Exam DM840 Cheminformatics (2018)

Time and Place

Time: January 28th, 2019, starting 10:00.
Place: The exam takes place in U158A

Even though the expected total examination time per student is about 27 minutes (see below), it is not possible to calculate the exact examination time from the placement on the list, since students earlier on the list may not show up. Thus, students are expected to show up plenty early. In principle, all students who are taking the exam on a particular date are supposed to show up when the examination starts, i.e., at the time the first student is scheduled. This is partly because of the way external examiners are paid, which is by the number of students who show up for examination. For this particular exam, we do not expect many no-shows, so showing up one hour before the estimated time of the exam should be safe.

Procedure

The exam is in English. 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 censor and the examiner 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, computer, etc.

You are allowed to bring 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 bring to the oral exam. Note that this list is expected to be put on the table during your oral exam.

We, the examinator and the censor, 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 between examinator and censor about the grade is included in the approx minutes, so do not count on more than 10-12 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 marked as "mandatory" in the Blackboard System. The following list of articles, book chapters, and slide sets is a superset of what you should know. Note that "mandatory" might of course not mean, that you need to know the complete content of the articles, nothing that has not been discussed during the lecture you do not have to know.

```
Slides / notes:

week36-2018.slideset1.final.pdf
week36-2018.slideset2.final.pdf
week37-2018.GenerativeChemistries.pdf
week37.2018.Isomorphism-Ullmann.mandatory.corrected.pdf
week38-2018.canon.pdf
week38-2018-mckay-notes.mandatory.pdf
week39-40.cycle-basis.mandatory-2.not.final.pdf
week39-40.MCB.horton.dePina.mandatory.pdf
week39-40.ring-perception-2017-notes(summary of mandatory article).pdf
week39-morgan.pdf
week39-SMILES-notes.mandatory.(or weininger article).pdf
week41-MOD-framework-addon.pdf
week43-2018-petri-final.mandatory.pdf
week45-AC-OGF-detailed.pdf
week46-AC-EGF-detailed.pdf
week47-AC-MGF-voluntary.pdf
week47-boltzmann.voluntary.pdf
week47-complex.justanaddon.pdf
week47-stereo-combinatorics.pdf
week48-mol-des.partially-mandatory.pdf
week48-pca-notes-voluntary(notes follow the article).pdf
week49-PA-final.pdf
week50.polya.pdf


Book Chapters:
--------------
week36-partially-mandatory--Chemoinformatics.Gasteiger.chapter2.pdf
week36-polya-enumeration-teaser_and_graph-theory-roots.voluntrary.pdf
week36-voluntary-Combinatorics-Ancient-and-Modern_-_Chapter8_-_Robin-Wilson_Early-Graph-Theory.pdf
week37.jla.thesis.GraphGrammars.mandatory.pdf
week38.jla.thesis.canonicalisation.mandatory.pdf
week47-partially-mandatory.pdf

Original Articles:
------------------
week36-polya-enumeration-teaser_and_graph-theory-roots.voluntrary.pdf
week36-weininger.87.smiles1-mandatory(or notes).pdf
week36-weininger.87.smiles2-mandatory(or notes).pdf
week37-fund-of-graph-recommended.pdf
```

week37-ullmann-mandatory.pdf
week38-McKay-nauty-Hartke-Radcliffe(here you find details for the handwritten notes,voluntary readi
week39-40.MCB-pina.runtime.voluntary.pdf
week39-40.ring-perception-mandatory(see notes).pdf
week39-canon.recommended.pdf
week40-berger-ring-perception-applications.voluntary.pdf
week41-downs-et-al-voluntary2(and wrong).MCB.pdf
week41-downs-et-al-voluntary(and wrong).MCB.pdf
week43-petrinet-murata89-examples_for_liveness_boundedness_reversability.pdf
week43-PetriNets-Complexity-Esparza-voluntary.pdf
week45-KShortest(voluntary-slides-should-be-good-enough).pdf
week48-eigenfaces-voluntary.pdf
week48-pca-mandatory-page_1-6.pdf
week49-intro-stoch-pa.voluntary.pdf
week49-PA-cardelli-Artificial-Biochemistry-2009.partially-mandatory.pdf
week50-Zombies-voluntary.pdf

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.)

