Database Theory
VU 181.140, WS 2020

0. General Information

Reinhard Pichler

Institute of Logic and Computation
DBAI Group
TU Wien

2 October, 2020

Outline

Classes

- Language. This lecture will probably be held in English.
- Time. Throughout the term: Fridays, 09:15 – 11:00.
- Place.
    - All classes will be online via Zoom
      see TISS and/or TUWEL for the Zoom link
    - Also the quiz at the beginning of the semester will be via Zoom

Prerequisites and Admission

- **Prerequisites**.
  - This course is designed for master's students.
  - It is highly recommended to attend this course after the course Formale Methoden der Informatik (185.291).

- **Knowledge and skills required**.
  - basic knowledge of databases helpful (e.g., VU Datenmodellierung)
  - basic knowledge in mathematical logic
  - introduction to complexity theory
  - in particular, the central concept of "problem reduction"

- **Admission**.
  - primarily for master's students!
  - positive assessment in a quiz is required
  - each student has at most two attempts

Quiz

- **Goal**.
  - ensure that students have the required knowledge and skills
  - basic knowledge in mathematical logic and complexity theory;
  - in particular, the central concept of "problem reduction".

- **Organization**.
  - closed book (no material allowed)
  - Being able to solve all questions of the exercise sheet of block 1 (complexity theory part) of the course "Formale Methoden der Informatik" clearly suffices for the quiz.
  - max. 10 credits; passed with ≥ 5.

- **Date and Time**.
  - Tuesday, 6 October, 09:00 - 11:00
  - Tuesday, 13 October, 09:00 - 11:00
  - 60 min actual working time

Organization of the Quiz (continued)

- the quiz will be remote

- make sure that you are alone in a room

- join the Zoom meeting

- the camera and microphone have to be switched on during the test

- you will receive the questions as PDF file

- print the file

- enter the solutions on the printout

- at the end: scan your solutions (or take pictures of sufficiently high quality) and upload the scan to TUWEL (in case of problems, send the scan via email to pichler@dbai.tuwien.ac.at).

Communication

- (during, after) classes
- Course Homepage:
  `http://www.dbai.tuwien.ac.at/staff/pichler/dbt`
- TISS: please check your mail address in TISS
- TUWEL (probably)

Course Overview (Tentative Plan)

## Fundamental aspects of (relational) query languages

- Relational algebra vs. relational calculus vs. SQL
- Introduction to datalog
- Codd's Theorem: relational algebra vs. First-Order logic
- Trakhtenbrot's Theorem: some undecidability results
- Complexity of Query Evaluation
- (Acyclic) Conjunctive Queries
- Inexpressibility Results (Ehrenfeucht-Fraïssé Games, Locality)
- Beyond relational data

Assessment

Components

1. Individual work on 1 article from database theory research
   - Details (e.g. assignment of articles) to be provided later
   - Short written report of this article (2–3 pages)
   - Oral presentation (ca. 20 min, depends on number of participants)
   - Sufficient to get mark 3 (or worse)
2. Oral exam
   - positive assessment of report & presentation required
   - exam not compulsory, but required for marks 1 and 2
3. Quiz
   - will be taken into account in case of intermediate marks
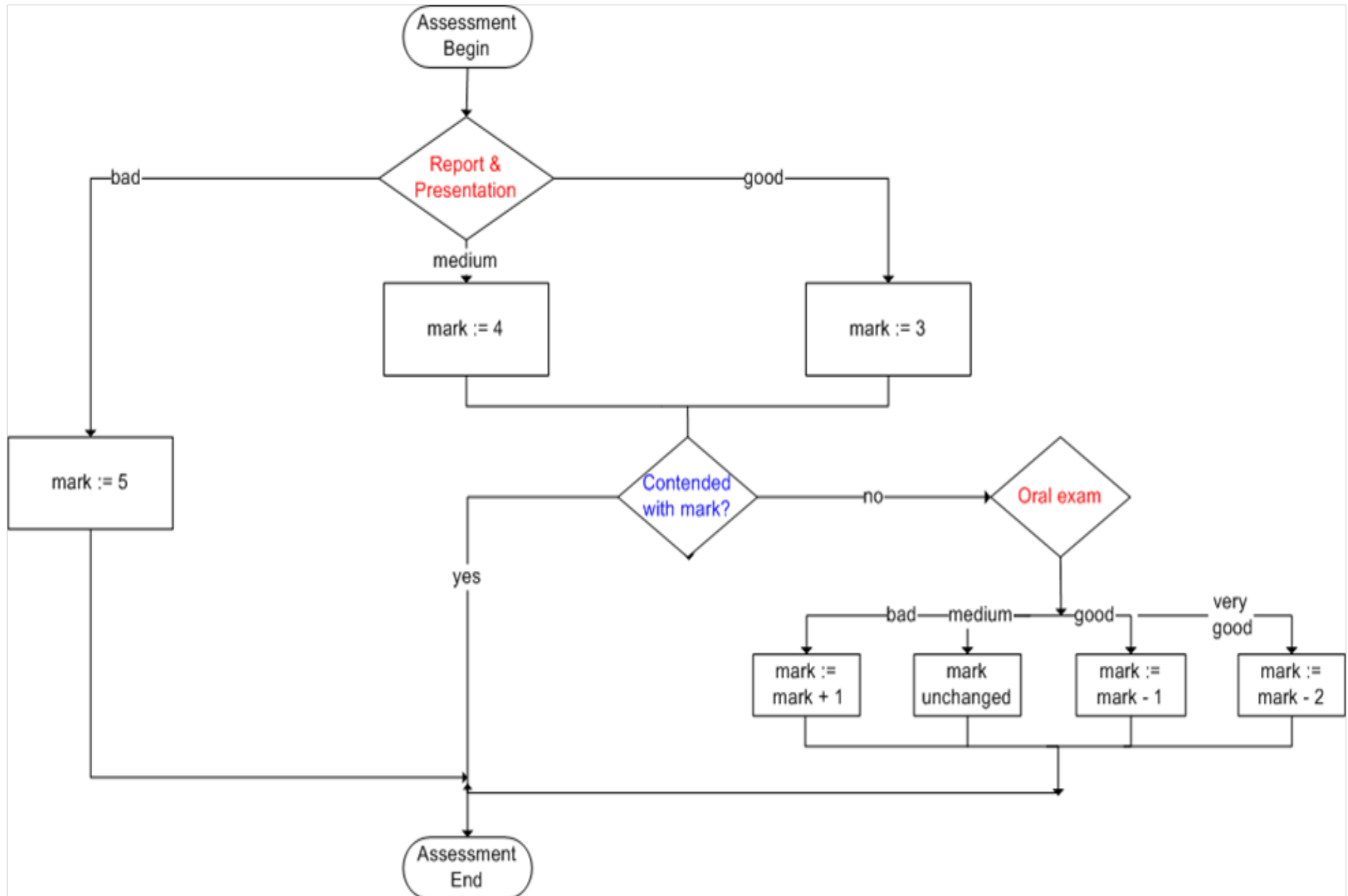
Assessment of Report & Presentation

