

Klausur Datenbanken
Wintersemester 2017/2018
Prof. Dr. Wolfgang May
26. Februar 2018, 11-13 Uhr
Bearbeitungszeit: 90 Minuten

Vorname:

Nachname:

Matrikelnummer:

Bei der Klausur sind **keine Hilfsmittel** (Skripten, Taschenrechner, etc.) erlaubt. Handies müssen ausgeschaltet sein. Papier wird gestellt. Benutzen Sie nur die **ausgeteilten**, zusammengehefteten **Blätter** für Ihre Antworten. Schreiben Sie mit blauem/schwarzem Kugelschreiber, Füller, etc.; Bleistift ist nicht erlaubt.

Zum **Bestehen** der Klausur sind **45** Punkte hinreichend.

- meine Note soll mit Matrikelnummer so bald wie möglich auf der Vorlesungs-Webseite veröffentlicht werden.
- meine Note soll nicht veröffentlicht werden; ich erfahre sie dann aus FlexNow oder beim zuständigen Prüfungsamt.

	Max. Punkte	Erreichte Punkte
Aufgabe 1 (ER-Modell)	15	
Aufgabe 2 (Transformation in das Relationale Modell)	18	
Aufgabe 3 (SQL und Relationale Algebra)	45	
Aufgabe 4 (Verschiedenes)	12	
Summe	90	

Note:

Themenstellung: GöVB

Alle Klausuraufgaben basieren auf einem gemeinsamen “Auftrag”: In der Klausur soll eine Datenbank für die Göttinger Verkehrsbetriebe entworfen werden, die deren Ressourcenplanung (Fahrer, Busse, Linien) unterstützt:

1. Für jeden Fahrer ist sein Name und sein Geburtsdatum gespeichert (es wird angenommen, dass der Name eindeutig ist).

Ernst Meier ist am 26.2.1960 geboren. *Heinz Müller* ist am 3.10.1989 geboren.

2. Jeder Bus hat eine Nummer. Außerdem ist gespeichert, von welcher Marke der Bus ist, wieviele Sitz- und Stehplätze er hat, und ob er ein Gelenkbus ist.

Bus Nr. 42 ist ein *Mercedes*-Gelenkbus, hat 39 Sitzplätze und 106 Stehplätze; Bus Nr. 43 ebenfalls. Bus Nr. 123 ist ein *MAN*-Bus, hat 25 Sitzplätze und 60 Stehplätze.

3. Jede Haltestelle hat einen eindeutigen Namen und eine Adresse.

Die Haltestelle *Nikolausberg* hat die Adresse *Auf der Lieth 5*. Die Haltestelle *Bahnhof* hat die Adresse *Bahnhofplatz*. Die Haltestelle *Zienterrassen* hat die Adresse *Grete-Henry-Strasse 8*.

4. Für jede Linie ist ihre Nummer gespeichert, ihre Start- und Endhaltestelle, die Uhrzeit, wann morgens der erste Bus an der Starthaltestelle abfährt, ihre Umlaufdauer (d.h., wie lange es dauert, bis ein Bus wieder an der Starthaltestelle angekommen ist, und die nächste Runde beginnt) und ob sie mit Gelenkbussen befahren werden darf: Linie 21 fährt von *Nikolausberg* zu den *Zienterrassen*, hat eine Umlaufdauer von 120 Minuten und darf mit Gelenkbussen befahren werden; der erste Bus fährt morgens um 6:25 Uhr. Linie 33 fährt vom *Holtenser Berg* zum *Klinikum*, hat eine Umlaufdauer von 60 Minuten, darf *nicht* mit Gelenkbussen befahren werden; erste Bus fährt morgens um 6:15 Uhr.

Außerdem ist gespeichert, an welchen Haltestellen eine Linie hält, und nach wievielen Minuten nach ihrer Abfahrt sie diese erreicht. Die Starthaltestelle ist dabei mit “0 Minuten” angegeben (es ist nur jeweils eine Richtung gespeichert; es wird angenommen, dass die Rückfahrt genau umgekehrt erfolgt):

Linie 21 fährt von *Nikolausberg* ab, hält (unter anderem) am *Faßberg* nach 5 Minuten, an der *Tammannstrasse* nach 10 Minuten, am *Bahnhof* nach 26 Minuten, und kommt an ihrer Endhaltestelle, den *Zienterrassen*, nach 45 Minuten an. Keine Linie kommt auf ihrer Strecke mehrmals an derselben Haltestelle vorbei.