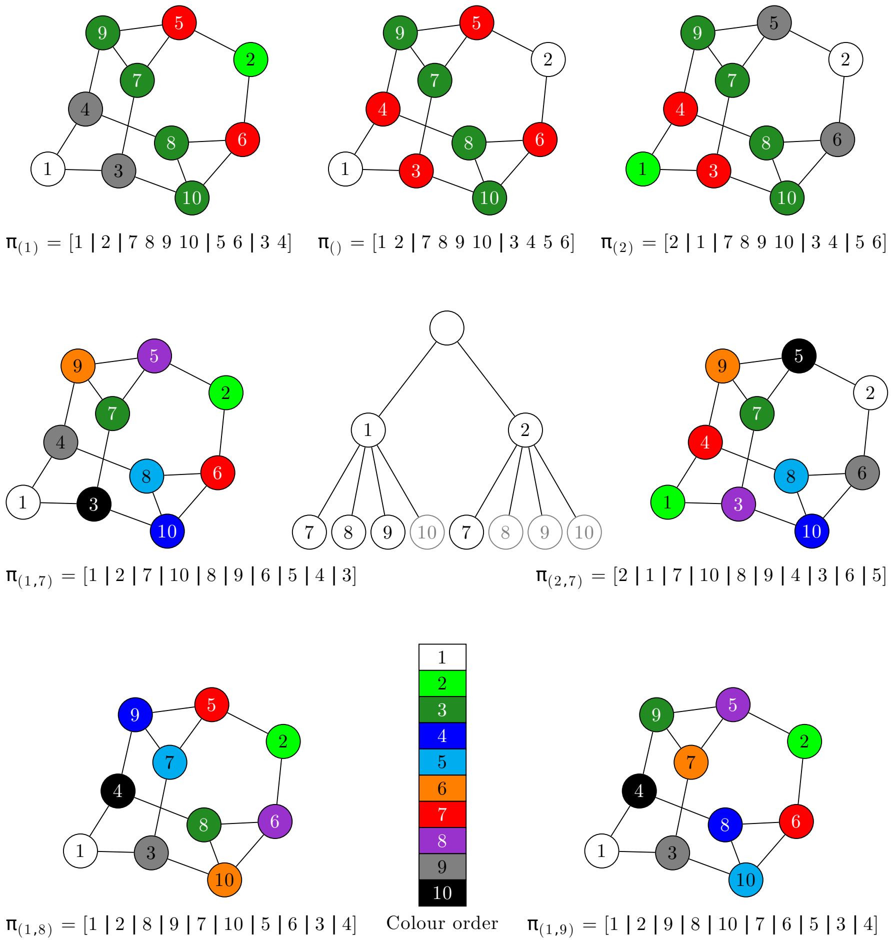
- **Canonical Representations**
  - Morgan's algorithm
  - SMILES notation
  - Canonical labeling of graphs (McKays algorithm (nauty), Traces)
- **Generative Chemistries**
  - Graph Grammars
  - Double pushout approach
  - ILP approach in hypergraphs
  - Autocatalysis
- **Ring Perception**
  - Hanser Algorithm
  - Cycle Bases in chemistry
  - Kirchhoff-fundamental cycle basis
  - Minimal cycle basis
  - Horton's algorithm
  - de Pina's algorithm
- **Graph Isomorphism**
  - Subgraph and graph isomorphism (Ullmanns algorithm)
  - Graph isomorphism (McKay's algorithm (nauty), traces)
  - Relation to (generative) chemistry
- **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
- **Combinatorics and Polya Counting**
  - Generating Functions

- Symbolic Method
- Ordinary and Exponential GF
- Multivariate GF
- Cylce Index Polynomial and Burnside's Lemma
- Connection to Chemistry (counting isomers, counting chlorinated compounds)

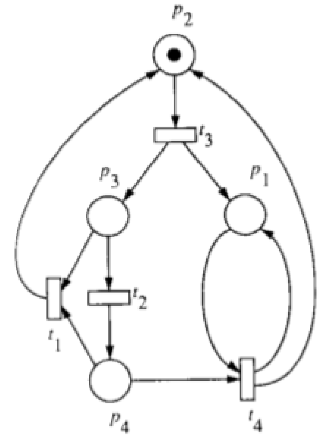
The following was discussed in the course, but it won't appear as a question in the oral exam.

