

Aufgabe 2: SQL-Anfragen

(4 Punkte)

Gegeben ist folgendes Relationenschema zur Verwaltung von Mietverträgen:

PERSON (PersID, Name, Vorname, GebDatum)

WOHNUNG (WohnID, Strasse, Ort, Stock, AnzZimmer)

VERTRAG (VertragsNr, Mieter, Vermieter, Mieta, Dauer, Beginn, Wohnung)

Die Primärschlüssel der Relationen sind wie üblich durch Unterstreichen gekennzeichnet. Mieter und Vermieter in VERTRAG sind Fremdschlüssel zu PersID in PERSON. Wohnung in VERTRAG ist Fremdschlüssel zu WohnID in WOHNUNG.

Formulieren Sie die folgenden Abfragen in SQL:

- Welche Mieter und Vermieter (die IDs der Mieter und Vermieter sind ausreichend) haben 10 Jahres-Verträge abgeschlossen?
- Wie lauten Vor- und Nachnamen aller Vermieter?
- In welchen Orten gibt es Wohnungen mit mehr als 4 Zimmern, deren Mieta weniger als 1200 DM beträgt?
- In welchem Stock wohnt der Mieter Hans Müller?

Aufgabe 3: Ersetzungsstrategien

(4 Punkte)

„In Systempuffer kann in seiner Pufferverwaltung logisch mehr Speicher verwalten, als physischer Hauptspeicher im System vorhanden ist. Aus diesem Grund muß die Pufferverwaltung mit Hilfe von Ersetzungsstrategien nicht mehr benötigte Systempufferseiten durch neu angeforderte Seiten ersetzen.“

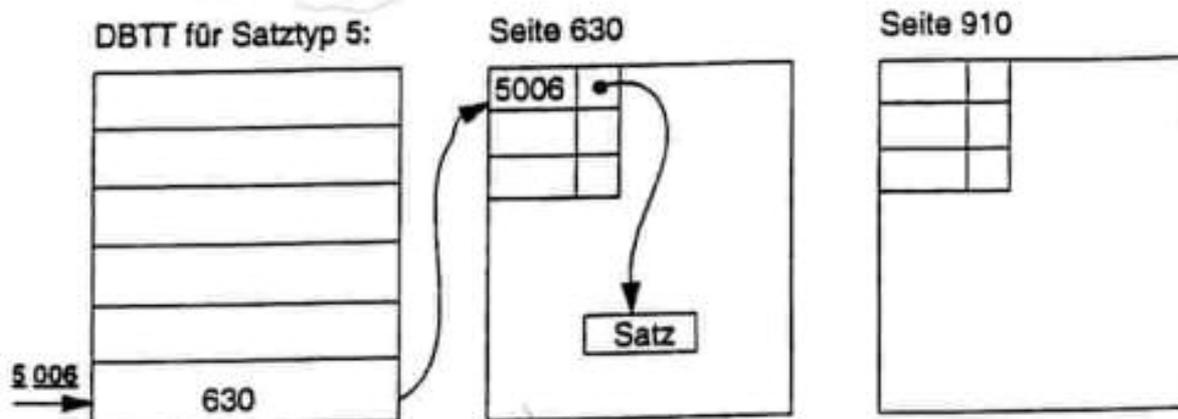
- Nennen Sie drei Ersetzungsstrategien.
- In welche zwei Gruppen lassen sich die realisierbaren Ersetzungsstrategien klassifizieren. Erklären Sie diese kurz. *S. 75!*
- Belady-Optimal* ist eine berühmte, nicht realisierbare Einbringstrategie. Skizzieren Sie kurz deren Vorgehensweise und erläutern Sie, wofür diese Strategie verwendet wird.
- Auch die in Betriebssystemen eingesetzten Puffermechanismen verwenden Einbringstrategien. Welche für die Datenkonsistenz wichtige Operation kommt bei Datenbanksystemen hinzu?

Aufgabe 4: Satzverwaltung

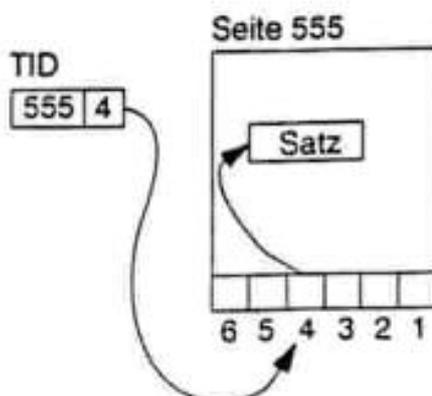
(8 Punkte)

Die Aufgabe der Satzverwaltung ist die physische Abspeicherung von (Daten-)Sätzen in Seiten, und die Unterstützung des Suchvorgangs beim Wiedereinlesen der Sätze. Jedem Datensatz, der gespeichert werden soll, wird von der Satzverwaltung eine eindeutige Satzadresse zugewiesen.

- a) Charakterisieren Sie die wesentlichen Merkmale der "Vollständigen Indirektion" bei der Satzadressierung.
- b) Beim Konzept der Data Base Key Translation Table (DBTT), stellt die DBTT eine Realisierungsform der vollständigen Indirektion dar. Zeichnen Sie analog zum Satz 5 006 die Adressierung der Sätze 5 001 (in Seite 630) und 5 004 (in Seite 910) ein, wobei 5 der Recordtyp und 001 und 004 Folgenummern sind.



- c) Das TID-Konzept stellt eine weitere Realisierungsform der vollständigen Indirektion dar. Gegeben sei die Situation in der untenstehenden Abbildung. Der Datensatz, auf den das Tupel mit der Identifikation 555 | 4 verweist, muß vergrößert werden. Da die Seite jedoch bereits voll ist, findet der vergrößerte Datensatz keinen Platz mehr. Skizzieren Sie anhand der Abbildung, was zu tun ist (Stichwort "Überlaufsatz").



- d) Was sind die wesentlichen Unterschiede zwischen dem TID- und dem DBTT-Konzept?

