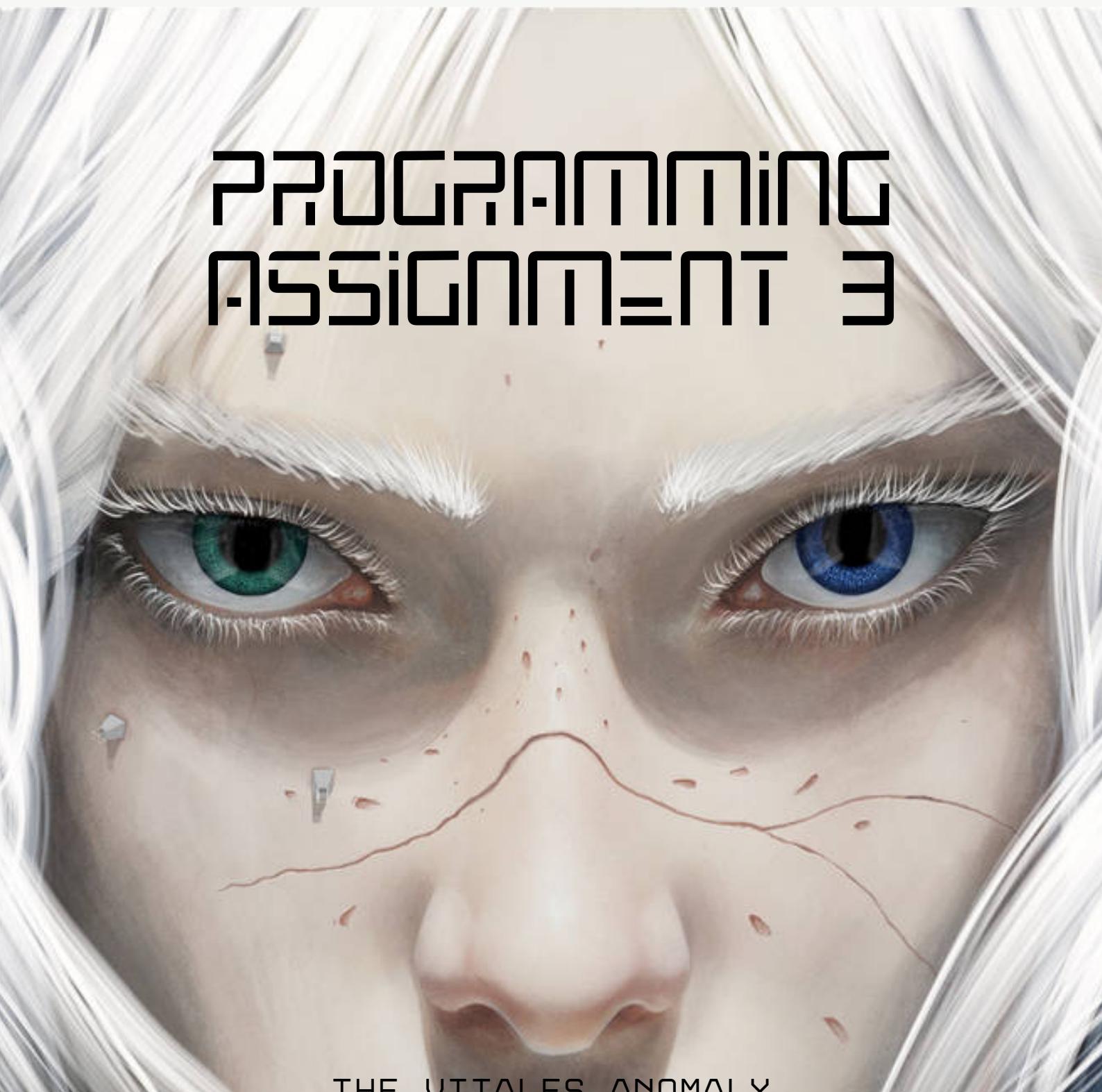




HACETTEPE UNIVERSITY  
COMPUTER ENGINEERING DEPARTMENT

BBM204 SOFTWARE PRACTICUM II  
SPRING 2024

# PROGRAMMING ASSIGNMENT 3

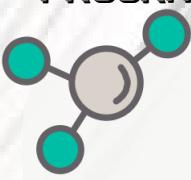


A close-up photograph of a person's eyes and nose. The person has long, light-colored hair. One eye is green and the other is blue. There are several small, dark red spots on the bridge of their nose and on their cheek, resembling blood or paint. A single tear is falling from the green eye. The background is blurred, showing more of the person's face and hair.

THE VITALES ANOMALY

# DIFFERENT EYES

assignment inspired by



**Topics:** Graphs - Connected Components, Minimum Spanning Trees

**Course Instructors:** Assoc. Prof. Dr. Erkut Erdem, Assoc. Prof. Dr. Hacer Yalm Kele, Assoc. Prof. Dr. Adnan Özsoy

**TAs:** Alperen Çakn, Dr. Selma Dilek

**Programming Language:** OpenJDK 11 - **You MUST USE this starter code**

**Due Date:** **Friday, 03.05.2024 (23:59:59)**

## Different Eyes: The Vitales Anomaly



A small, isolated community known as the **Vitales**, for their vitality and health, has captured the world's attention. This group of individuals possesses distinct physical characteristics: an ethereal fair complexion, striking white hair, and most notably, **heterochromia** - one blue and one green eye - giving them an almost otherworldly appearance. But it's not just their looks that have drawn interest. The Vitales have an unusually high resistance to diseases that have plagued humanity for centuries.

The discovery of the Vitales came at a critical time. Humanity was grappling with a series of global health crises, and traditional medical solutions were falling short. The Vitales, seemingly untouched by these ailments, represented a beacon of hope. Early analyses hinted at profound differences in their DNA variations and blood composition, which could hold the key to new treatments or even cures.