- **NOT A QUESTION: Process Algebras**
  - π-Calculus
  - Modeling of Chemistry with stochastic π-calculus
  - Modeling of epidemic systems with stochastic π-calculus

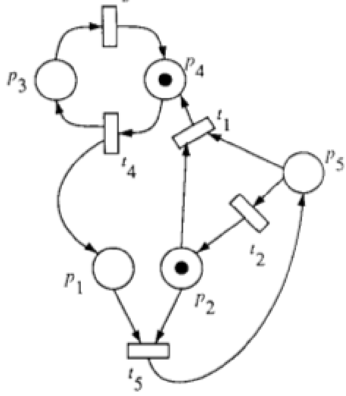
The following two figures will be provided as slides for the oral exam. You can (but don't have to) use them during the second part in order to explain.



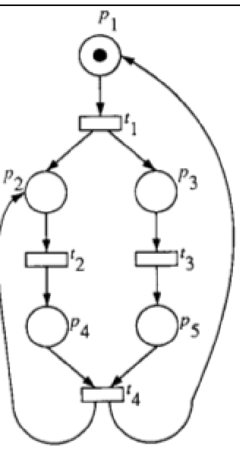
$\pi_{(1)} = [1 \mid 2 \mid 7\ 8\ 9\ 10 \mid 5\ 6 \mid 3\ 4]$
$\pi_{()} = [1\ 2 \mid 7\ 8\ 9\ 10 \mid 3\ 4\ 5\ 6]$
$\pi_{(2)} = [2 \mid 1 \mid 7\ 8\ 9\ 10 \mid 3\ 4 \mid 5\ 6]$
$\pi_{(1,7)} = [1 \mid 2 \mid 7 \mid 10 \mid 8 \mid 9 \mid 6 \mid 5 \mid 4 \mid 3]$
$\pi_{(2,7)} = [2 \mid 1 \mid 7 \mid 10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

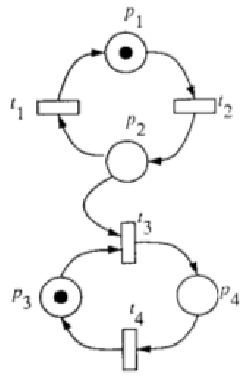
(a)


t_3
p_3
p_4
t_1
t_4
p_5
t_2
p_1
p_2
t_5

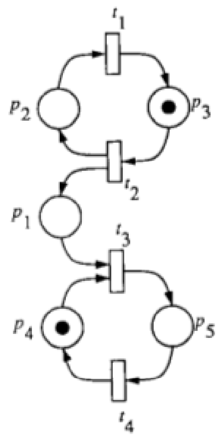
(b)


p_1
t_1
p_2
p_3
t_2
t_3
p_4
p_5
t_4

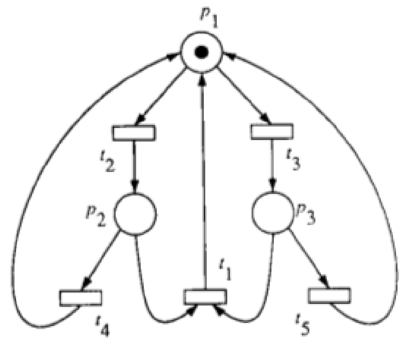
(c)


p_1
t_1
t_2
p_2
t_3
p_3
p_4
t_4

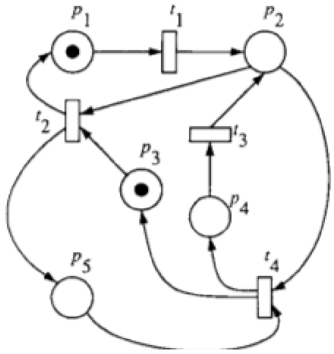
(d)

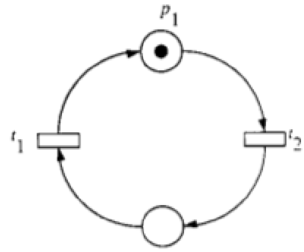

t_1
p_2
p_3
t_2
p_1
t_3
p_4
p_5
t_4

(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