Programming Expertise
University of Potsdam SS 2025
Detlef Groth, Christian Kappel
1st Test-Exam July 3rd, 2025

You have 90 minutes time for the implementation of the programming tasks. Make after the steps 1-3 intermediate versions: GooboParser1.cpp, GooboParser2.cpp and GooboParser3.cpp . Submit at the end of the exam those single versions to the folder *Test-Exam-01*. 75% of the tasks will be on execute correctness and 25% of the tasks will be weighted by usefulness and clearness of the implementation.
The theory questions in 4-6 are to be answered first without any aids and the sheet with the answers is to be handed in after about 15 minutes. For the computer tasks 1-3 all aids are allowed during the exam. This does not include personal support from fellow students or other persons.

Clarification

With this I state, that I will not take and give any not allowed support during the exam.

Name, Matrikel-Number.: Signature:

1. 2 points (layout console application)
2. 4 points (implementation of console application
3. 4 points (advanced terminal task)
4. 2 points (theory – C)
5. 2 points (theory – C)
6. 2 points (theory - C++)

Sum: 16 points

Good luck !!

1. Layout of console application and utilizing command line arguments (2 points):

Create the basic outline of a console application with main function, help function and checking of command line arguments. Save the possible command line arguments in variables or use a command line processor like popl or argparse. If the right number of arguments was not given call the help function and exit the application. The two arguments and the optional third should be a GO-obo input filename and a taskname like *--get--metacyc* or *--tab-metacyc* and the optional third a *MetaCyc* identifier for the *--get-metacyc* parameter.

You find the Obofiles on Moodle. Check if the given filename points to a valid filename, check if a valid argument was entered and check, if --get-metacyc was provided, that as well an MetaCyc id argument was given and that this id contains either a RXN or a PWY substring.

So the usage line shoud be something like this:

```
$ appname [--help,--get-metacyc MetaCyc-ID,--tab-metacyc] OBOFILE
```

C++ filename first task: _____

2. Opening and searching in the GO-obo file (4 points): Our program should work with any GO-obo file. Please don't hardcode the filename in your application, if you do, you get a minus point. Implement the *--get-metacyc* function for a specific MetaCyc-ID. If the user gives, for example, the three arguments: *--get-metacyc BETA-LACTAMASE-RXN filename.obo* on the command line, then the GO-obo file is parsed and the ID, the name, the namespace and and the entry itself should displayed on the terminal. Your function has to return the data to the function caller, for instance as a vector of tuples, print the result outside of the function, for instance in main.
Hint: If your code is very slow use limit your search to the first 1000 entries first to save your programming time during development. You should un-comment this in your final program. Here is a possible example invocation with output result:

```
$ appname --get-metacyc  BETA-LACTAMASE-RXN gene_ontology-2022-01.obo
GO:0008800      beta-lactamase activity molecular_function      MetaCyc:BETA-LACTAMASE-RXN
GO:0033250      penicillinase activity  molecular_function      MetaCyc:BETA-LACTAMASE-RXN
GO:0033251      cephalosporinase activity       molecular_function      MetaCyc:BETA-LACTAMASE-RXN
```

C++ filename(s) second task: _____

3. Advanced Terminal Application (4 points):

1 (1 point): Extend your application as a C++ class and so that it can handle more than 1 MetaCyc-ID at the same time given on the terminal.

Hint: It might be easier to handle multiple ids in the main function only, not within the class function/method!

2) (3 points): Create a mapping using the command argument *--tab-metacyc* between all RXN-Metacyc entries, GO-ids and EC-ids, only if all three entries are within the same go-id display them, here an example invocation:

```
$ appname --tab-metacyc go-2020.obo
GO:0000016      EC:3.2.1.108    LACTASE-RXN
GO:0000034      EC:3.5.4.2      ADENINE-DEAMINASE-RXN
GO:0000048      EC:2.3.2.12     PEPTIDYLTRANSFERASE-RXN
...
```

Hint: You can assume, that EC comes first in the OBO file before the MetaCyc-ID if there is any. In case there is no EC-mapping write NA in the EC column.

Deadline for submission is as usually Wednesday morning.