Intrigued by the potential within the Vitales genetic makeup and blood properties, a renowned biomedical research institute launched "**Project Vitae**," inspired by the Latin word for life, reflecting the project's aim to explore the life-enhancing genetic factors and unique molecular structures of the Vitales. The goal was ambitious: to delve into the Vitales' genome and blood tests to identify the genetic factors and molecular structures behind their disease immunity. Scientists hoped that understanding these elements could lead to breakthroughs in biomedical engineering, offering new strategies for disease prevention and treatment across the global population.

*Your mission is to help the scientists by designing algorithms that will be used in "**Project Vitae**".*





## 1 Part I - Mission Genesis

The **Vitales Anomaly** refers to the unique characteristics of the Vitales people, marked by their distinctive physical traits and exceptional disease resistance. Extensive data from blood analyses and tests on Vitales individuals have been compiled. In this mission, your task is to aid scientists in developing algorithms to discern molecular structures from this data. The project aims to analyze data from both Vitales and typical humans to identify unique molecular markers associated with the Vitales' enhanced immunity.

### 1.1 Background Information and Objectives

The collected data includes molecular profiles for both typical humans and Vitales individuals, cataloged separately. For each molecule identified in the test samples, the dataset provides its ID, the strength of its bonds with other molecules, and a list of molecules to which it is bonded. Molecules are assigned unique IDs for identification. **Despite the incomplete data on each molecule's bonds, it is established that any two molecules that are bonded directly or via other molecule(s) belong to the same molecular structure.**



#### Mission Objective:

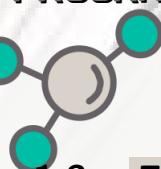
Analyze the test data of molecules identified in blood samples from both typical humans and Vitales individuals. Your task is to develop an algorithm that accurately determines the total number of unique molecular structures and identifies which molecules comprise each structure.