Aufgabe 5: Baumstrukturierter Schlüsselvergleich (6 Punkte)

- a) Durch welche Merkmale unterscheiden sich B-Bäume von B*-Bäumen?
- b) Die Zahlen von 1 bis 7 sind in einen zunächst leeren B*-Baum aufsteigend einzufügen. Die maximale Anzahl der Einträge in Blattknoten ist $2k$ ($k=1$), die maximale Anzahl der Einträge in inneren Knoten ist $2k^*$ ($k^*=1$).
Zeichnen Sie die Zustände des B*-Baums nach jedem Einfügen. Äste, die sich nach dem Einfügen nicht ändern, müssen nicht jedesmal vollständig gezeichnet werden.

Aufgabe 6: Protokollierung (4 Punkte)

Um die ACID-Eigenschaften von Transaktionen sicherzustellen, muß das Datenbanksystem Protokollinformationen sammeln, die zur Wiederherstellung im Fehlerfall verwendet werden können. Die dabei verwendeten Protokollierungsverfahren werden in *physische* und *logische Protokollierungsverfahren* eingeteilt.

- a) Charakterisieren Sie beide Verfahren kurz, und nennen Sie auch jeweils ein Beispiel dafür, wie ein Protokollsatz aussehen könnte.
- b) Welche Vor- und Nachteile haben die beiden Verfahren jeweils?

Aufgabe 7: STEAL/NO STEAL-Strategie (4 Punkte)

- a) Erläutern Sie die STEAL-Strategie und den Unterschied zur NO-STEAL-Strategie.
- b) Welche der vier Wiederherstellungsmaßnahmen werden bei einer NO-STEAL-Strategie überflüssig?
- c) Welchen Nachteil hat die NO-STEAL-Strategie jedoch?

Aufgabe 8: Sicherung und Wiederherstellung (4 Punkte)

Sicherungs- und Wiederherstellungsmaßnahmen in einem Datenbanksystem dienen der Realisierung von Transaktionen mit deren ACID-Eigenschaften. Das zentrale Ziel dabei ist die Erhaltung der physischen und logischen Konsistenz der Daten.

- a) Was ist mit den Begriffen physische und logische Konsistenz gemeint, und wodurch werden sie jeweils sichergestellt?
- b) Nehmen Sie an, eine Transaktion wird beim Schreiben ihres EOT-Satzes (End-of-Transaction) durch einen Systemfehler abgebrochen. Welche Wiederherstellungsmaßnahme muß angewendet werden?

Aufgabe 9: Granulate**(3 Punkte)**

- a) Bei der physischen Protokollierung gibt es einen unbedingt notwendigen Zusammenhang zwischen Sperrgranulat und Protokollgranulat. Welchen?
- b) Erläutern und begründen Sie den Zusammenhang aus Teilaufgabe a) anhand einer Beispielskizze.

Aufgabe 10: Zwei-Phasen-Freigabeprotokoll**(3 Punkte)**

Geben Sie eine kurze Beschreibung der beiden Phasen des Zwei-Phasen-Freigabeprotokolls.

1. Phase:

2. Phase:

Aufgabe 11: Phasen der Anfrageverarbeitung**(4 Punkte)**

Die Anfrageverarbeitung in Datenbanksystemen läßt sich in Phasen unterteilen. Nennen Sie diese Phasen und beschreiben Sie sie kurz.

- Als Schmierpapier ist nur das vom Aufsichtspersonal verteilte erlaubt!
 - In den Antworten ist nur das ausdrücklich Gefragte verlangt - nicht mehr und nicht weniger.
 - Nur solche Schmierpapierblätter, die mit Ihrem Namen versehen und zusammen mit Ihrer Klausur abgegeben wurden, können bei der Bewertung berücksichtigt werden.
 - Die Bearbeitungszeit beträgt 90 Minuten.
-

Aufgabe 1: Datenmodelle**(4 Punkte)**

- a) Nennen Sie die drei Grundelemente des CODASYL-Datenmodells!
- b) Auf welcher Ebene des Schichtenmodells ist der Datenzugriff im CODASYL-Datenmodell anzusiedeln?
- c) Welchen Vorteil bezüglich der Anfrageformulierung hat das relationale Modell gegenüber dem CODASYL-Datenmodell?

Aufgabe 2: SQL-Anfragen**(4 Punkte)**

Gegeben ist folgendes Relationenschema zur Verwaltung von Daten aus der Fußballbundesliga:

VEREIN (Vereinsname, Kapitän)

SPIELER (SpielerID, Name, Vorname, Wohnort, Vereinsname)

MATCH (MatchID, Datum, Heimverein, Gastverein, ToreHeim, ToreGast)

Die Primärschlüssel der Relationen sind wie üblich durch Unterstreichen gekennzeichnet.

Kapitän in VEREIN ist Fremdschlüssel zu SpielerID in SPIELER.

Vereinsname in SPIELER ist Fremdschlüssel zu Vereinsname in VEREIN.

Heimverein in MATCH ist Fremdschlüssel zu Vereinsname in VEREIN.

Gastverein in MATCH ist Fremdschlüssel zu Vereinsname in VEREIN.

- a) Formulieren Sie die folgenden Abfragen in SQL:

- Wo wohnt Stefan Reuter?
- Wie heißt der Kapitän des '1.FC Nürnberg'?

- b) Formulieren Sie folgendes SQL-Statement umgangssprachlich!

```
SELECT S.Name, S.Vorname
FROM VEREIN V, SPIELER S, MATCH M
WHERE M.Datum = '17.7.97'
      AND M.Heimverein = V.Vereinsname
      AND V.Kapitän = S.SpielerID
      AND M.Gastverein = 'FC Bayern München'
```

Aufgabe 3. Paging

(4 Punkte)

Erklären Sie die Phänomene "Page Fault", "Database Fault" und "Double Page Fault" anhand der folgenden Diagramme.