Criteria of a good report & presentation

- reasonable effort (30 - 40 hours)
- basic understanding of the article
- honestly identify parts which you did not understand
  (give a justification: which prerequisites were missing?)
- relate the article to the contents of the course
- quickly check important background articles
- (presentation) being able to answer questions
  (in particular, those relating the article to the course)

ECTS Breakdown

3 ECTS corresponds to 75h of work for "standard students" fulfilling the prerequisites (i.e., VU Formale Methoden der Informatik - 185.291).

```
10 classes (including preparation):        25h
research article (report, presentation):   35h
preparation for oral exam:                 15h
===========================================
in total:                                  75h
```

DB Literature

Most Important DB-Conferences

- ACM SIGMOD:
  - "International Conference on Management of Data"
  - 2020 in Portland: `https://sigmod2020.org/`
- VLDB:
  - "International Conference on Very Large Data Bases"
  - 2020 in Tokyo: `https://vldb2020.org/`
- ICDE:
  - "IEEE International Conference on Data Engineering"
  - 2020 in Dallas: `https://www.utdallas.edu/icde/`
- EDBT:
  - "International Conference on Extending Database Technology"
  - 2020 in Copenhagen:
    `https://diku-dk.github.io/edbticdt2020/`

Most Important DB Theory Conferences

- ACM PODS:
  - "Symposium on Principles of Database Systems"
  - always in conjunction with SIGMOD
  - 2020 in Portland: `https://sigmod2020.org/`
- ICDT:
  - "International Conference on Database Theory"
  - since 2009 in conjunction with EDBT
  - 2020 in Copenhagen:
    `https://diku-dk.github.io/edbticdt2020/`

Most Important DB-Journals

- **ACM TODS:**
  - "ACM Transactions on Database Systems"
  - free access from TUWIEN domain via ACM digital library
  - `http://portal.acm.org/dl.cfm` → "Transactions" →
    "ACM Transactions on Database Systems (TODS)"

- **VLDB Journal**
  - free access from TUWIEN domain via University library
  - `http://www.ub.tuwien.ac.at/` → "E-Journals" →
    "Universitätsbibliothek der TU Wien" → search for "VLDB Journal"

- **IEEE TKDE:**
  - "IEEE Transactions on Knowledge and Data Engineering"
  - free access from TUWIEN domain via IEEE Xplore
  - `http://ieeexplore.ieee.org/Xplore/guesthome.jsp` →
    "Journals & Magazines"

Access To Articles

Access from the TUWIEN domain is free to (almost) all major conferences and journals.

- SIGMOD and PODS proceedings
  - free access from TUWIEN domain via ACM digital library
  - `http://portal.acm.org/dl.cfm` → "Proceedings" → {SIGMOD, PODS}
- VLDB proceedings:
  - free access from anywhere
  - `http://www.vldb.org/` → "VLDB Conferences"

Access To Articles (continued)

- ICDE proceedings
  - free access from TUWIEN domain via IEEE Xplore Digital Library
  - `http://www.ieee.org/web/publications/xplore/` $\rightarrow$
    search for "ICDE"
- EDBT and ICDT proceedings:
  - originally Springer LNCS Series or ACM proceedings
    $\Rightarrow$ free access from TUWIEN domain via University library "eBooks"
    or ACM digital library
  - EDBT: open proceedings since 2014
  - ICDT: LIPIcs (open access) since 2015

Comfortable Search & Access via DBLP

- DBLP "Computer Science Bibliography"

- contains information on (almost) all relevant publications

- Overview: `http://www.informatik.uni-trier.de/~ley/db`

- Google-search, e.g., "DBLP $<$author$>$" or "DBLP $<$conference$>$"

- access to the article: via "electronic edition' icon

- free access from TUWIEN-domain as described above
  (e.g., ACM digital library, Springer Verlag, etc.)

Alternative Search Methods

- Citeseer, e.g., Google-search: "citeseer <title of article>"
- `http://scholar.google.com` (keyword search)
- Authors' Homepages