### 1.2 Input File Format

The input file with data for all discovered molecules and bonds will be given in the XML format as the *first command-line argument*. Your program should parse all the molecules into an instance of MolecularStructure class for further processing. The input file format is given below:

```
<MolecularData>
  <HumanMolecularData>
    <Molecule>
      <ID>M145</ID>
      <BondStrength>21</BondStrength>
      <Bonds>
        <MoleculeID>M46</MoleculeID>
        <MoleculeID>M88</MoleculeID>
      </Bonds>
    </Molecule>
    ...
  </HumanMolecularData>
  <VitalesMolecularData>
    <Molecule>
      <ID>M15</ID>
      <BondStrength>110</BondStrength>
      <Bonds>
        <MoleculeID>M14</MoleculeID>
      </Bonds>
    </Molecule>
    ...
  </VitalesMolecularData>
</MolecularData>
```

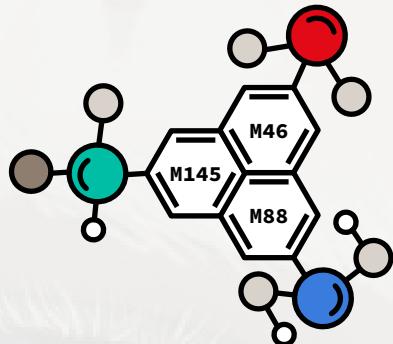




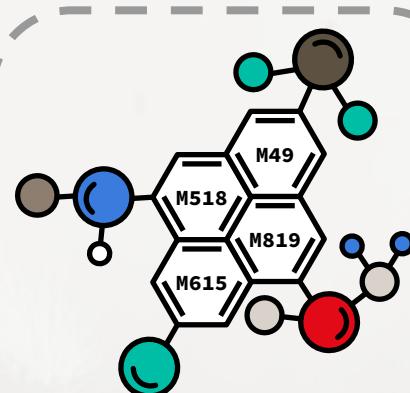
### 1.3 Expected Solution and Output Format

For the given sample input file `molecularData.xml`, the expected solution is illustrated below. We observe that the expected solution reveals the identification of three molecular structures within typical human samples and three within Vitales samples. Notably, two of these structures are unique to Vitales individuals, with no counterparts found in typical humans.

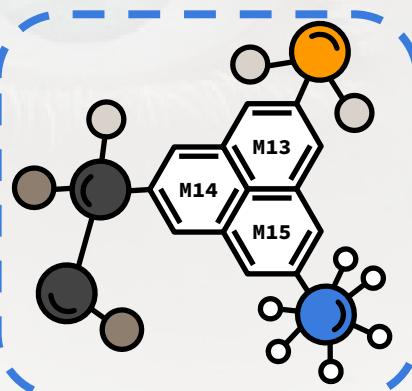
Typical Human Molecular Structures



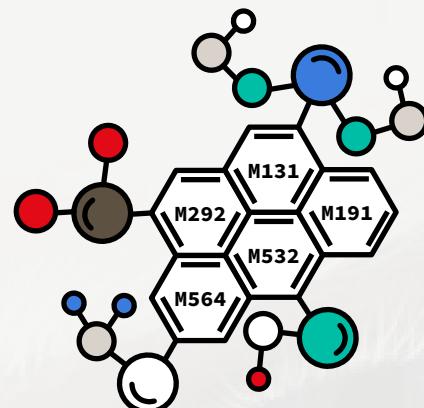
Matching Molecular Structure



Vitales Molecular Structures



Molecular Structure Unique to Vitales



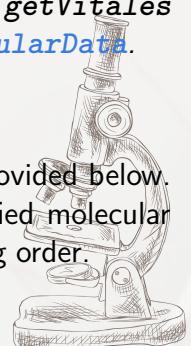
Molecular Structure Unique to Vitales



You are expected to complete:

- `readXML()` function from the class `MissionGenesis`,
- `identifyMolecularStructures()`, `printMolecularStructures()`, `getVitalesAnomaly()` and `printVitalesAnomaly()` functions from the class `MolecularData`.

