

Einführung in Datenbanken

— Übungsblatt 10 (Aggregationen, UNION, RA) —

Organisatorische Hinweise:

- **Einsendeschluss ist Montag, der 25.01.2021, 18⁰⁰.**
- Ihre Lösungen laden Sie bitte in die Übungsplattform in StudIP hoch ([StudIP-Eintrag der Vorlesung], Reiter „Übungsplattform“, dann auf „Weiter zu ... Übungsplattform“).
- Bitte geben Sie die Hausaufgaben als einfache Textdatei mit der Endung `.sql` ab, bzw. mit der Endung `.txt` für Teil c) (Relationale Algebra). Alle Zusatzangaben müssen mit `--` als Kommentar gekennzeichnet werden.
- Hausaufgaben müssen einzeln bearbeitet werden. „Zu ähnliche Lösungen“ führen automatisch zu 0 Punkten für alle Beteiligten. Das gilt auch dann, wenn Sie nicht direkt abgeschrieben haben, sondern nur zufällig die gleiche Quelle benutzt haben.
- Probieren Sie Ihre Anfragen mit dem Adminer aus:

```
[https://dbs.informatik.uni-halle.de/edb?pgsql=db&
    username=student.gast&db=postgres&ns=]
```

Die Zugangsdaten unserer Installation stehen in StudIP, Reiter „Adminer“.

- Anfragen, die Syntaxfehler enthalten, werden höchstwahrscheinlich mit 0 Punkten bewertet. Nutzen Sie rechtzeitig das Forum in StudIP, um Hilfe zu bekommen.
- Verwenden Sie in Ihren Anfragen nur Informationen, die in der Aufgabenstellung erwähnt sind. Ihre Anfrage muss mit beliebigen DB-Zuständen funktionieren, nicht nur dem Beispiel-Zustand.
- Schreiben Sie keinen Schema-Namen in die Anfrage. Bei der automatischen Vorkorrektur heißt das Schema eventuell anders.
- Es können auch Punkte für schlechten Stil abgezogen werden. Z.B. sollten Sie sinnvolle Namen für die Tupelvariablen wählen. Bemühen Sie sich auch um gute Formatierung. Vermeiden Sie unnötige Verbunde oder andere unnötigen Komplikationen.
- Ihre Anfragen sollen keine Duplikate liefern (in keinem Zustand, der die Integritätsbedingungen erfüllt). Ein überflüssiges `DISTINCT` kostet aber auch einen Punkt.

Hausaufgabe 10 (15 Punkte)

Wählen Sie im Adminer das Schema „empdept_public“. Die ist eine bekannte Beispiel-Datenbank von Oracle und enthält Informationen über Angestellte und Abteilungen:

- `dept(deptno, dname, loc)`
- `emp(empno, ename, job, mgro→emp, hiredate, sal, commo, deptnoo→dept)`

Sie bekommen jeweils 5 Punkte für jede der folgenden drei Anfragen.

- a) Numerieren Sie die Angestellten nach Abteilungen (`deptno`), und in jeder Abteilung alphabetisch nach dem Namen durch. Sie dürfen den Angestellten mit einem Nullwert in `deptno` ignorieren. Es gibt einen Bonuspunkt, wenn Sie Angestellte ohne Abteilung ganz hinten einsortieren (innerhalb dieser Gruppe wieder alphabetisch). Sie sollen also (mit Bonuspunkt) diese Tabelle erzeugen:

id	deptno	ename
1	10	CLARK
2	10	KING
3	10	MILLER
4	20	ADAMS
5	20	FORD
6	20	JONES
7	20	SCOTT
8	30	ALLEN
9	30	BLAKE
10	30	JAMES
11	30	MARTIN
12	30	TURNER
13	30	WARD
14	NULL	SMITH

14 Datensätze

Die letzte Zeile darf auch fehlen (dann kein Bonuspunkt). Sortieren Sie die Ausgabe nach der so erzeugten Spalte „id“.

Sie dürfen in dieser Aufgabe nur Sprachkonstrukte verwenden, die in der Vorlesung behandelt wurden. Tipp: Zählen Sie, wie viele Angestellte vor dem aktuellen Angestellten kommen müssen (oder gleich sind). Sie müssen die verlangte Sortierung also durch eine logische Bedingung ausdrücken.

Sie dürfen voraussetzen, dass es keine zwei Angestellten mit gleichem Namen gibt.