M = Realer Hauptspeicher

N = Systempuffer der Datenbank

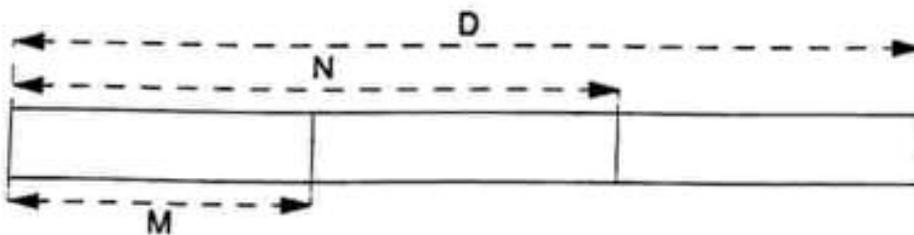
D = Datenbank

Die Reihenfolge der Vorgänge ist durch nummerierte Pfeile darzustellen. Die benötigte Seite und die zu ersetzende Seite sind wie folgt zu markieren:

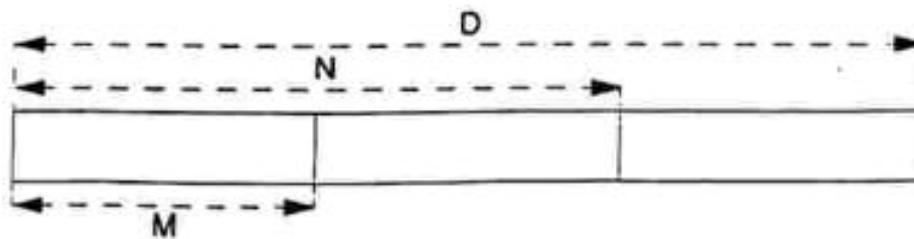
 = benötigte Seite

 = zu ersetzende Seite

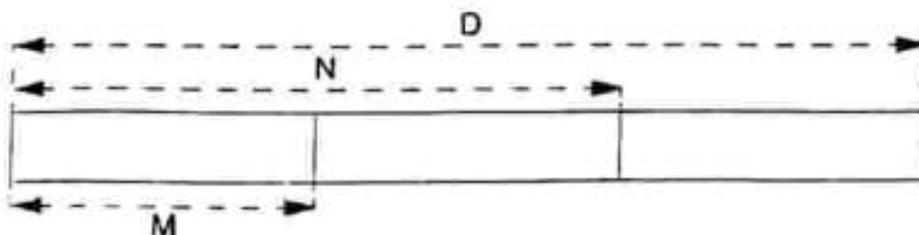
a) Page Fault



b) Database Fault



c) Double Page Fault



Aufgabe 4: DBTT-Konzept

(3 Punkte)

Erläutern Sie kurz unter Zuhilfenahme einer Zeichnung das DBTT-Konzept (Data Base Key Translation Table). Beschriften Sie Ihre Zeichnung, und zeichnen Sie den Weg vom Datenbank-schlüssel bis zum Datensatz ein.

Aufgabe 5: B-Baum

(5 Punkte)

Gegeben sei ein anfangs leerer B-Baum (maximale Anzahl der Einträge pro Knoten = $2k$, $k=1$). Die Zahlen 4, 7, 1, 3, 2, 8 und 9 sind in dieser Reihenfolge in den B-Baum einzufügen. Zeichnen Sie die Zustände des B-Baums nach jedem Einfügevorgang. Falls sich durch das Einfügen die Struktur des Baumes nicht ändert, muß er nicht neu gezeichnet werden, sondern der einzufügende Wert kann einfach in den Baum eingetragen werden.

Aufgabe 6: Transaktionskonzept

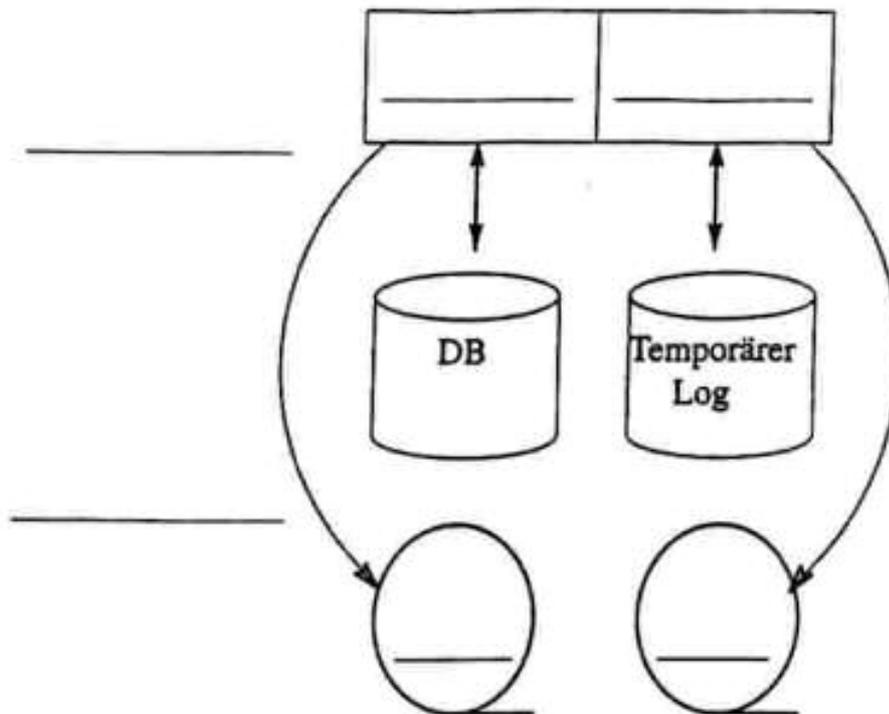
(7 Punkte)

- a) Benennen und erläutern Sie die klassischen Eigenschaften, die von einer Transaktion gefordert werden.
- b) Welche der Eigenschaften ist nur durch den Benutzer bestimmt? Begründen Sie Ihre Entscheidung!

Aufgabe 7: Speicherbereiche für Wiederherstellung

(8 Punkte)

- a) Ergänzen Sie die in der Abbildung gezeigten Speicherbereiche, durch Ausfüllen der Platzhalter.



- b) Auf welche der Bereiche muß in den Recovery-Fällen R1, R2, R3, R4 jeweils zugegriffen werden?

