Exercise sheet 1 (WS 2022)

3.0 VU Data Modelling / 6.0 VU Database Systems

About the exercises

General information

In this part of the exercises you are asked to create a small database using EER-diagrams, transform EER-diagrams into a relational schema and make yourself familiar with relational algebra and relational calculus.

We recommend you to solve the problems **on your own** (it is a good preparation for the exam – and also for your possible future job – to carry out the tasks autonomously). Please note that if we detect duplicate submissions or any plagiarism, both the "original" and the "copy" will be awarded 0 points.

Your submission must consist of a single, typset PDF document (max. 5MB). **We do not accept PDF files with handwritten solutions.**

In total there are 8 tasks and at most 15 points that can be acchieved on the entire sheet.

Deadlines

until 07.11.	**12:00pm**	Upload your solutions to TUWEL
24.11.	13:00pm	Evaluation and feedback is provided in TUWEL

Further questions – TUWEL forum

If you have any further questions concerning the contents or organization, do not hestitate to ask them on TUWEL forum. **Under no circumstances should you post (partial) solutions on the forum!**

We also recommend that you get involved in the forum and actively discuss with your colleagues on the forum. From experience we believe that this helps all parties in the discussion greatly to improve their understanding of the material.

Exercise: EER-diagrams

Exercise 1 (Creating an EER-diagram) [3 points]

You are part of a development team for an online free-to-play role-playing game. The game is a multiplayer video game where two teams which both consist of 5 players compete against each other in matches to collect points. Players collect points plus several different achievements by winning matches, and there are rankings where where players of different leagues are listed and ranked by their points.

Your task is to develop a relational database for this game.

Draw an EER-diagram based on the available information (see next page). Use the notation presented in the lecture and the (min, max) notation. NULL values are not allowed and redundancies should be avoided. Sometimes it might be necessary to introduce additional keys.

A possible software for creating the EER-diagram is DIA (`http://wiki.gnome.org/Apps/Dia`, binaries at `http://dia-installer.de`; Attention: select ER in the diagram editor!). Of course you are also allowed to create the diagram with any other suitable software.

Description of the issues:

A player has a unique name (NAME), a unique ID number (ID), a password (PASSWORD), and an email address (EMAIL). As the game is a role-playing game, there are several character classes, where each character class has a unique name (NAME), and depending on the skills of the individual class points in the categories power (POWER), intelligence (INTELLIGENCE), and agility (AGILITY). Every player has to decide to choose exactly one character class when registering.

Players can participate in multiple teams, wher each team consists of exactly 5 players. Teams are identified by their unique team id (ID), and additionally the team's creation date (DATE) and the teamname (NAME) are stored. Two teams compete against each other in a Match, where one team takes the role of the hosting team and the second team takes the role of the visiting team. A match is identified by its hosting team, the date of the match (DATE), and the starting time (STARTTIME). The same two teams can compete against each other in multiple matches. For every match and team the result is stored (Win or Loss). Additionally, there is a difference between regular matches and tournament matches, where tournament matches additionally have an assigned tournament number (TNr).

Matches are being played on different servers and there is always exactly one referee watching the match. Each match is played on exactly one server, where a server is identified by its IP address (IP) and additionally has a given name (NAME). Referees have a unique ID number (SID) plus a name (SNAME). Teams and matches are further organized in different leagues, where a league is identified by a unique name (NAME) and must have at least two participating teams. Every team is associated to exactly one league and every match is associated to a single league, however, not all matches that are played by a team has to be associated to the same league as the team itself. Each team has a rank (RANK) in the ranking of its league. It is possible that multiple teams have the same rank in case of an equal number of collected points.

Playes can also collect several achievements by beating challenges in matches, where each achievement is identified by a unique name (NAME). Further, there are two special achievement types: Medals and team trophies, where team trophies are associated to exactly one team.

Exercise 2 (Semantics of EER diagrams) [1 point]

Consider the EER-diagram shown in Figure 1.

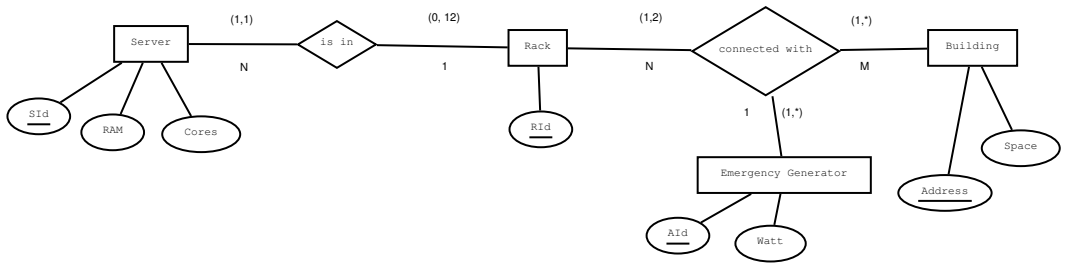

Figure 1: EER-diagram for exercise 2

In the ER diagram, both the notation by means of functionalities, as well as the (min, max) notation is used.
(*note:* this is not done in practice.)
Therefore, the diagram contains more information compared to the use of only one notation.

