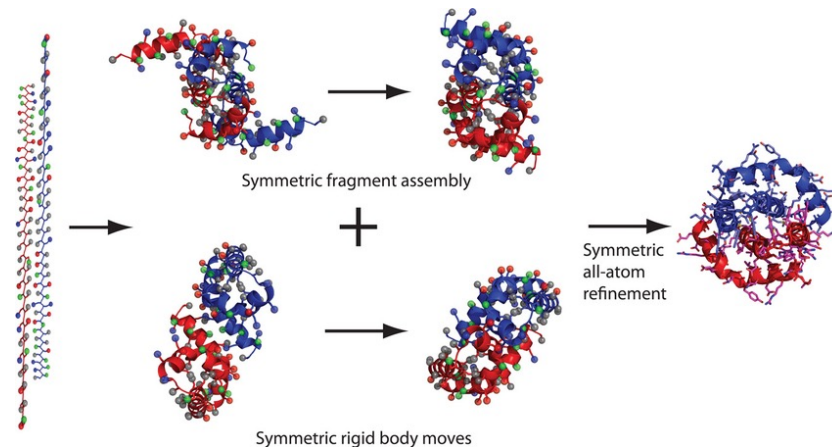
- Protein Structure Prediction

  - Predicting protein structures requires algorithms that can simulate folding pathways and determine the lowest energy conformation among countless possibilities.

- Protein Structure Prediction
  - *How predicting the 3D structure of proteins can reveal their function*
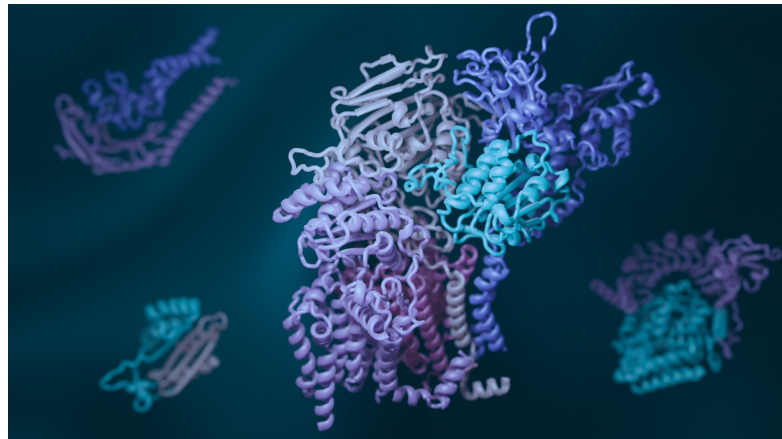
- Gene Editing (CRISPR-Cas9)
  - CRISPR-Cas9 is a groundbreaking technology that allows for the editing of genes within organisms in a precise manner.

- Gene Editing (CRISPR-Cas9)
  - Algorithms help design the guide RNA sequences that target the right spot in the genome for editing, minimizing off-target effects
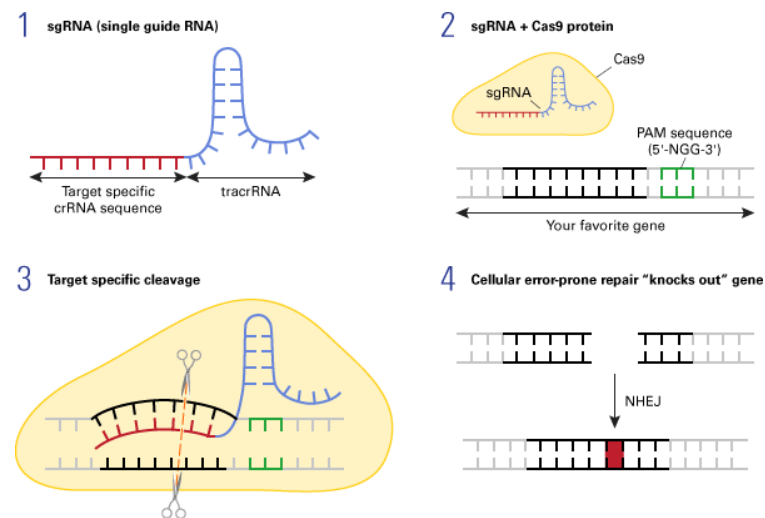
- Gene Editing (CRISPR-Cas9)
  - *As editing text in a document — finding and replacing specific letters (sequence)*

- Can you think of a problem in your daily life that could be solved with an algorithm? How would you approach it?

- Instructions: Think individually for 2 minutes, discuss with a partner for 3 minutes, be ready to share with the class.

- Propose an algorithm to play the game of "Tic-Tac-Toe"

- Do not forget to:
  - Identify the context
  - Show the sequential steps developed to tackle this problem
  - Describe how the algorithm may be successfully applied and its impact

- An *algorithm* can be expressed in different manners, but additionally, it must address the following characteristics:
    1. Unambiguity/clearness
    2. Effectiveness
    3. Finiteness/termination
    4. Correctness.

- <u>Unambiguity and clearness</u>: this means that the algorithm details exactly what needs to be detailed to solve the problem at hands, in a clear, not ambiguous manner.

- <u>Effectiveness</u>: this means that the algorithm is effective, this is to say, it provides an output / solution to the problem.

- <u>Finiteness and termination</u>: as stated before, the algorithm should run and have an end in an acceptable amount of time.

- <u>Correctness</u>: this means that the output is the expected and is acceptable considering the given input.
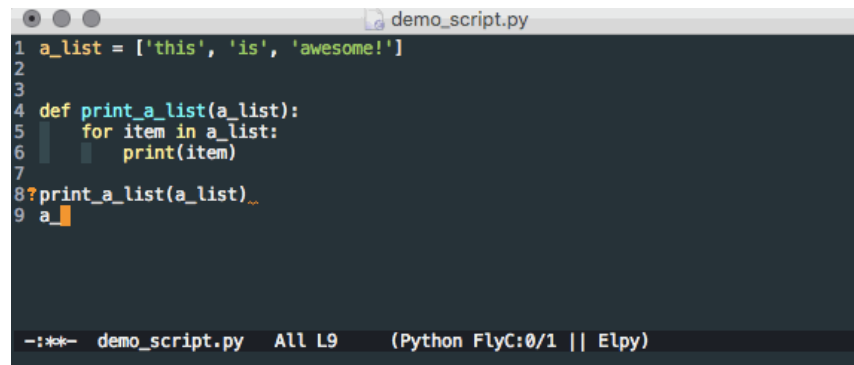
- We can also think of an algorithm as a function, in the sense that the algorithm receives some data (**input**), **processes** the data and returns a suitable **output**.

- An algorithm is a <u>blueprint</u>; a computer program is the <u>construction</u>.

- Algorithms are theoretical sequences of steps to solve problems, while programs are these steps <u>translated</u> into a language a computer can understand and execute.

- Converting an algorithm into a program involves choosing the right programming language and environment, then writing code that follows the logic of the algorithm precisely.

- Let's code together! Follow along as we go from a human-readable algorithm to a computer-executable program.

Challenge#1: Define an algorithm and its program to add two numbers.

Challenge#2: Define an algorithm and its program to identify the biggest number in a list.

Challenge#3: Define an algorithm and its program to identify the mean of a list of numbers.