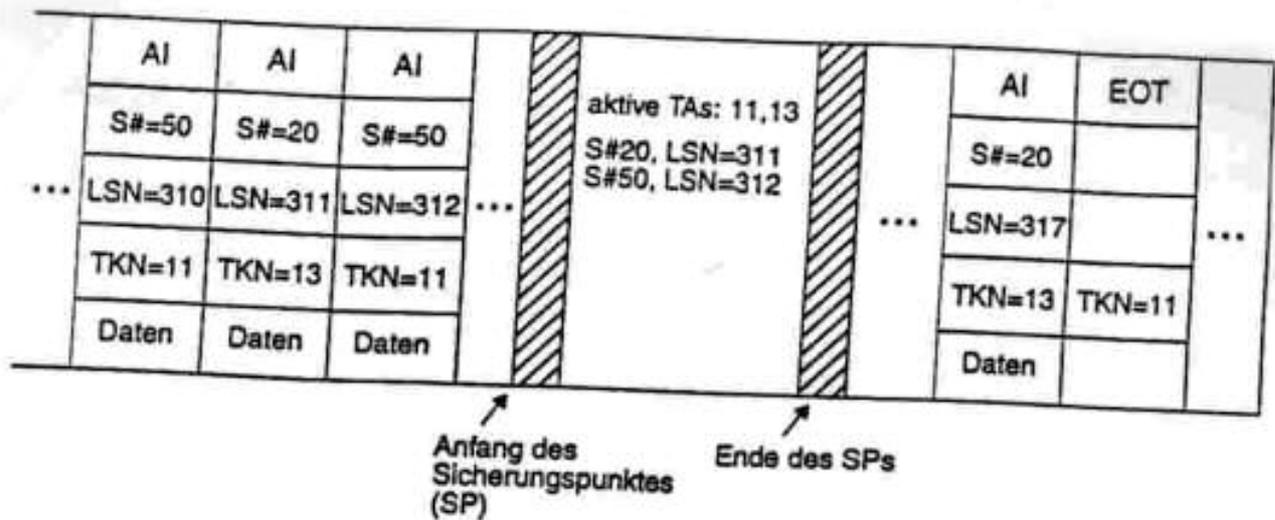
Aufgabe 8: Indirekte Sicherungspunkte

(2 Punkte)

Gegeben sei ein Auszug aus einer temporären Protokolldatei mit einem indirekten Sicherungspunkt (SP).

- AI = After Image
- S# = Seitennummer
- LSN= Log-Sequence-Number
- TKN= Transaktionskennung
- EOT= End of Transaction
- TA = Transaktion

Wie findet man heraus, daß der grau unterlegte Protokollsatz das jüngste After-Image der Seite 50 der Transaktion T11 ist?



Aufgabe 9: Zwei-Phasen Freigabeprotokoll

(3 Punkte)

Geben Sie eine kurze Beschreibung der beiden Phasen des Zwei-Phasen-Freigabeprotokolls.

Aufgabe 10: Fehler im Mehrbenutzerbetrieb

(6 Punkte)

In einem Datenbanksystem mit Mehrbenutzerbetrieb können Fehler auftreten, die es im Einbenutzerbetrieb nicht gibt. Erläutern Sie jeweils anhand eines Beispiels die Fehlerarten, die durch physische Sperrverfahren verhindert werden können. Benutzen Sie für Ihre Beispiele folgende Syntax:

$R_i(x)$: Lesen von Transaktion i auf Datenobjekt x

$W_i(x)$: Schreiben von Transaktion i auf Datenobjekt x

Aufgabe 11: Phasen der Anfrageverarbeitung

(4 Punkte)

Die Anfrageverarbeitung in Datenbanksystemen läßt sich in vier grundsätzliche Phasen unterteilen. Nennen und beschreiben Sie die einzelnen Phasen kurz!

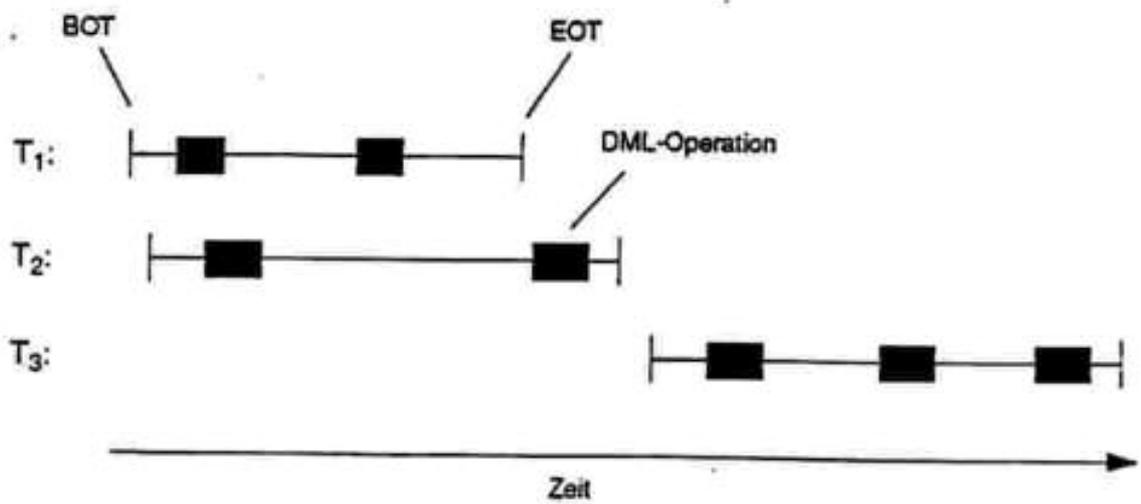
Aufgabe 1: Sicherungspunkte

(5 Punkte)

- a) Erläutern Sie den Unterschied zwischen einem direkten und einem indirekten Sicherungspunkt.

Gegeben sei der unten dargestellte Transaktionsablauf der drei Transaktionen T_1 , T_2 und T_3 .

