Exam Algorithmic Cheminformatics (2025)

Time and Place

Registration necessary (mail to: nwi-office@uni-bielefeld.de, deadline for registration: Thursday July 17th, 16:00)

Time: Tuesday, July 22, 2024 or Thursday July 24th, between 9am and 1pm.
Place: The exam takes place physically U10-116.

The expected total examination time per student is about 30 minutes (see below).

Procedure

The exam is in German.

The procedure is as follows:

You will first randomly chose a question and give a presentation (technically you will decide to turn around a specific piece of paper with the question number on it). This will be followed by asking question on the topic you chose. After that, the second part of the exam starts, where we will ask questions on all other topics listed below.

When it is your turn for examination, you will draw a question. Note that you have no preparation time. The list of questions can be found below. Then the actual exam takes place.

The whole exam (without the time for agreeing on a grade) lasts approximately 25-30 minutes. You should start by presenting material related to the question you drew. Aim for a reasonable high pace and focus on the most interesting material related to the question. You are not supposed to use note material, textbooks, transparencies, etc. Aim for a reasonable high pace and focus on the most interesting material related to the question.

You are allowed to use keywords for each question, such that you can remember what you want to present during your presentation. As a guideline you are expected to not have more than 10 keywords per question on the list, that you are allowed to use during the oral exam. Note that this list is expected to be put on your table during your oral exam. We will not accept that you use any material during the (second) part of the exam, where we ask questions.

We will supplement with specific questions when appropriate, and after a while, we will end the discussion of the exam question that you drew and turn to material from other parts of the curriculum. Note that all of this as well as discussion about the grade is included in the approx minutes, so **do not count on more than 10-14 minutes for your own presentation**.

Some of the questions below are quite broad, so you must select the material you choose to cover. You will of course also be evaluated based on your selection of material. If you only present the simplest material, you limit the grade you can obtain. On the other hand, a good presentation of the simple material is better than a very poor presentation of the harder material. We might of course still ask you questions about material that you have decided to skip.

Curriculum

The curriculum consists of all documents listed at `https://ac2025.algochem.techfak.de/Material/` (after the last lecture on July 18th). The list of articles, book chapters, and slide sets is a superset of what you should know. Note that "mandatory" in the title of the slide set might not necessarily mean, that you need to know the complete content of the articles/slides, nothing that has not been discussed during the lecture you do not have to know. Furthermore, there is also mandatory material not explicitly marked as mandatory.

In addition the curriculum contains the text of the required assignments and the Weekly Notes/Exercises.

Exam Questions

In the following the list questions that you draw from is given (in bold face), (the list of subquestions is incomplete and just a suggestion.)

- **Canonical Representations and Graph Isomorphism**
  - Morgan's algorithm
  - SMILES notation
  - Canonical labeling of graphs
  - Subgraph and graph isomorphism (Ullmanns algorithm)
  - Graph isomorphism (automorphism-pruning-based approaches like traces, nauty)
  - Relation to (generative) chemistry

- **Graph Grammars / Generative Chemistries**
  - Graph Grammars
  - Double pushout approach
  - ILP approach in hypergraphs
  - Autocatalysis
- **Ring Perception**
  - Hanser Algorithm
  - Cycle Bases in chemistry
  - Kirchhoff-fundamental cycle basis
  - Minimal cycle basis
  - Horton's algorithm
  - de Pina's algorithm
- **Petri Nets**
  - Properties (liveness/boundedness/reversibility)
  - Reachability graph, coverability graph
  - Invariants and how to determine them (Farkas algorithm)
  - NPc proof of reachable marking
  - Petri nets to model chemical/biological networks

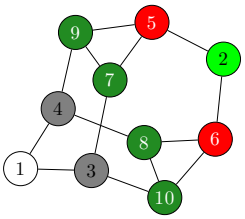
- **Molecular Descriptors and QSAR**
  - Molecular descriptors
  - QSAR
  - Principal component analysis (PCA) / PCR

 Algorithmic Cheminformatics, Exam Questions

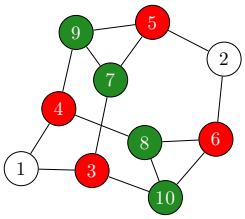
The following is an (incomplete) list of topics that have also been discussed in the course, but they won't appear as an individual presentation question in the oral exam. You of course might still be asked questions on those topics in the second part of the exam.

- **ADDON (examples)**
  - Law of Mass Action
  - Gillespie simulation
  - k shortest path finding
  - DPO rule composition
  - anything else that was discussed

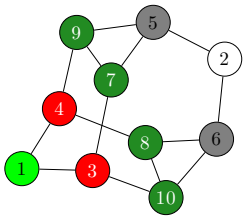
The following two figures is material you can use. You can (but don't have to) use them during the second part (not in the first part) in order to explain.