5. Es ist gespeichert, welche Fahrer welche Busse steuern dürfen; d.h. dass sie eine entsprechende Schulung absolviert haben, und wann dies war.

Ernst Meier hatte Schulungen für die Busse mit den Nummern 42, 43 (jeweils am 10.3.2015) und 123 (am 24.10.2016).

6. Der Einsatzplan soll gespeichert werden: welcher Fahrer an welchem Tag mit welchem Bus welche Linie fährt, ob er für die Frühschicht oder für die Spätschicht eingeteilt ist, und wieviele Minuten nach dem ersten Bus auf dieser Linie er fährt. (Kommentar: bei einer halbstündig verkehrenden Linie mit einer Umlaufdauer von 120 Minuten sind dann z.B. pro Schicht 4 Fahrer mit 0, 30, 60, 90 Minuten Abstand eingeteilt). Ein Fahrer darf nicht am selben Tag zur Früh- und zur Spätschicht eingeteilt werden.

Ernst Meier und *Heinz Müller* sind beide am 26.2.2018 für die Frühschicht auf der Linie 21 eingeteilt: Meier fährt morgens als erster (mit Bus Nr. 42), Müller 30 Minuten nach ihm (mit Bus Nr. 43). Am 27.2.2018 ist *Ernst Meier* für die Spätschicht auf der Linie 21 mit Bus Nr. 123 als erster eingeteilt.

Aufgabe 1 (ER-Modell [15 Punkte])

Entwickeln Sie ein ER-Modell für das Szenario. Geben Sie darin die Schlüsselattribute sowie die Beziehungskardinalitäten an.

- Gelenkbus kann anstatt als (true/false-)Attribut auch als Subklasse von Bus modelliert werden (hat aber dann keine speziellen neuen Attribute, also nicht sehr interessant).
- Wenn man “Einsatzplan” nicht als dreistellige Beziehung, sondern als schwachen Entitätstyp modelliert, muß man an dieser Stelle schon über den (lokalen) Schlüssel und die identifizierenden Beziehungen entscheiden (ansonsten erst in Aufgabe 2; siehe Diskussion dort).

Aufgabe 2 (Transformation in das Relationale Modell [18 Punkte])

- a) Lösen Sie diesen Aufgabenteil auf dem *letzten* Blatt und trennen dieses ab (und geben es am Ende mit ab!). Dann haben Sie dieses Blatt separat zugreifbar um später damit die Aufgaben 2b, 3 und 4 (SQL, Relationale Algebra+SQL, Diverses) zu lösen.

Geben Sie an, welche Tabellen (mit Attributen, Schlüsseln etc.) Ihre Datenbank enthält (keine SQL CREATE TABLE-Statements, sondern einfach grafisch). (12 P)

Markieren Sie dabei auch Schlüssel (durch unterstreichen) und Fremdschlüssel (durch überstreichen).

Geben Sie die Tabellen mit jeweils mindestens zwei Beispieldupeln (z.B. denen, die sich aus dem Aufgabentext ergeben, und weiteren erfundenen) an.

Vorgegeben wird bereits die Tabelle, in der gespeichert ist, für welche Busse die Fahrer Schulungen absolviert haben (sie wird in Aufgabe 4 wieder gebraucht):

Schulung		
<u>Fahrer</u>	<u>Bus</u>	Datum
Ernst Meier	42	10.3.2015
Ernst Meier	43	10.3.2015
Ernst Meier	123	24.10.2016
:	:	:

- b) Geben Sie das CREATE TABLE-Statement für diejenige Tabelle, in der die Halte der einzelnen Linien gespeichert sind, so vollständig wie möglich an (6 P).

Aufgabe 3 (SQL und Relationale Algebra [45 Punkte])

Verwenden Sie für diese Aufgabe die von Ihnen entworfene relationale Datenbasis. Keine der Antworten soll Duplikate enthalten.

