Aufgabe 1: Lösung

Die 3 Datenebenen heißen: (1 Punkt)
- externes Modell
- konzeptuelles Modell
- internes Modell

1) Das externe Modell: (3 Punkte)
Das externe Modell beschreibt die einzelnen Sichten der Benutzer. Jeder Benutzer soll nur die für ihn relevanten Daten sehen. Die Objekttypen und Beziehungstypen müssen nicht mit denen des konzeptuellen Modells identisch sein, sie müssen lediglich „inhaltlich" im konzeptuellen Modell enthalten sein. (2 Punkte)
Die Erstellung des entsprechenden Schemas hängt von den Anforderungen und auch von den Befugnissen (Zugriffsrechten usw.) des jeweiligen Benutzers ab (1 Punkt).

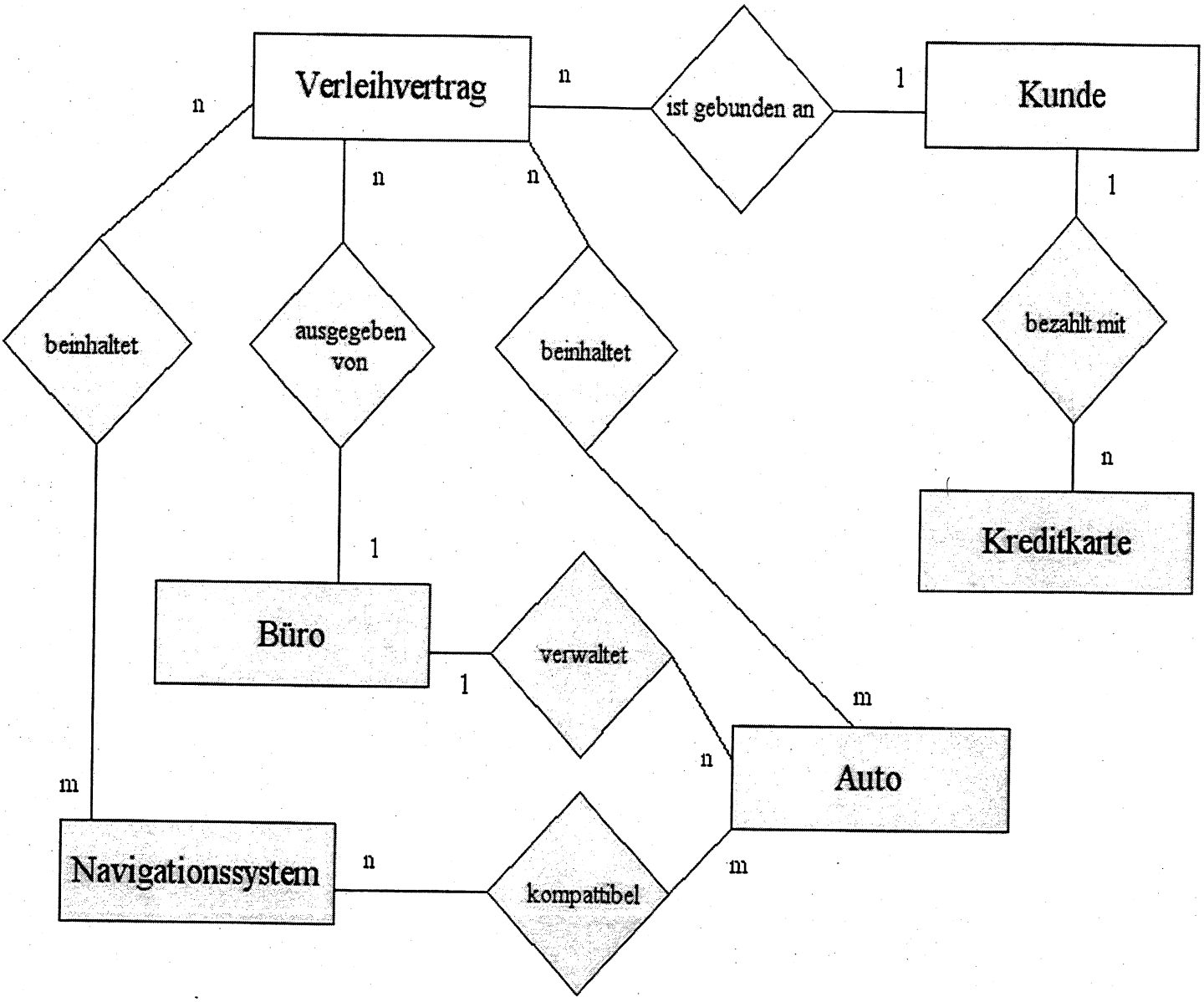
2) Das konzeptuelle Modell: (3 Punkte)
Das konzeptuelle Modell beschreibt die Gesamtheit aller Daten, die innerhalb der Datenbank verwaltet werden (1 Punkt). Dazu gehören die Festlegungen von Beziehungen, Integritätsbedingungen und Operationen (1 Punkt).
Das zu erstellende Schema hängt einzig und allein von den in der Realität existierenden Strukturen (die wesentlichen Daten eines Unternehmens und die wesentlichen Beziehungen zwischen den Daten) ab (1 Punkt).