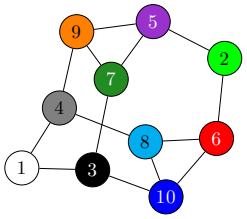
$\pi_{(1)} = [1 \mid 2 \mid 7\ 8\ 9\ 10 \mid 5\ 6 \mid 3\ 4]$



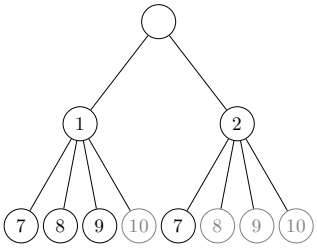
$\pi_{()} = [1\ 2 \mid 7\ 8\ 9\ 10 \mid 3\ 4\ 5\ 6]$

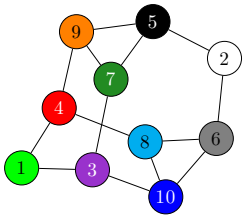


$\pi_{(2)} = [2 \mid 1 \mid 7\ 8\ 9\ 10 \mid 3\ 4 \mid 5\ 6]$

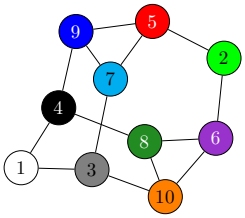


$\pi_{(1,7)} = [1 \mid 2 \mid 7 \mid 10 \mid 8 \mid 9 \mid 6 \mid 5 \mid 4 \mid 3]$



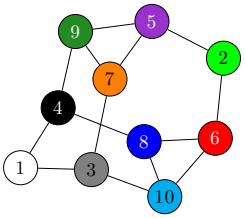


$\pi_{(2,7)} = [2 \mid 1 \mid 7 \mid 10 \mid 8 \mid 9 \mid 4 \mid 3 \mid 6 \mid 5]$

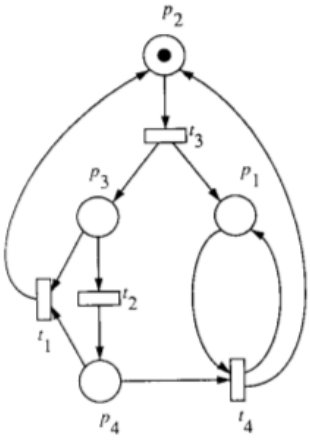


$\pi_{(1,8)} = [1 \mid 2 \mid 8 \mid 9 \mid 7 \mid 10 \mid 5 \mid 6 \mid 3 \mid 4]$

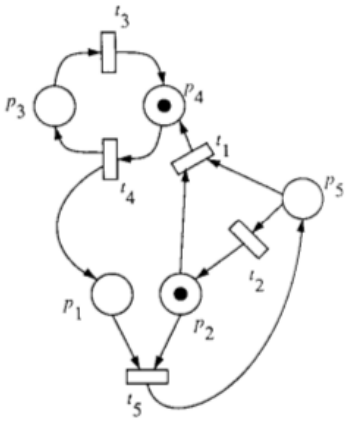
Colour order



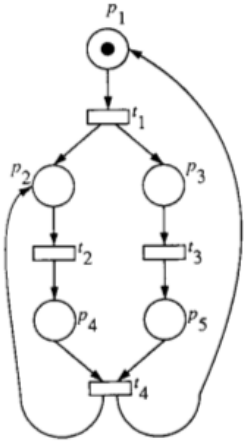
$\pi_{(1,9)} = [1 \mid 2 \mid 9 \mid 8 \mid 10 \mid 7 \mid 6 \mid 5 \mid 3 \mid 4]$


p_2
t_3
p_3
p_1
t_2
t_1
p_4
t_4

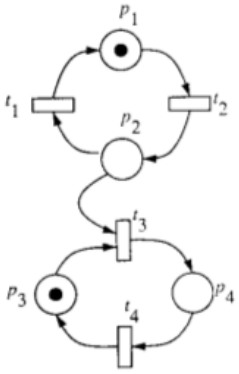
(a)


t_3
p_3
p_4
t_1
t_4
p_5
t_2
p_1
p_2
t_5

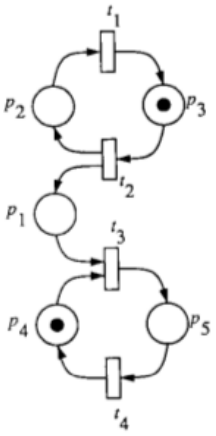
(b)


p_1
t_1
p_2
p_3
t_2
t_3
p_4
p_5
t_4

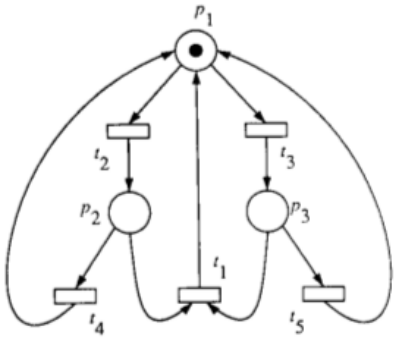
(c)


p_1
t_1
t_2
p_2
t_3
p_3
p_4
t_4

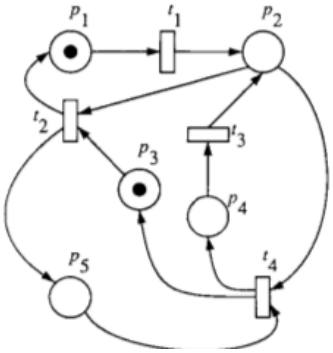
(d)

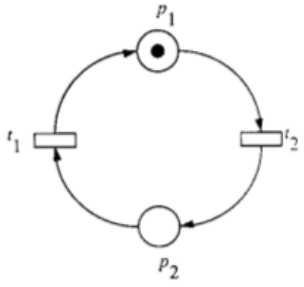

t_1
p_2
p_3
t_2
p_1
t_3
p_4
p_5
t_4

(e)


p_1
t_2
t_3
p_2
p_3
t_1
t_4
t_5

(f)


p_1
t_1
p_2
t_2
t_3
p_3
p_4
p_5
t_4


p_1
t_1
t_2
p_2