The expected STDOUT output format for the given sample input file `molecularData.xml` is provided below. For full credit, ensure your output matches this format exactly. While the numbering of identified molecular structures may vary, the lists of molecules within each must be sorted by molecule IDs in ascending order.





### MISSION GENESIS START ###

3 molecular structures have been discovered in regular humans.

Molecules in Molecular Structure 1: [M46, M88, M145]

Molecules in Molecular Structure 2: [M49, M518, M615, M819]

Molecules in Molecular Structure 3: [M131, M191, M292, M532, M564]

3 molecular structures have been discovered in Vitales individuals.

Molecules in Molecular Structure 1: [M13, M14, M15]

Molecules in Molecular Structure 2: [M49, M518, M615, M819]

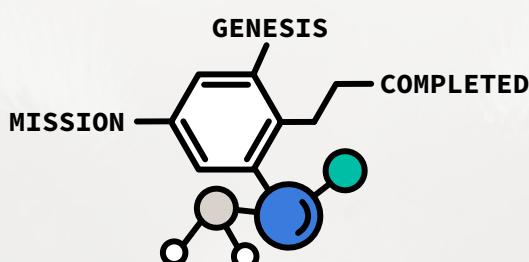
Molecules in Molecular Structure 3: [M4, M12, M532, M564]

Molecular structures unique to Vitales individuals:

[M13, M14, M15]

[M4, M12, M532, M564]

### MISSION GENESIS END ###



## 2 Part II - Mission Synthesis

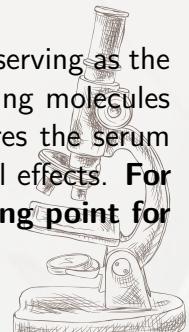
In the second mission, your task is to develop an algorithm for synthesizing an immunizer serum, leveraging unique Vitales molecular compounds to enhance the immune response in typical humans.

### 2.1 Background Information and Objectives

Initial attempts to inject typical humans with Vitales' unique molecular compounds resulted in severe allergic reactions. However, it was found that synthesizing a serum using typical human molecular structures as a base, with the addition of Vitales' immunizing compounds, significantly improved the human body's acceptance.

A critical challenge emerged: if the bond between typical human and Vitales molecular structures is overly strong, the serum's efficacy is compromised. This is because the Vitales compounds fail to detach from the human compounds, rendering them ineffective.

To address this, the serum must be synthesized by linking individual human molecular structures (serving as the base) with Vitales structures known for their immunizing properties. The focus will be on utilizing molecules with the lowest bond strength within each structure as the bonding points. This strategy ensures the serum can easily break down once administered, allowing the Vitales compounds to exert their beneficial effects. **For this reason, the molecule with the lowest bond strength should be chosen as the bonding point for each molecular structure.**