3) Das interne Modell: (4 Punkte)
Das interne Modell beschreibt die vom Datenbankadministrator festgelegte physische Datenorganisation (1 Punkt). Es enthält u.a. alle Informationen über den Aufbau der abgespeicherten Daten, die Speicherung der Daten und über die Zugriffspfade (1 Punkt). Es muß genaue Festlegungen der folgenden Punkte enthalten (1 Punkt):
- Repräsentation von Attributwerten
- Aufbau gespeicherter Sätze
- Zugriffsmethoden auf Sätze
- zusätzliche Zugriffspfade (Indizes, Verkettungen, usw.)

Die Erstellung des internen Schemas hängt von den statischen Informationen über die Häufigkeit von Anwendungen, von Zugriffen auf Objekten, über Zeitbeschränkungen für Anwendungen, Verteilung von Werten für Attribute usw. ab (1 Punkt).

Datenunabhängigkeit: (3 Punkte)
Datenunabhängigkeit bedeutet, daß Anwendungsprogramme von Änderungen auf der internen und der konzeptuellen Ebene unberührt bleiben. Man unterscheidet physische und logische Datenunabhängigkeit. (1Punkt)
Physische Datenunabhängigkeit bedeutet Isolierung der Anwendungsprogramme vor Änderungen der physischen Datenorganisation. (1Punkt)
Logische Datenunabhängigkeit bedeutet Isolierung der Anwendungsprogramme vor Änderungen des konzeptuellen Modells. (1Punkt)

Aufgabe 2: Lösung



Anmerkungen: Alternativ können Verleihvertrag, Auto und Navigationssystem in einer dreistelligen Beziehung stehen. Auch ist es möglich, die Navigationssysteme und die Autos zu zu entsprechenden Typklassen in Beziehung zu setzen und zwischen diesen Klassen die Kompatibilitätsrelation aufzuspannen.

Aufgabe 3: Lösung

(kleingeschriebene Attribute sind optional und nur zur Verdeutlichung)

Mitarbeiter (M_NR, L_NR, K_NR, name, adresse,...)
Lehrgebiet (L_NR, F_NR, M_NR, bezeichnung,...) // M_NR ist die Relation zum Professor
Fachbereich(F_NR, bezeichnung,...)
Projekt (P_NR, bezeichnung, auftraggeber,...)
Kurs (K_NR, F_NR, titel,...)

Mitarbeiter_bearbeitet_Projekt (M_NR,P_NR)

Aufgabe 4:Lösung

Listen Sie Personalnummer und Name der C4-Professoren. (1 Punkt)

```
select PersNr, Name
from   Professoren
where  Rang = 'C4'
```

Zählen Sie alle Studenten mit Semesterzahlen zwischen 1 und 4. (1 Punkt)

```
select count(*)
from   Studenten
where  Semester >= 1 and Semester <= 4
```

alternativ:

```
where  Semester between 1 and 4
```

alternativ:

```
where  Semester in (1,2,3,4)
```

Listen Sie alle Vorlesungen, die im Titel den String `Daten` enthalten, klein oder groß geschrieben. (2 Punkte)

```
select *
from   Vorlesungen
where  upper(Titel) like '%DATEN'
```

Listen Sie Professorennummer, Name und Rang aller Professoren, absteigend sortiert nach Rang, innerhalb des Rangs aufsteigend sortiert nach Name. (2 Punkte)

```
select   ProfNr, Name, Rang
from     Professoren
order by Rang desc, Name asc
```