- b) Zeichnen Sie in das folgende Diagramm jeweils einen *speicherinkonsistenten* (1), einen *speicherkonsistenten* (2) (aber semantisch *inkonsistenten*) und einen *semantisch konsistenten* (3) Sicherungspunkt ein.



Aufgabe 2: Paging

(4 Punkte)

Erklären sie die Phänomene "Page Fault", "Database Fault" und "Double Page Fault" anhand der folgenden Diagramme.

M = Realer Hauptspeicher

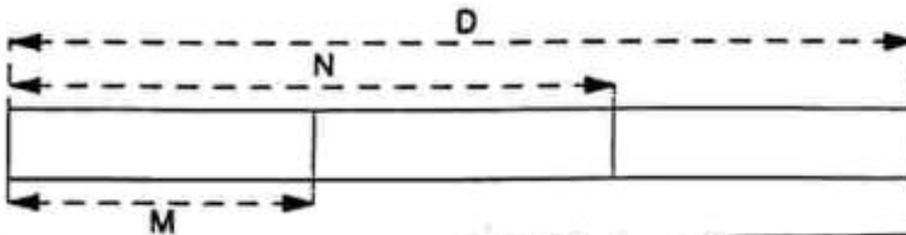
N = Systempuffer der Datenbank

D = Datenbank

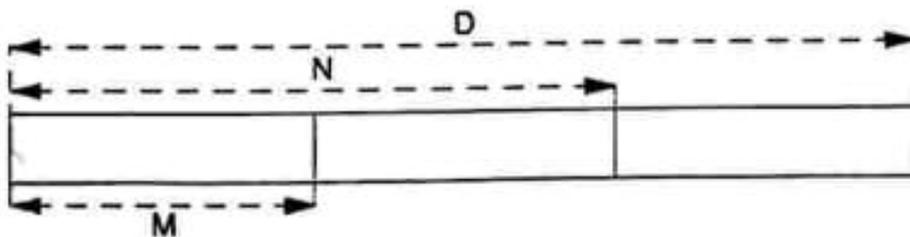
Die benötigte Seite und die zu ersetzende Seite ist wie folgt zu markieren:

 = benötigte Seite  = zu ersetzende Seite

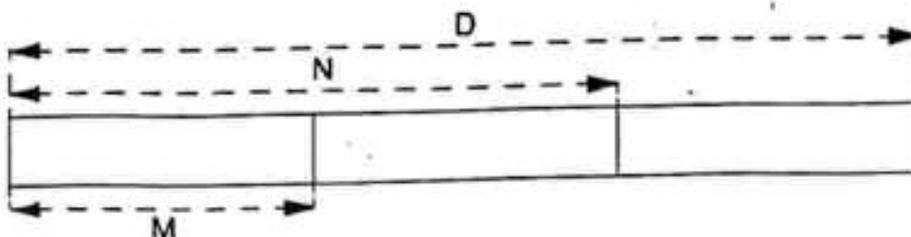
a) Page Fault



b) Database Fault



c) Double Page Fault



Aufgabe 3: TID-Konzept und DBTT-Tabelle

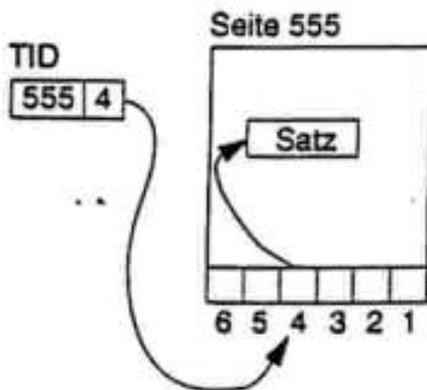
(4 Punkte)

a) TID-Konzept

Gegeben sei die Situation in dem untenstehenden Diagramm. Der Datensatz, zu dem das Tupel mit der Identifikation

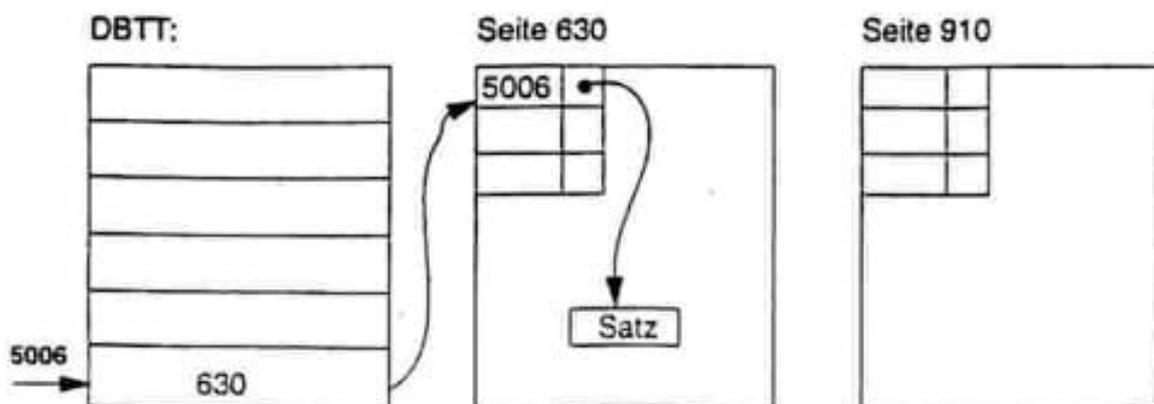
555	4
-----	---

 verweist, muß vergrößert werden. Da die Seite jedoch bereits voll ist, findet der vergrößerte Datensatz keinen Platz mehr. Skizzieren Sie anhand eines Diagramms, was zu tun ist (Stichwort "Überlaufsatz")



b) DBTT-Tabelle

Zeichnen Sie analog zum Satz 5 006 die Adressierung der Sätze 5 001 (in Seite 630) und 5 004 (in Seite 910) ein, wobei 5 der Recordtyp und 001 und 004 Folge-nummern sind.