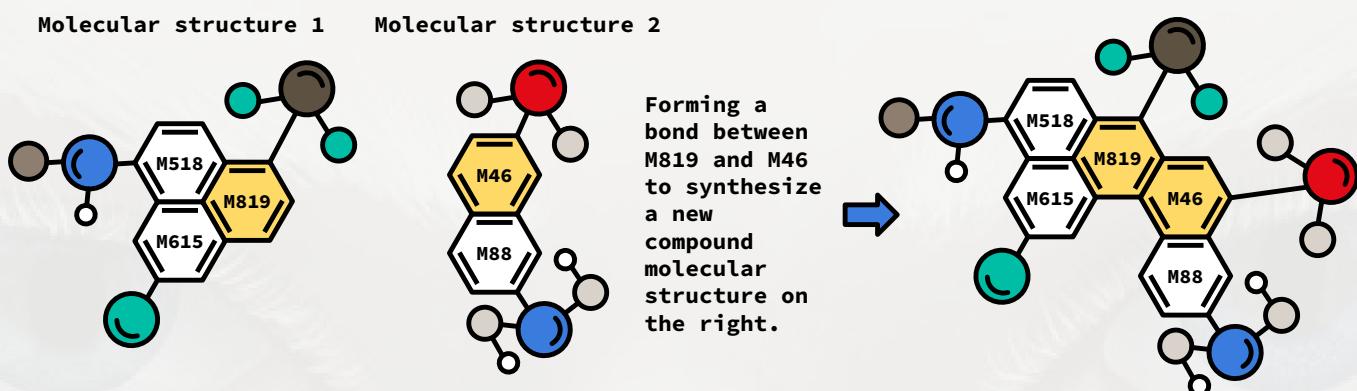




When two molecules,  $i$  and  $j$ , form a bond, the resulting bond strength is calculated as the average of their individual bond strengths:

$$BondStrength_{i,j} = \frac{BondStrength_i + BondStrength_j}{2}$$

Your mission is to develop an algorithm that will synthesize a new serum by identifying which bonds should be formed among the identified typical human molecular structures and the unique Vitales molecular structures. In other words, your algorithm will determine which bonds will be chosen among the molecules with the lowest bond strength in each molecular structure to comprise the newly formed molecular compound, i.e., serum, such that the overall bond strength will be minimum. **The serum will be fully synthesized when each molecular structure is linked to this new compound structure via a bond to some other molecular structure in this serum.** Note that by forming a bond between two molecules in two separate molecular structure, we obtain a new compound molecular structure, as illustrated in the figure below.



#### Mission Objective:

Use the information about the identified typical human molecular structures and the unique Vitales molecular structures, obtained in the previous mission, to synthesize the lowest-bond-strength compound serum from molecular structures at hand.

## 2.2 Expected Solution and Output Format

Following the sample I/O from the previous mission, remember that three typical human molecular structures and two unique Vitales molecular structures have been identified. Now, they will be used for serum synthesis. Among each of the molecular structure at hand, a molecule with the lowest bond strength will be selected:

From the structure [M46, M88, M145] in typical humans, M145 will be selected.

From the structure [M49, M518, M615, M819] in typical humans, M49 will be selected.

From the structure [M131, M191, M292, M532, M564] in typical humans, M292 will be selected.

From the structure [M13, M14, M15] in Vitales, M13 will be selected.