Listen Sie die Namen der Studenten mit ihren Vorlesungstiteln. (3 Punkte)

```
select s.Name, v.Titel
from       Studenten s, hoeren h, Vorlesungen v
where      s.MatrNr = h.MatrNr
and        h.VorlNr = v.VorlNr
```

Listen Sie Namen und Summe der SWS pro C4-Professor, sofern seine Durchschnitts-SWS größer als 3 ist. (4 Punkte)

```
select   p.Name, sum(SWS)
from     Vorlesungen, Professoren p, lesen l
where    VorlNr = l.VorlNr
    and l.ProfNr = p.ProfNr
    and p.Rang = 'C4'
group by l.ProfNr, p.Name
having   avg(SWS) > 3
```

Listen Sie alle Professoren, die keine Vorlesung halten. (3 Punkte)

```
select Name
from   Professoren
where  ProfNr not in ( select ProfNr from lesen )
```

alternativ:

```
select Name
from   Professoren p
```

```
where  not exists ( select *
                    from lesen
                    where ProfNr = p.ProfNr )
```

Listen Sie die Studenten mit der größter Semesterzahl. (3 Punkte)

```
select Name
from   Studenten
where  Semester >= all ( select Semester
                         from Studenten )
```

Listen Sie solche Studenten, die alle 4-stündigen Vorlesungen hören. (4 Punkte)

```
select s.*
from   Studenten s
where  not exists
  (select *
  from Vorlesungen v
  where v.SWS = 4 and not exists
    (select *
     from hoeren h
  where h.VorlNr = v.VorlNr and h.MatrNr = s.MatrNr))
```

Aufgabe 5:Lösung

a) (5 Punkte)
Die Relation Ware ist in 2NF, da die Nichtschlüsselattribute Lager_Bez und Lager_Leiter voll funktional abhängig sind vom Schlüssel Waren_Bez.

Ware ist nicht in 3NF, da es einen Schlüssel Waren_Bez gibt, einen Nichtschlüssel Lager_Bez und ein Nichtschlüsselattribut Lager_Leiter gibt, derart dass

Waren_Bez	-> Lager_Bez	gilt,
Lager_Bez	-> Lager_Leiter	gilt und
Lager_Bez	-> Waren_Bez	nicht gilt.

Damit ist Ware auch nicht in BCNF. (3 Punkte)

Die Relation Bestellung ist nicht in 2NF, da das Nichtschlüsselattribute L_Anschrift nicht voll funktional abhängig ist vom Schlüssel (Lieferanten_Bez, Waren_Bez, Best_Datum).
Damit ist Bestellung auch nicht in 3NF und somit auch nicht in BCNF. (2 Punkte)

b) (4 Punkte)
<u>Einfüge-Anomalie:</u> In der Relation BESTELLUNG kann kein neuer Lieferant mit Anschrift gespeichert werden, bei dem nicht zumindest eine Ware bestellt wurde. (2 Punkte)

<u>Lösch-Anomalie:</u> Wird die derzeit einzige Bestellung bei einem Lieferanten storniert, so gehen mit dem Löschen dieses Datensatzes alle Informationen über diesen Lieferanten verloren. (1,5 Punkte)

<u>Änderungs- Anomalie:</u> Ändern sich die Adresse eines Lieferanten, so muss in allen Einträgen von Bestellungen, die diesen Lieferanten enthalten, diese Änderung durchgeführt werden. Entspr: wird ein neuer Lagerleiter für ein Lager eingesetzt, so muss in allen WAREN-Sätzen, die Ware aus diesem Lager beschreiben, diese Änderung durchgeführt werden(1,5 Punkte)