Aufgabe 4: B-Baum

(5 Punkte)

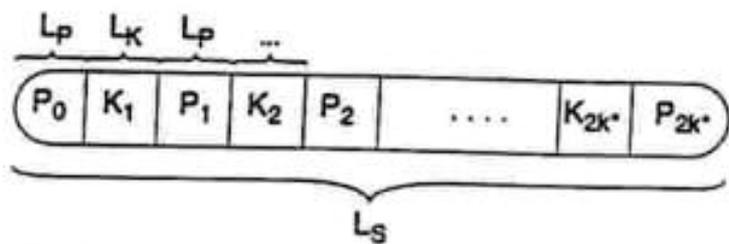
Die Zahlen von 1 bis 10 sind in einen B-Baum (maximale Anzahl der Einträge = $2k$, $k=1$), zu Anfang leer ist, aufsteigend einzufügen. Zeichnen Sie die Zustände des B-Baums nach jedem Einfügen.

Aufgabe 5: B*-Baum

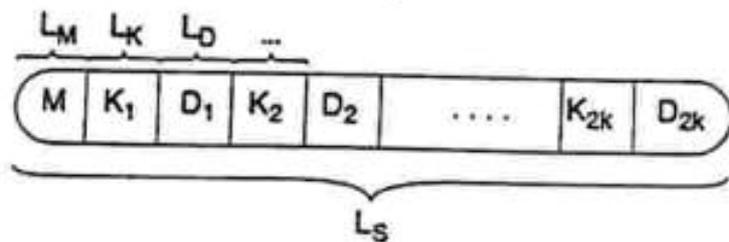
(2 Punkte)

Der B*-Baum hat bekanntlich zwei Knotenformate

innere Knoten:



Blattknoten:



Die Felder haben die folgenden Längen:

L_p (Länge der Pointerfelder) = 4 Byte

L_k (Länge des Schlüsselfeldes) = 4 Byte

L_m (Länge des Markenfeldes) = 4 Byte

L_d (Länge des Datenfeldes) = 80 Byte

Man bestimmt die Parameter k und k^* , wenn die Seitenlänge $L_s = 2$ Kilobyte = 2048 Byte beträgt.

Aufgabe 6: Anfragen und Integritätsbedingungen**(4 Punkte)**

Gegeben sei die folgende Datenbank einer Universität

STUDENT(STNR, NAME, ALTER, FBNR)

4711	Müller	23	1
4712	Schulze	24	11
:	:	:	:

FACHBEREICH(FBNR, BEZEICHNUNG, ORT)

1	Theologie	Erl
11	Technik	Erl
5	Wirtschaft	Nbg
:	:	:

a) Formulieren Sie in SQL die folgende Anfrage:

In welchem Ort befindet sich der Fachbereich des Studenten mit der Studentenummer STNR = 4711?

b) Spezifizieren Sie formal die referentielle Integritätsbedingung zwischen der FBNR in STUDENT und der FBNR in FACHBEREICH.**Aufgabe 7: Ersetzungsstrategie im DB-Puffer****(3 Punkte)**Ein Puffer der Größe $N = 4$ wird mit der Ersetzungsstrategie FIFO bzw. LRU verwaltet. Der logische Referenzstring sieht wie folgt aus:

20, 30, 10, 40, 20, 50, 20, 30, 10, 40
—————→
Zeit

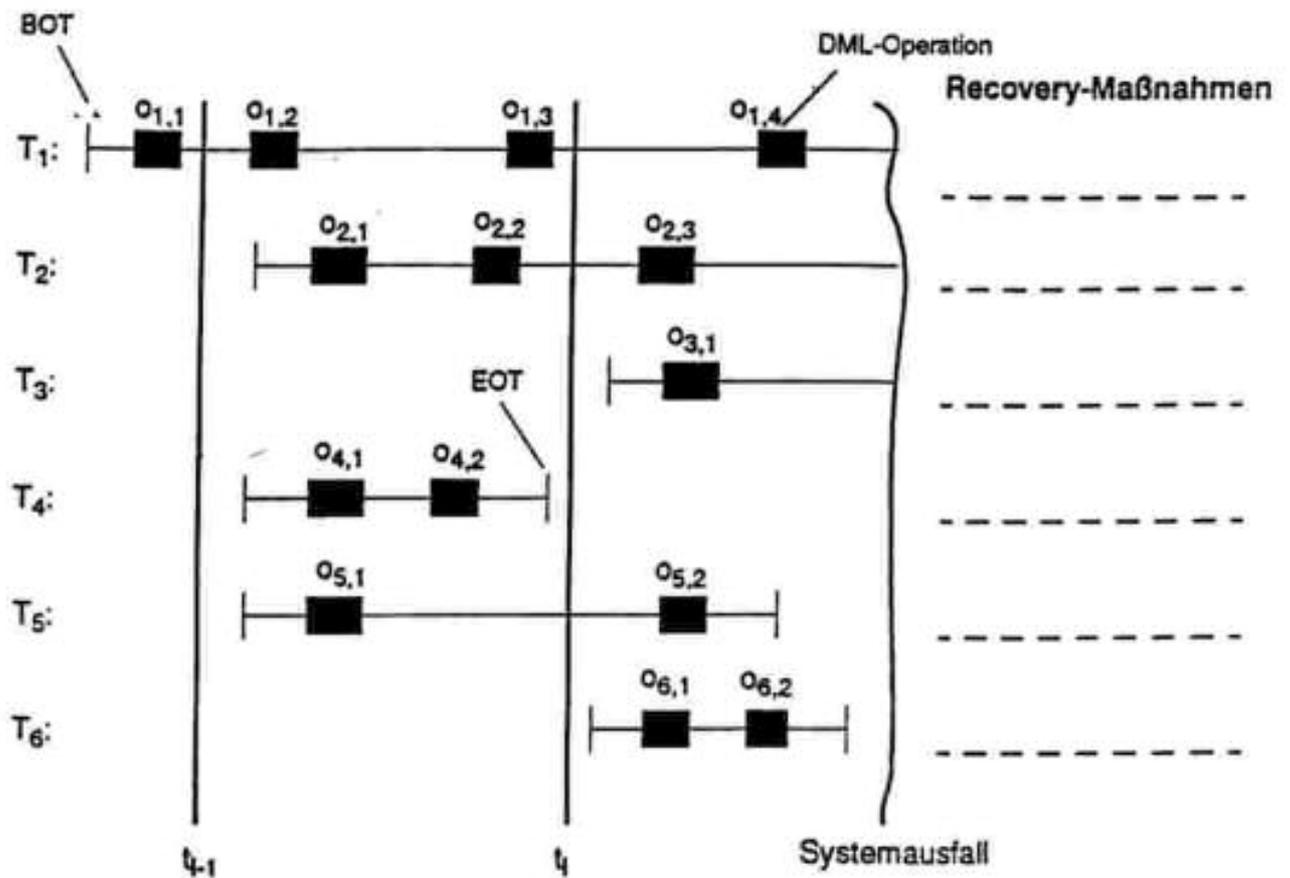
Welcher Pufferinhalt liegt nach Abarbeiten der letzten Referenz vor, und wieviele Puffer-Faults gibt es jeweils?