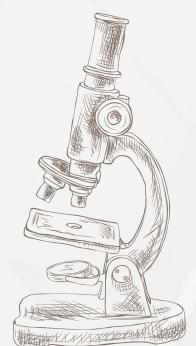
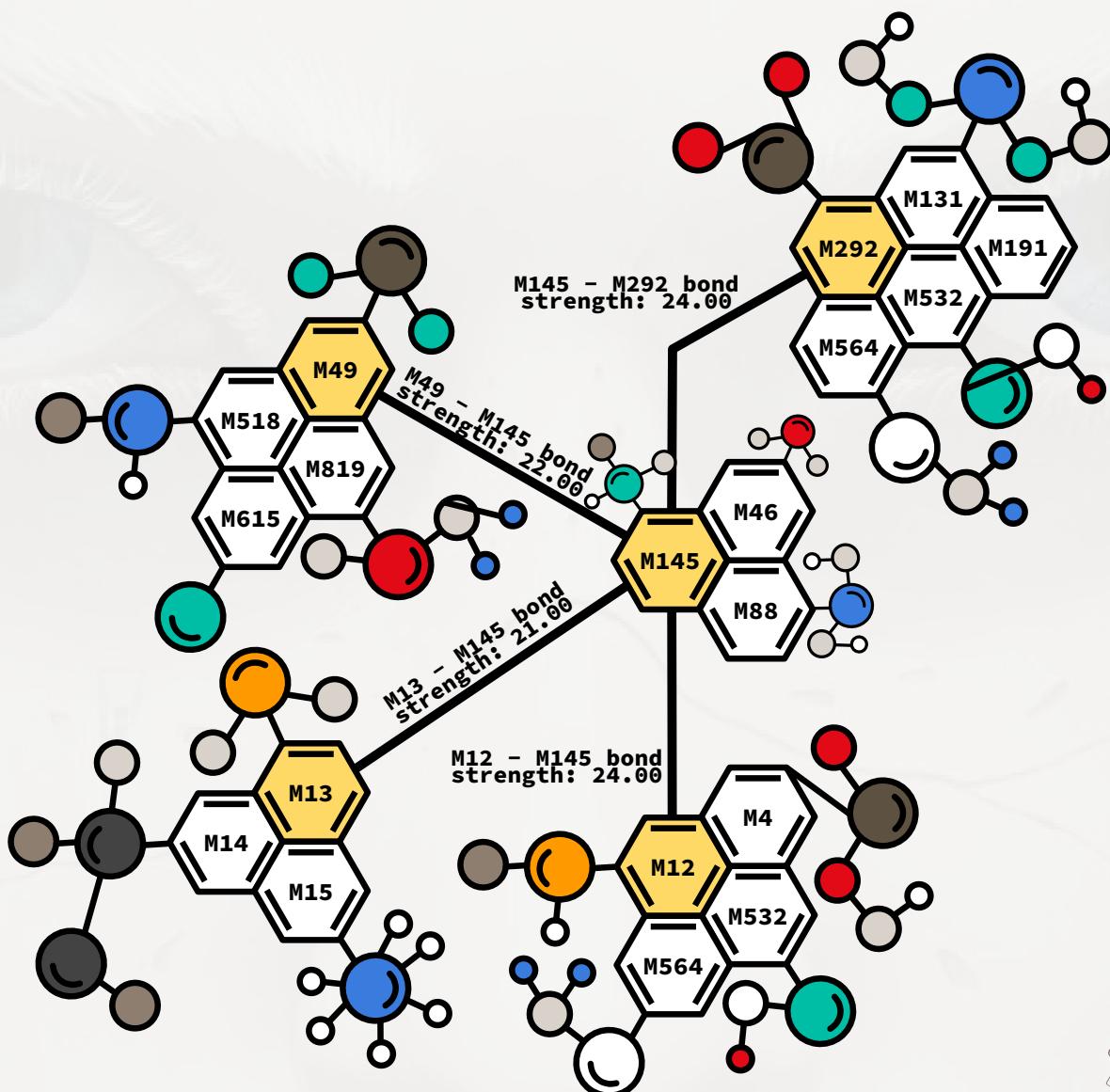
From the structure [M4, M12, M532, M564] in Vitales, M12 will be selected.

Using the formula for calculating the bond strength between two molecules, we obtain the strength of each potential bond that can be considered for synthesizing the serum:



Potential bond between M145 – M49 with strength 22.0  
Potential bond between M145 – M292 with strength 24.0  
Potential bond between M145 – M13 with strength 21.0  
Potential bond between M145 – M12 with strength 24.0  
Potential bond between M49 – M292 with strength 25.0  
Potential bond between M49 – M13 with strength 22.0  
Potential bond between M49 – M12 with strength 25.0  
Potential bond between M292 – M13 with strength 24.0  
Potential bond between M292 – M12 with strength 27.0  
Potential bond between M13 – M12 with strength 24.0

It is sufficient for each molecular structure to be bonded to at least one other structure within the serum. Therefore, we should avoid forming additional bonds that would increase the serum's overall bond strength unnecessarily. For instance, in the given sample scenario, a bond between M292 and M12 with a strength of 27 is unnecessary since both molecules can bond to M145 at lower bond strengths, thus preventing an increase in the serum's total bond strength. The expected solution is illustrated below.





You are expected to complete:

- `synthesizeSerum()` and `printSynthesis()` functions from the class `MissionSynthesis`.