- a) Geben Sie eine SQL-Anfrage *und* einen Algebra-Ausdruck oder -Baum an, die die Namen der Fahrer und die Nummern der Busse ausgeben, so dass derjenige Fahrer am 26.2.2018 mit einem Gelenkbus an der Haltestelle “Tammannstrasse” hält. (3+3P)
- b) Geben Sie eine SQL-Anfrage an, die zu jeder Haltestelle ausgibt, wieviele Linien an dieser Haltestelle halten. (2 P)
- c) Geben Sie eine SQL-Anfrage an, die den/die Namen derjenigen Haltestelle/Haltestellen ausgibt, an der/denen die *meisten* Linien halten. (4 P)
- d) Geben Sie eine SQL-Anfrage *und* einen Algebra-Ausdruck oder -Baum an, die die Namen aller Fahrer ausgeben, die nach ihrem Einsatzplan *nie* an der Haltestelle *Bahnhof* halten. (3+4 P)
- e) Geben Sie eine SQL-Anfrage *und* einen Algebra-Ausdruck oder -Baum an, die die Namen aller Haltestellen angeben, an denen am 26.2.2018 *alle Mercedes-Gelenkbusse* mindestens einmal vorbeikommen. (5+4P)
- f) Geben Sie eine SQL-Anfrage an, die ergibt, um welche Uhrzeit *Heinz Müller* am 26.2.2018 zum ersten Mal mit einem Bus am Bahnhof hält, falls er Frühschicht hat. Nehmen Sie für diese Aufgabe an, dass man eine Zahl (als Minuten) einfach zu einem Zeit-Wert addieren kann. (4 P)
- g) Wie sieht das Ergebnis der Anfrage aus Teil (f) aus, wenn *Heinz Müller* an diesem Tag frei hat? (2 P)
- h) Geben Sie eine SQL-Anfrage an, die für jedes Paar von Linien, die mindestens eine Haltestelle gemeinsam haben, angibt, wieviele Haltestellen sie gemeinsam haben. (5 P)
- i) Ein bisschen Theorie (6 P).
 Sei $R(\bar{X})$ eine Relation mit Primärschlüssel $\bar{Y} \subseteq \bar{X}$.
 Seien $\bar{V} \subseteq \bar{X}$ und $\bar{W} \subseteq \bar{X}$ Teilmengen der Attribute mit $\bar{V} \cup \bar{W} = \bar{X}$ und $\bar{V} \cap \bar{W} \neq \emptyset$.
- i) Zeigen Sie, dass im allgemeinen **nicht** $\pi[\bar{V}](R) \bowtie \pi[\bar{W}](R) = R$ gilt, sondern meistens $\pi[\bar{V}](R) \bowtie \pi[\bar{W}](R) \supsetneq R$ ist. (3 P)
- ii) Unter welchen Bedingungen (an \bar{V} und \bar{W}) gilt garantiert $\pi[\bar{V}](R) \bowtie \pi[\bar{W}](R) = R$?
 (mit Begründung) (3 P)

Name:

MatNr.:

Aufgabe 4 (Verschiedenes [12 Punkte])

- a) *Jeden Abend* sollen die Daten des vergangenen Tages im Einsatzplan gelöscht werden. Wie macht man das am besten (SQL-Statement, und was könnte man weitergehend dazu anmerken?) (3 P)
- b) Bus Nr. 123 bekommt als Pilotprojekt eine neue, teilautonome Fahrautomatik. *Ernst Meier*, der diesen Bus ja schon fuhr, bekam am 24.2.2018 eine neue Schulung für diesen Bus.

Jemand führt das folgende SQL-Statement aus (siehe die in Aufgabe 2b vorgegebene Tabelle "Schulung"):

```
INSERT INTO Schulung VALUES ('Ernst Meier', 123, '24.2.2018');
```

Was passiert – begründen Sie ihre Aussage? (2 P)

- c) Es darf nicht passieren, dass in den Einsatzplan eine Fahrer-Bus-Kombination eingetragen werden kann, die nicht zulässig ist (d.h., wo der Fahrer keine Schulung für diesen Bus hatte). Wie kann man das datenbankseitig *sicher verhindern*? (4 P)
- d) Während des Updates aus Aufgabenteil (b) gibt es einen Stromausfall. Der Computer läßt sich danach wieder problemlos starten, physikalisch sind alle Bauteile in Ordnung.

Skizzieren Sie, wie sich die Datenbank danach wieder (vom Admin entsprechend auszuführen) in einen regulären Zustand bringen kann. (3 P)

Name:

MatNr.:

Name:

MatNr.:

[Trennen Sie dieses Blatt am besten vor Beginn der Bearbeitung ab]

Lösen Sie hier Aufgabe 2a