a) bei FIFO**b)** bei LRU

Aufgabe 9: Recoverymaßnahmen

(6 Punkte)

Gegeben sei der folgende zeitliche Verlauf von sechs Transaktionen. Zu den Zeitpunkten t_{-1} und t_1 wurden speicherkonsistente direkte Sicherungspunkte genommen. Welche Recoverymaßnahmen (R1 bis R4) sind nach einem Systemausfall jeweils erforderlich? Welche DML-Operationen sind betroffen? Dabei sei angenommen, daß eine unterbrechbare Einbringstrategie mit Update-In-Place verwendet wird.



Aufgabe 10: Zwei-Phasen Freigabeprotokoll**(4 Punkte)**

Geben Sie eine kurze Beschreibung der beiden Phasen des Zwei-Phasen-Freigabeprotokolls.

1. Phase:

2. Phase:

Aufgabe 11: Serialisierungsgraph**(3 Punkte)**Transaktion T_1 konfigiert mit Transaktion T_2 **UND** Transaktion T_2 konfigiert mit Transaktion T_3 .

a) Zeichnen Sie hierzu einen möglichen Serialisierungsgraphen.

b) Warum gilt nicht allgemein die "Transitivität"

WENN T_1 mit T_2 konfigiert**UND** T_2 mit T_3 konfigiert**DANN** konfigiert T_1 mit T_3 ?**Aufgabe 12: Kompatibilität hierarchischer Sperren****(3 Punkte)**

Bestimmen Sie, ob die Sperren A und B in der untenstehenden Tabelle kompatibel sind.

A	B	kompaübel (ja/nein)
SIX	IX	
IX	IS	
IX	IX	
SIX	S	
IX	S	
IS	S	

Aufgabe 13: Hierarchische Sperrketten

(4 Punkte)

Gegeben seien die hierarchischen Sperrgranulate "Datenbank (DB)", "Segment (S)", "Relation (R)" und "Tupel (T)".

Welche hierarchischen Sperren sind zu setzen, um bei größtmöglicher Nebenläufigkeit

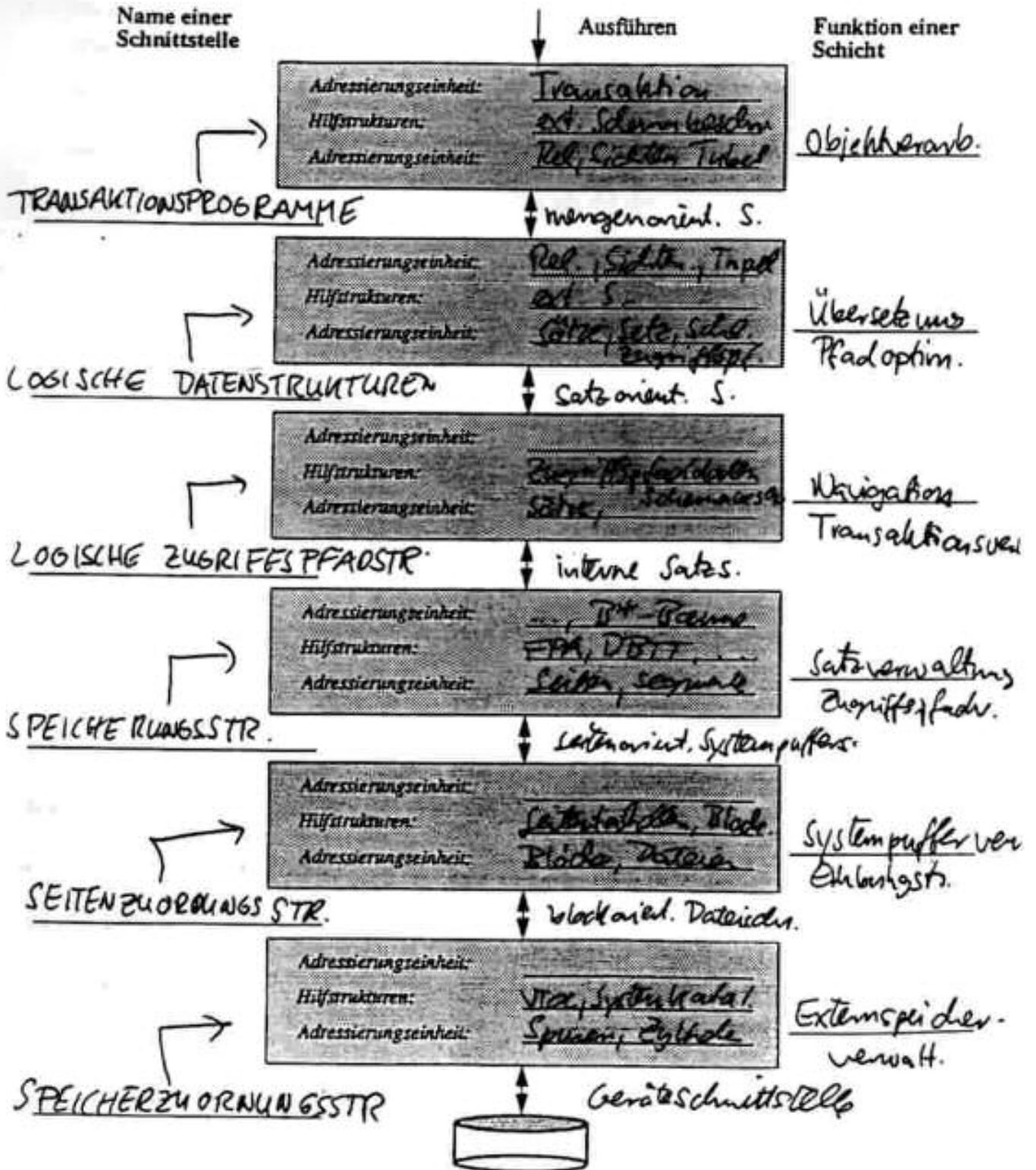
a) ein Tupel t zu lesen?

b) ein Tupel t zu modifizieren?