c) (4 Punkte)
Ware (Waren_Bez, Lager_Bez, Lager_Leiter) -->

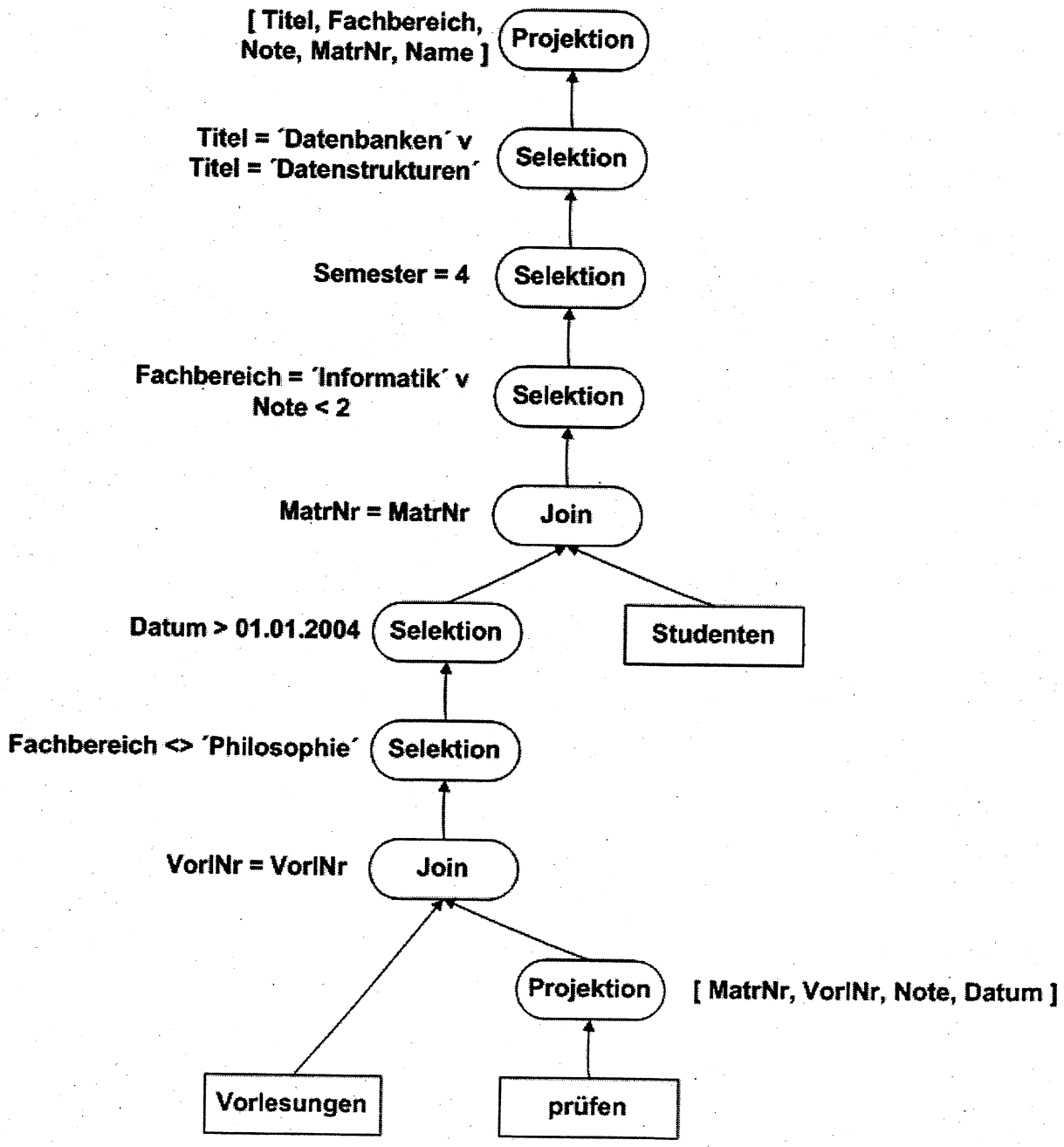
Ware (Waren_Bez, Lager_Bez)
Lager (Lager_Bez, Lager_Leiter)

Bestellung (Lieferanten_Bez, L_Anschrift, Waren_Bez, Menge, Best_Datum) -->

Lieferant (Lieferanten_Bez, L_Anschrift)
Bestellung (Lieferanten_Bez, Waren_Bez, Menge, Best_Datum)

Aufgabe 6:Lösung

a) Operatorbaum:

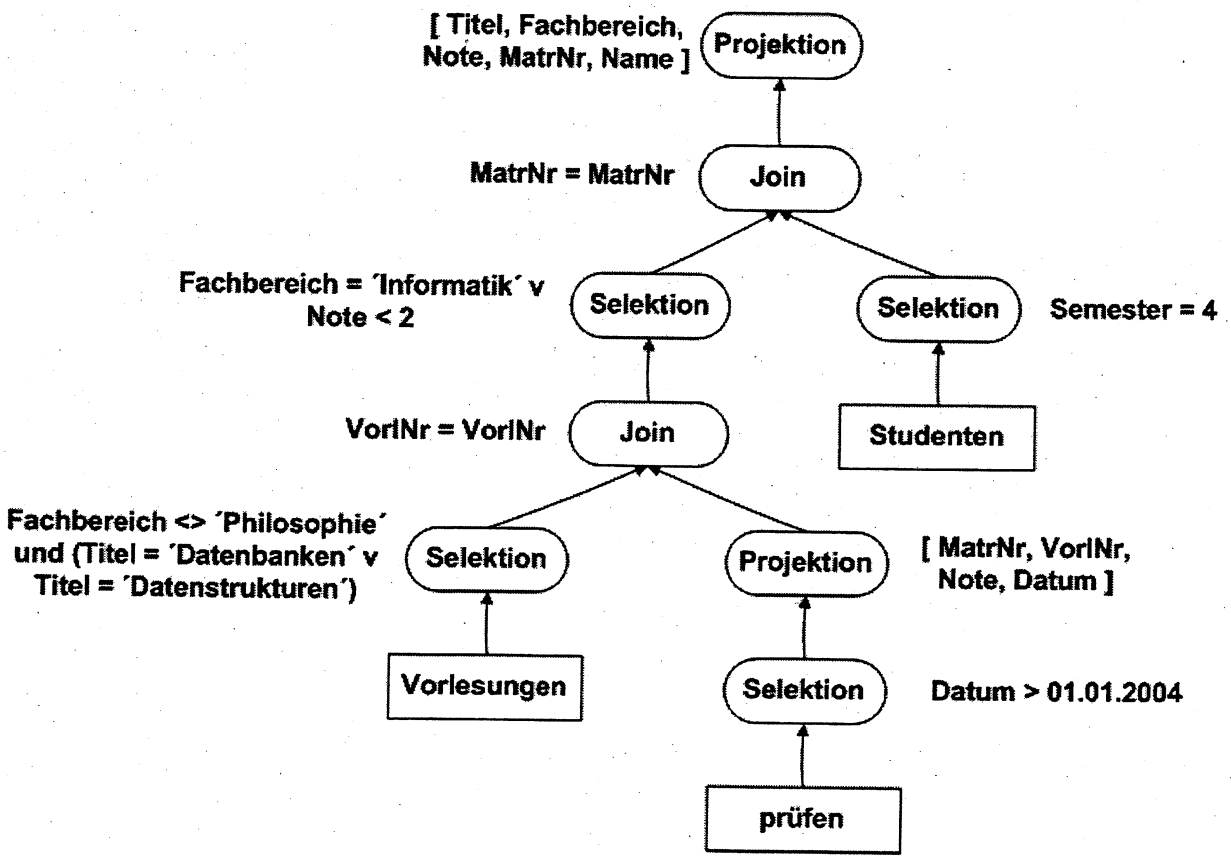


b) Die Selektion **[Datum > 01.01.2004]** bezeiht sich nur auf die Relation **prüfen** und kann daher vor dem Verbund **[VorlNr = VorlNr]** ausgeführt werden.
[Titel = 'Datenbanken' v Titel = 'Datenstrukturen'] und **[Fachbereich <> 'Philosophie']** sind Selektionen, die sich beide auf die Relation **Vorlesungen** beziehen. Deshalb ist es möglich, diese zu einer Selektion zusammenzufassen und sie vor dem Verbund **[VorlNr = VorlNr]** auszuführen.
Die Selektion **[Semester = 4]** bezieht sich nur auf die Relation **Studenten** und kann daher vor dem Verbund **[MatrNr = MatrNr]** ausgeführt werden.
Die Selektion **[Fachbereich = 'Informatik' v Note < 2]** kann, da es sich um eine ODER-Verknüpfung handelt, nicht aufgespalten werden. Sie bezieht sich jedoch nur auf die beiden Relationen **Vorlesungen** und **prüfen** und kann daher auch vor dem Verbund
[MatrNr = MatrNr] ausgeführt werden.