The expected STDOUT output format for the given sample input is given below. Note that the molecules at each end of a bond must be printed sorted according to molecule IDs in ascending order, but the order in which you print bonds is not important.

```
### MISSION SYNTHESIS START ###
```

```
Typical human molecules selected for synthesis: [M145, M49, M292]
```

```
Vitales molecules selected for synthesis: [M13, M12]
```

```
Synthesizing the serum...
```

```
Forming a bond between M13 - M145 with strength 21.00
```

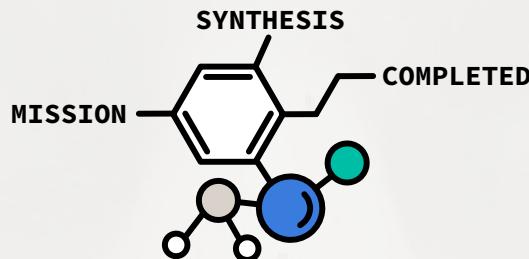
```
Forming a bond between M49 - M145 with strength 22.00
```

```
Forming a bond between M145 - M292 with strength 24.00
```

```
Forming a bond between M12 - M145 with strength 24.00
```

```
The total serum bond strength is 91.00
```

```
### MISSION SYNTHESIS END ###
```



## Must-Use Starter Codes

You MUST use [this starter code](#). All classes should be placed directly inside your **zip** archive. Feel free to create other additional classes if necessary, but they should also be directly inside **zip**.

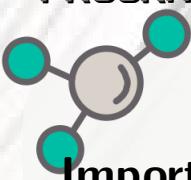
## Grading Policy

- Implementation of the algorithms: 90%
  - Mission Genesis: 45%
  - Mission Synthesis: 45%
- Output test: 10%



Note that you need to get a NON-ZERO grade from the assignment to get the submission accepted. Submissions graded with 0 will be counted as NO SUBMISSION!





## Important Notes

- Do not miss the deadline: **Friday, 03.05.2024 (23:59:59)**.
- Save all your work until the assignment is graded.
- The assignment solution you submit must be your original, individual work. Duplicate or similar assignments are both going to be considered as cheating.
- You can ask your questions via Piazza (<https://piazza.com/hacettepe.edu.tr/spring2024/bbm204>), and you are supposed to be aware of everything discussed on Piazza.
- You must test your code via **Tur<sup>6</sup>Bo Grader (does not count as submission!)**.
- You must submit your work via <https://submit.cs.hacettepe.edu.tr/> with the file hierarchy given below:
  - **b<studentID>.zip**
    - \* Bond.java <FILE>
    - \* Main.java <FILE>
    - \* MissionGenesis.java <FILE>
    - \* MissionSynthesis.java <FILE>
    - \* MolecularData.java <FILE>
    - \* MolecularStructure.java <FILE>
    - \* Molecule.java <FILE>
- The name of the main class that contains the main method should be **Main.java**. **You MUST use this starter code**. The main class and all other classes should be placed directly in your **zip** archive. Feel free to create other additional classes if necessary, but they should also be inside the **zip**.
- This file hierarchy must be zipped before submitted (not .rar, only .zip files are supported).
- **Usage of any external libraries is forbidden.**

## Run Configuration

Your code will be compiled and run as follows:

```
javac *.java  
java Main <molecularDataXMLFile>
```

## Academic Integrity Policy

All work on assignments **must be done individually**. You are encouraged to discuss the given assignments with your classmates, but these discussions should be carried out in an abstract way. That is, discussions related to a particular solution to a specific problem (either in actual code or in pseudocode) **will not be tolerated**. In short, turning in someone else's work (including work available on the internet), in whole or in part, as your own will be considered as a **violation of academic integrity**. Please note that the former condition also holds for the material found on the web as everything on the web has been written by someone else.



The submissions will be subjected to a similarity check. Any submissions that fail the similarity check will not be graded and will be reported to the ethics committee as a case of academic integrity violation, which may result in suspension of the involved students.