Aufgabe 1: Schichtenmodell

(12 Punkte)

Skizzieren Sie die wichtigsten Elemente des Schichtenmodells, indem Sie alle Platzhalter korrekt ausfüllen (Bei den Hilfsstrukturen und Adressierungseinheiten genügt jeweils eine Nennung pro Platzhalter).



Aufgabe 2: SQL-Anfragen

(3 Punkte)

Für die Verwaltung ihrer Bestände in einer Datenbank verwendet eine kleine Buchhandlung die folgenden Relationen:

BUCH (ISBN, Autor_Vorname, Autor_Nachname, Titel, Verlags_Id, Lagerbestand)

VERLAG (Verlags_Id, Verlagsname, Verlagsadresse)

Die Primärschlüssel der Relationen sind unterstrichen. Verlags_Id in BUCH ist Fremdschlüssel zu Verlags_Id in VERLAG.

Formulieren Sie die folgenden Datenbankoperationen in SQL.

- Welche ISBN hat das Buch "Kaufmännische Datenbanken" von Hartmut Wedekind?
- Welche Adresse hat der Verlag, in dem das Buch mit der ISBN 3-411-16541-3 erschienen ist?

Aufgabe 3: Einbringstrategien

(4 Punkte)

- Was versteht man unter einer *direkten Einbringstrategie*?
- Was versteht man unter einer *indirekten Einbringstrategie*?
- Warum werden indirekte Einbringstrategien verwendet?
- Nennen Sie zwei Beispiele für indirekte Einbringstrategien.

Aufgabe 4: Baumstrukturierter Schlüsselvergleich

(5 Punkte)

- Durch welche Merkmale unterscheiden sich B-Bäume von B*-Bäumen?
- Die Zahlen von 1 bis 7 sind in einen zunächst leeren B*-Baum, aufsteigend einzufügen. Die maximale Anzahl der Einträge in Blattknoten ist $2k$ ($k=1$), die maximale Anzahl der Einträge in inneren Knoten ist $2k^*$ ($k^*=1$).
Zeichnen Sie die Zustände des B*-Baums nach jedem Einfügen. Äste, die sich nach dem Einfügen nicht ändern, müssen nicht jedesmal vollständig gezeichnet werden.

Aufgabe 5: Recovery

(6 Punkte)

Nennen Sie die vier klassischen Recovery-Maßnahmen. Für welche Art von Fehlern werden diese jeweils benötigt und welche Aktionen müssen jeweils durchgeführt werden?

Aufgabe 6: Logging

(2 Punkte)

- Definieren Sie die Begriffe "Before Image" und "After Image".
- Wie kann verhindert werden, daß während der Ausführung einer Transaktion mehrere Before Images derselben Seite angelegt werden?

Aufgabe 7: FORCE-Strategie

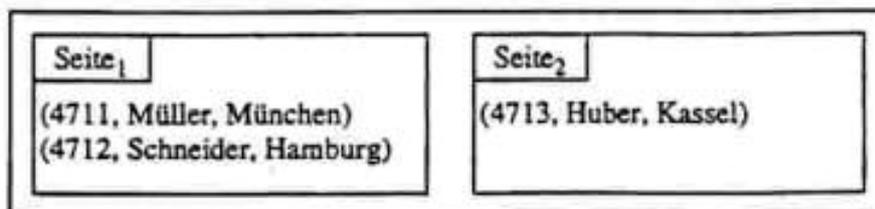
(3 Punkte)

- Erläutern Sie die FORCE-Strategie.
- Welcher Recovery-Operator wird bei einer FORCE-Strategie überflüssig?
- Welchen Nachteil hat die FORCE-Strategie?

Aufgabe 8: Protokollierung

(6 Punkte)

Gegeben sei der in der Abbildung gezeigte Datenbankzustand.



Es werden die folgenden Operationen innerhalb derselben Transaktion (TID=ta) ausgeführt:

```
BOT ta
INSERT Satz (4714, Weber, Nürnberg)
READ Satz 4712
MODIFY Satz 4712 (Schneider, Bremen)
DELETE Satz 4711
EOT ta
```

- Definieren Sie die Begriffe "physische Protokollierung" und "logische Protokollierung".
- Geben Sie die temporäre Protokolldatei für das gegebene Szenario bei logischer Protokollierung an.

Aufgabe 9: Granulate

(2 Punkte)

- a) Bei der physischen Protokollierung gibt es einen unbedingt notwendigen Zusammenhang zwischen Sperrgranulat und Protokollgranulat. Welchen?
- b) Erläutern und begründen Sie den Zusammenhang aus Teilaufgabe a) anhand einer Beispielskizze.

Aufgabe 10: Zwei-Phasen-Sperrprotokoll

(3 Punkte)

- a) Geben Sie eine kurze Beschreibung der beiden Phasen des Zwei-Phasen-Sperrprotokolls.
1. Phase:
2. Phase:
- b) Was ist ein *strikt*es Zwei-Phasen-Sperrprotokoll?

Aufgabe 11: Phasen der Anfrageverarbeitung

(4 Punkte)

Die Anfrageverarbeitung in Datenbanksystemen läßt sich in Phasen unterteilen. Nennen Sie diese Phasen und beschreiben Sie sie kurz.