Als Resultat der Optimierung ergibt sich folgender Ausdruck:

Vorlesung [Fachbereich <> 'Philosophie' und (Titel = 'Datenbanken' v
Titel = 'Datenstrukturen')] [VorlNr = VorlNr] (prüfen [Datum > 01.01.2004]
[MatrNr, VorlNr, Note, Datum]) [Fachbereich = 'Informatik' v Note < 2]
[MatrNr = MatrNr] (Studenten [Semester = 4])
[Titel, Fachbereich, Note, MatrNr, Name]

Operatorbaum:



c) Die notwendigen Verbund-Operationen (z.B. **R [A = B] S**) lassen sich auf physischer Ebene noch optimieren durch:

- Ausnutzung von Sekundärindizes

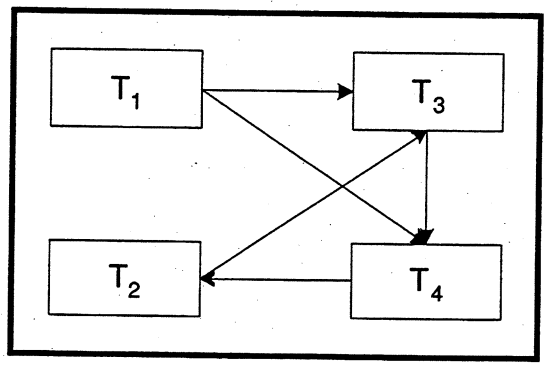
Existiert ein Index für **B**, dann ist für jedes Tupel in **R** sofort die zugehörige Tupelmenge in **S** feststellbar – und damit die Tupelmenge des Verbundes konstruierbar. Die Zahl der notwendigen Zugriffe auf Tupel wird damit linear zur Größe der beiden Relationen.

- *Sortieren von **R** nach **A** und von **S** nach **B***

Bei effizienten Sortierverfahren werden O(n log n) Schritte benötigt. Das anschließende Erstellen des Verbundes ist dann sehr schnell: **R** und **S** werden parallel durchlaufen, wobei jeweils auf gleich Werte für **A** und **B** zu prüfen ist.

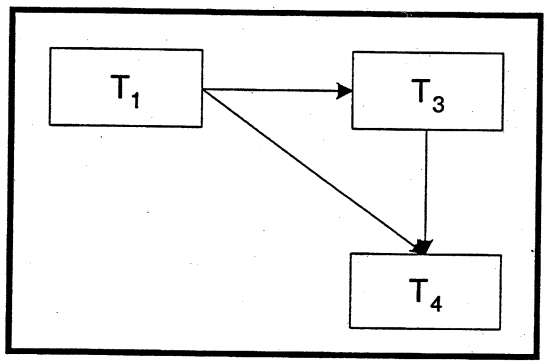
Aufgabe 7 : Lösung

a)



b) Der Abhängigkeitsgraph ist zyklisch (T2 → T3 → T4 → T2).

c) Die Entfernung von T2 ist ausreichend:



d) T_1, T_3, T_4

e) Ein Transaktionssystem kann serialisierbar sein, obwohl der zugehörige Abhängigkeitsgraph nicht zyklenfrei ist. Dabei wird eine lokale Inkonsistenz durch eine spätere Transaktion korrigiert (Wirkungslosigkeit durch Nur-schreibende Transaktionen). Beispiel:

R1(a), R2(a), W1(a), W2(a), W3(a)
T1 vor T2 da R1(a) und W1(a) vor W2(a)
T2 vor T1 da R2(a) vor W1(a)
Also Zyklus
Serielle Ausführung: T1, T2, T3 oder T2, T1, T3, da in jedem Fall das Ergebnis von W3(a) überlebt und dies nicht von einem vorherigen Wert von a abhängig sein kann (da ja nicht gelesen wurde)

Aufgabe 8 : Lösung

(a)

⇒ T1 muss nicht wiederholt werden, da sie bereits vor dem Checkpoint beendet wurde. (1 Punkt)

⇒ Für T2 müssen die After-Images aus dem Logfile in die Datenbank geschrieben werden. Diese brauchen nur bis zum letzten Checkpoint zurückreichen, da alle davorliegenden "After-Images" bereits in die Datenbank eingetragen worden sind. (1 Punkt)

⇒ Für T3 muss ein Rollback ausgeführt werden. (1 Punkt)

⇒ Für T4 müssen die After-Images aus dem Logfile in die Datenbank geschrieben werden. (1 Punkt)

⇒ Für T5 muss ein Rollback ausgeführt werden. (1 Punkt)

(b)

⇒ T1 muss nicht wiederholt werden. (1 Punkt)

⇒ T2 muss nicht wiederholt werden, da das Herausschreiben der veränderten Daten bei EOT geschah. (1 Punkt)

