EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

BIO-4372
CHEMINFORMATICS
Winter Semester 2022-23

Dr. Philipp Thiel
Institute for Bioinformatics
and Medical Informatics

Assignment Sheet 6

Submission deadline: December 5, 2022, 9 a.m.

A6.1: Implementing MCS Detection [12 points]

In this task you shall implement detection of pairwise MCS(s). To do so, please implement functionality to generate the compatibility graph from a pair of molecules and use this graph as the input to an algorithm for detection of the maximum clique(s). Please use the following molecules to develop and test your implementation:

- m11: `CC(=O)OC1=C(C=CC=C1)C(O)=O`
- m12: `NC(=O)OS`

For clique detection you are allowed to use a Python graph library and you are not required to implement an algorithmic solution yourself. You may consider the Python library NetworkX to solve this subtask.

A6.2: Analyzing Clique-based MCS [8 points]

Please apply the implemented MCS detection onto the following molecules:

- m21: `CC(=O)OSC1C(CCCC1)C(O)=O`
- m22: `C1CCCCC1NC(=O)OS`

Based on the detected MCSs, please deduce important properties of the underlying MCS implementation. As a guideline which properties are meant, please have a look at slide 19 of the corresponding lecture. Please identify interesting MCS cases that support your findings and briefly discuss them in your **report**.

If you have not solved A6.1: The maximum cliques of the compatibility graph between m21 and m22 can be found in the supporting file `chin-a6-mcs-list.csv`. It thus provides you all the information you need to solve this subtask without haven solved A6.1.

Hints to A6.1 and A6.2:

A graphical representation of the molecules labeled with atom indices and highlighting of the MCS substructures can be very helpful. The linked resources explain how you can use Jupyter Notebooks together with RDKit to draw molecules and to highlight substructures. Thus, it is probably a good idea to implement both tasks in a Jupyter Notebook and you are allowed to hand in a Jupyter Notebook instead of a plain Python file.

- `https://www.rdkit.org/docs/Cookbook.html#drawing-molecules-jupyter`
- `https://www.rdkit.org/docs/Cookbook.html#highlight-a-substructure-in-a-molecule`

Also very helpful is the RDKit function `MolsToGridImage()`, which is part of the `Draw` module. Just do a search on the Cookbook-Page (the links above) and you will find what you need!

▷ Please use Slack to discuss problems in the first place
▷ If you have confidential questions, don't hesitate to drop by or write an e-mail