- b) In der Angestellten-Tabelle `emp` enthält die Spalte `mgr` die Nummer (`empno`) des direkten Vorgesetzten. Es sollen jetzt auch die indirekten Vorgesetzten-Beziehungen berechnet werden unter der Voraussetzung, dass es maximal vier Hierarchie-Ebenen gibt, d.h. `e1` hat den direkten Vorgesetzten `e2`, der wiederum hat den direkten Vorgesetzten `e3`, und dieser hat schließlich den Vorgesetzten `e4`.

Dies wäre eigentlich eine gute Anwendung für rekursive Sichten, aber Sie dürfen keine Rekursion einsetzen, da diese erst in der Fortsetzungs-Vorlesung behandelt wird. Sie können aber z.B. `UNION` oder `UNION ALL` verwenden.

Die Ausgabe soll vier Spalten haben:

- `chef`: Name des Chefs,
- `ang`: Namen des Untergebenen (ggf. indirekt),
- `diff`: den Abstand in Hierarchiestufen (1, 2 oder 3),
- `path`: eine Zeichenkette, die mit dem Namen des Chefs beginnt und den Namen des Untergebenen endet und dazwischen durch senkrechten Strich „|“ getrennt ggf. die Zwischen-Chefs enthält.

Sie können natürlich voraussetzen, dass es in der Hierarchie keine Zyklen gibt. Sie können außerdem voraussetzen, dass es keine zwei Angestellten mit gleichem Namen gibt. Die erwartete Antwort ist:

chef	ang	diff	path
FORD	SMITH	1	FORD SMITH
BLAKE	ALLEN	1	BLAKE ALLEN
BLAKE	WARD	1	BLAKE WARD
KING	JONES	1	KING JONES
⋮	⋮	⋮	⋮
JONES	SMITH	2	JONES FORD SMITH
KING	ALLEN	2	KING BLAKE ALLEN
KING	WARD	2	KING BLAKE WARD
KING	MARTIN	2	KING BLAKE MARTIN
KING	SCOTT	2	KING JONES SCOTT
KING	TURNER	2	KING BLAKE TURNER
JONES	ADAMS	2	JONES SCOTT ADAMS
KING	JAMES	2	KING BLAKE JAMES
KING	FORD	2	KING JONES FORD
KING	MILLER	2	KING CLARK MILLER
KING	SMITH	3	KING JONES FORD SMITH
KING	ADAMS	3	KING JONES SCOTT ADAMS

c) Die folgende Anfrage ist in relationaler Algebra zu schreiben. Sie bezieht sich auf die Beispiel-Datenbank der Vorlesung:

- $\text{STUDENTEN}(\underline{\text{SID}}, \text{VORNAME}, \text{NACHNAME}, \text{EMAIL}^o)$
- $\text{AUFGABEN}(\underline{\text{ATYP}}, \underline{\text{ANR}}, \text{THEMA}, \text{MAXPT})$
- $\text{BEWERTUNGEN}(\underline{\text{SID}} \rightarrow \text{STUDENTEN}, (\underline{\text{ATYP}}, \underline{\text{ANR}}) \rightarrow \text{AUFGABEN}, \text{PUNKTE})$

Es ist empfohlen, die Anfrage in Relax auszuprobieren:

[<http://dbis-uibk.github.io/relax/calc/gist/8dc2652578ee12ae756a234c4cf21b3f>]

Welche Studierenden haben für irgendeine Aufgabe die volle Punktzahl bekommen? Geben Sie Vorname und Nachname des Studierenden sowie Aufgabentyp und Aufgabennummer aus:

STUDENTEN.VORNAME	STUDENTEN.NACHNAME	BEWERTUNGEN.ATYP	BEWERTUNGEN.ANR
'Lisa'	'Weiss'	'H'	1

Der Präfix der Spaltennamen mit dem Relationenname darf auch anders sein oder fehlen (Relax druckt einen solchen Präfix).

Tipps:

- Es ist hier empfohlen, einen natürlichen Verbund zu nutzen, z.B.:

`STUDENTEN join BEWERTUNGEN.`

- Ein Beispiel für die Syntax der Selektion ist:

`sigma SID = 101 (STUDENTEN).`

Leider ist der Index in Relax nicht speziell markiert. In der Vorlesung würde es $\sigma_{\text{SID} = 101}(\text{STUDENTEN})$ geschrieben.

- Ein Beispiel einer Projektion ist:

`pi SID, EMAIL (STUDENTEN).`

Dies entspricht: $\pi_{\text{SID,EMAIL}}(\text{STUDENTEN})$.

- Denken Sie daran, dass die Groß-/Kleinschreibung der Spalten- und Tabellennamen bei Relax wichtig ist.