- Specify a specific relationship in the diagram where omitting one of the two notations causes a loss of information.
- For the chosen relationship type explain which notation, when omitted, leads to the loss of information.
- Explain briefly in your own words which information can no longer be represented.
- Provide a concrete example of the lost information, i.e. for the type of relationship you have chosen, specify an instance that violates (at least) one condition expressed by the omitted notation, but satisfies all requirements by the remaining notation.

Exercise 3 (Construct a relational schema) [2 points]

Construct a relation schema according to the EER-diagram given in Figure 2. `NULL` values are not allowed (you can assume that all attributes specified for an entity type exist for all entities of this type, i.e., the definedness of all attributes is 100%). Create as few relations as possible without introducing any redundancies. For each relation clearly mark the primary keys by underlining the corresponding attributes and display foreign keys in italics.

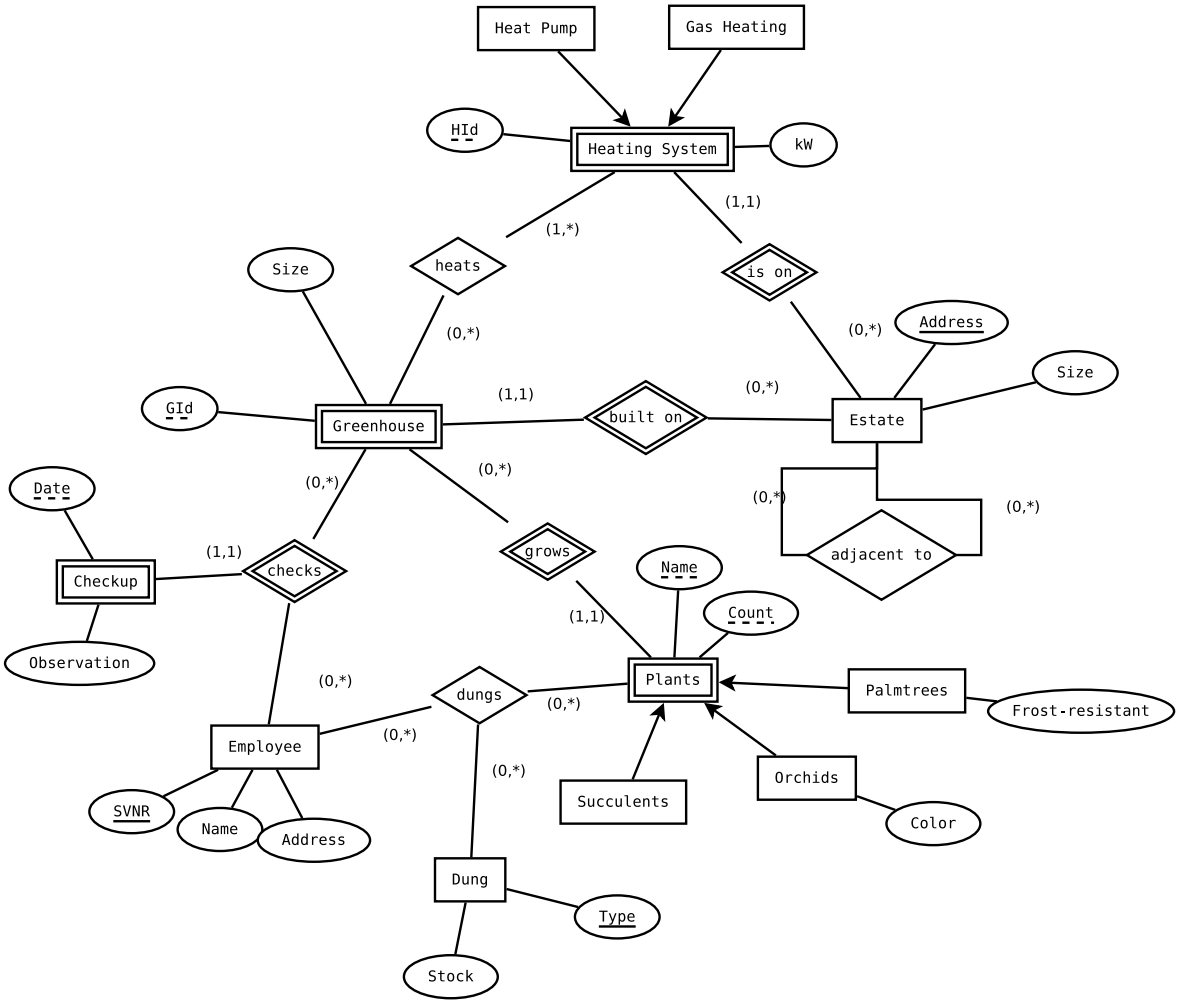

Figure 2: EER-diagram for exercise 3

Exercises: Relational Algebra - Relational Calculus

To help with typesetting the solutions to the following exercises, we compiled a list of the most important symbols for the Relational Algebra at `http://dbai.tuwien.ac.at/education/dm/resources/symbols.html`. You can copy and paste them into your Word/ LibreOffice/ OpenOffice/... document. In addition, the corresponding LaTeX commands are listed as well.

Exercise 4 (Evaluation) [0.5 points]

Consider the four relations below.

Manufacturer		
name	**country**	**founded**
BMW	Germany	1916
Ford	USA	1903
Renault	France	1898
Fiat	Italy	1899

Car		
name	**horsepower**	**year**
Renault 10	46	1965
BMW M1	277	1978
Fiat Nuova 500	13.5	1957
Ford Model A	40	1928
Fiat 1500 Spider	67	1963

producedBy	
by	**cname**
Ford	Ford Model A