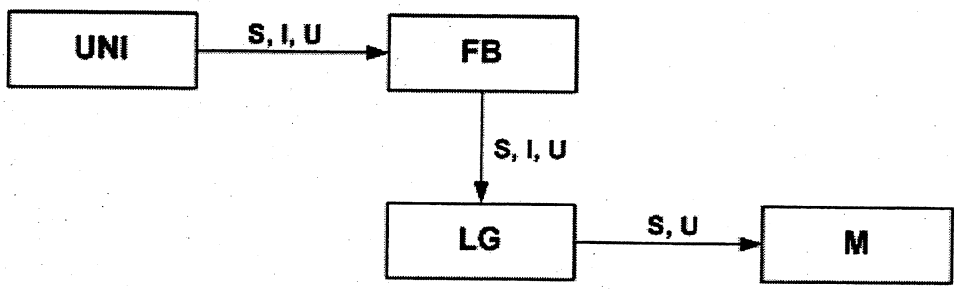
⇒ Für T3 muss ein Rollback ausgeführt werden. (1 Punkt)

⇒ T4 muss nicht wiederholt werden, da das Herausschreiben der veränderten Daten bei EOT geschah. (1 Punkt)

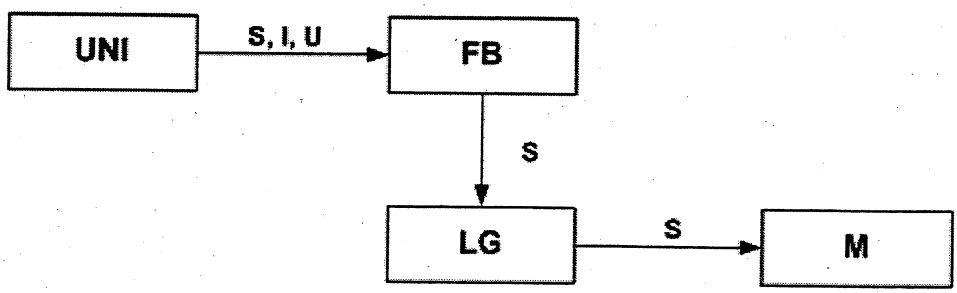
⇒ Für T5 muss ein Rollback ausgeführt werden. (1 Punkt)

Aufgabe 9 : Lösung

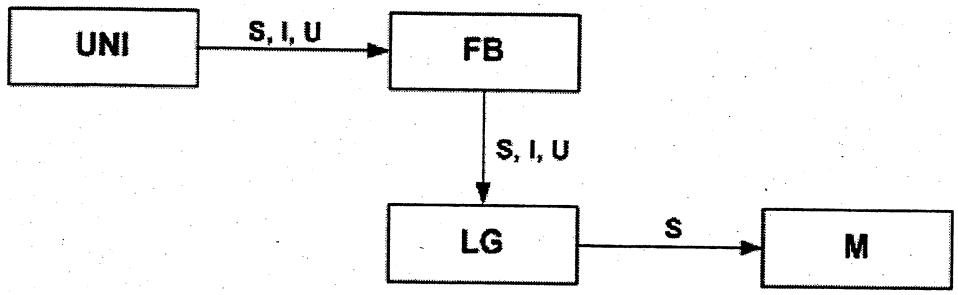
a) Abhängigkeitsgraph zur Relation VORLESUNGEN: (3 Punkte)



Abhängigkeitsgraph zur Relation PROFESSOREN: (3 Punkte)



Abhängigkeitsgraph zur Relation LESEN: (3 Punkte)



b) (je 0,5 Punkte)

FB: GRANT SELECT, INSERT, UPDATE ON VORLESUNGEN TO LG WITH GRANT OPTION

FB: GRANT SELECT ON PROFESSOREN TO LG WITH GRANT OPTION

FB: GRANT SELECT, INSERT, UPDATE ON LESEN TO LG WITH GRANT OPTION

LG: GRANT SELECT, UPDATE ON VORLESUNGEN TO M WITH GRANT OPTION

LG: GRANT SELECT ON PROFESSOREN TO M WITH GRANT OPTION

LG: GRANT SELECT ON LESEN TO M WITH GRANT OPTION

c) (je 1 Punkt)

FB: REVOKE INSERT, UPDATE ON VORLESUNGEN FROM LG

FB REVOKE INSERT, UPDATE ON LESEN FROM LG

Damit müssen die Änderungsrechte der Mitarbeiter an den Lehrgebieten zurückgenommen werde. Also:

LG: REVOKE UPDATE ON VORLESUNGEN FROM M