BMW	BMW M1
Fiat	Fiat 1500 Spider
Renault	Renault 10
Fiat	Fiat Nuova 500

Sale		
sname	**date**	**insurance**
Renault 10	2002-09-09	full
BMW M1	1999-01-03	classic
Fiat Nuova 500	2010-06-08	full
Ford Model A	2014-10-08	full
Fiat Nuova 500	1995-04-03	classic

Provide the results of the following queries over these relations.

(a) (1)

$$q_1 = \Big(\pi_{name}\big(\sigma_{founded<1900}\big(\texttt{Manufacturer}\big)\big) \rtimes \rho_{name\leftarrow by}\big(\texttt{producedBy}\big)\Big)$$

(2)

$$q_2 = \rho_{sname\leftarrow name}\Big(\pi_{name,horsepower}\big(\texttt{Car}\big) \bowtie \rho_{name\leftarrow sname}\Big(\sigma_{insurance=full}\big(\texttt{Sale}\big)\Big)\Big)$$

(3)

$$q_3 = \sigma_{cname=sname}\big(q_1 \times q_2\big)$$

(b)

$$\{m.country, c.horsepower \mid\ m \in \texttt{Manufacturer} \wedge c \in \texttt{Car} \wedge$$
$$\exists p \in \texttt{producedBy}\Big(p.by = m.name \wedge$$
$$\exists o \in \texttt{Car}(o.name = p.cname \wedge o.horsepower > 100)\Big)\}$$

Exercise 5 (Equivalences) [2 points]

Consider the following relation schemata $R(\underline{A}CE)$, $S(BC\underline{D})$, $T(BE\underline{F})$ and the pairs q_i, q_j of expressions in relational algebra below. For every pair of expressions:

- Verify, whether the two expressions are equivalent (i.e. whether they produce the same result on all possible instances of the relation schemata). You may assume that no `NULL`-values occur in any instance of the schemata.
- Justify your answer with a brief **explanation**.
- In case the two expressions are *not* equivalent but valid expressions, additionally provide a **counterexample**. (A counterexample consists of the concrete instances of the affected relation schemata and the results of both expressions over these instances.)
  In case one of the expressions is not a valid expression in relational algebra you do not have to provide a counterexample, hence in this case an explanation suffices.

(a) $q_1 \colon \sigma_{B>3}(\pi_B(S) - (\pi_B(S) - \pi_B(T)))$ and
$q_2 \colon \sigma_{B>3}(\pi_B(S)) \cap \sigma_{B>3}(\pi_B(T))$

(b) $q_3 \colon \sigma_{F>3}(\rho_{F\leftarrow A}(R)) \cup (\rho_{A\leftarrow E}(S \bowtie T) \rtimes \rho_{F\leftarrow A}(R))$ and
$q_4 \colon \sigma_{F>3}(\rho_{F\leftarrow A}(R) \cup \rho_{A\leftarrow E}(S \bowtie (T \rtimes \rho_{F\leftarrow A}(R))))$

(c) $q_5 \colon \sigma_{B=3}(S \overline{\bowtie} T) \bowtie \pi_{BCE}(\sigma_{B=3}(\rho_{B\leftarrow A}(R)))$ and
$q_6 \colon \sigma_{B=3}(S) \overline{\bowtie} \pi_{BCE}(\sigma_{B=3}(T) \bowtie \rho_{B\leftarrow A}(\sigma_{A=3}(R)))$

Exercise 6 (Answer Sizes) [1.5 points]

Consider the relational schemas $R(\underline{A}BC)$, $S(AB\underline{CE})$, and $T(\underline{A}C\underline{D}E)$ as well as an instance of every schema, where there are $|R|$ tuples in the instance of R, $|S|$ tuples in the instance of S, and $|T|$ tuples in the instance of T.

- Provide the minimal and maximal size (= number of tuples) of the following expressions in Relational Algebra for the given values of $|R|$, $|S|$, $|T|$.
- Justify your answer.
- For both, the smallest and biggest possible answer size, provide concrete instances of the schemas (with R, S, and T having $|R|$, $|S|$, and $|T|$ tuples, respectively) over which the query returns an answer with the minimal/maximal number of tuples.

(a) $q_1 \colon \left(\rho_{D\leftarrow C}(R) \overline{\bowtie} S\right) \bowtie \rho_{D\leftarrow C}(R)$ (mit $|R| = 2$ und $|S| = 5$)

(b) $q_2 \colon \Big(\big(\pi_{C,E}(S) - \pi_{C,E}(T)\big) \cap \pi_{C,E}(S)\Big) \overline{\bowtie} \pi_{A,E}\Big(\rho_{E\leftarrow C}\big(R\big)\Big)$ (mit $|R| = 2$, $|S| = 4$ und $|T| = 5$)

(c) $q_3 \colon \sigma_{A=1}(R) \rtimes \Big(\pi_{ADE}\big(\rho_{B\leftarrow C}(T)\big) \times \pi_{BC}(S)\Big)$
(mit $|R| = 4$, $|S| = 1$ und $|T| = 4$)

Exercise 7 (Query Languages) [1 point]

Consider the relational schema $R(\underline{A}B)$, $S(AB\underline{CD})$, and $T(AC\underline{E}F)$.

In the following exercises you are given a query in one of the query languages from the lecture. Your task is to translate the query into the two other query languages that were discussed in the lecture.

(a) Translate the query

$$S \ltimes \pi_A(\sigma_{B \neq -1}(R))$$

into tuple relational and domain relational calculus.

(b) Translate the query

$$\{[a, b] \mid \exists c, d([a, b, c, d] \in S \ \wedge [c, d] \in R)\}$$

into tuple relational calculus and relational algebra.

(c) Translate the query

$$\{[s.A, s.D] \mid s \in S \wedge \exists t(t \in T \wedge t.A > s.B)\}$$

into domain relational calculus and relational algebra.

Exercise 8 (Formalizing Queries) [4 points]

A home-automation software uses a database to organize devices and automation system. The database has the following schema (primary keys are underlined, foreign keys are written in *italics*):

```
Device      (DId, Name, Model)
ZigbeeDevice(DId :  Device.DId, RSSI, PowerSource, LastSeen)
LightDevice (DId :  ZigbeeDevice.DId, Brightness, ColorTemperature)
Coordinator (DId :  Device.DId, Name)
ZigbeeGroup (DId :  Coordinator.DId, GId, Name)
connected   (Coordinator:  Coordinator.DId, Device:  ZigbeeDevice.DId)
link        (Dev1:  ZigbeeDevice.DId, Dev2:  ZigbeeDevice.DId, quality )
grouped     (GroupCoordinator:  ZigbeeGroup.DId, GroupId:
   ZigbeeGroup.GId, Device:  ZigbeeDevice.DId)
```

(In the following you may use suitable (unique) abbreviations for relations and table names.)

Express the queries described below in **relational algebra**, the **tuple relational calculus** and the **domain relational calculus**.

(a) Give out the RSSI and Brightness of all LightDevices which are a part of the ZigbeeGroup with the name 'Deckenleuchten Badezimmer'.

(b) Give out the names of all ZigbeeDevices that are connected (via link) with at least two other ZigbeeDevices

(c) Give out the name of all ZigbeeDevices which are connected with the coordinator that has the name 'Conbee II' and where PowerSource = 'Battery'