
Prüfungsteilnehmer

Prüfungstermin

Einzelprüfungsnummer

Kennzahl: _____

Kennwort: _____

Arbeitsplatz-Nr.: _____

**Frühjahr
2012**

66114

**Erste Staatsprüfung für ein Lehramt an öffentlichen Schulen
— Prüfungsaufgaben —**

Fach: Informatik (vertieft studiert)

Einzelprüfung: Datenbank- und Betriebssysteme

Anzahl der gestellten Themen (Aufgaben): 4 Aufgaben, von denen zwei gemäß untenstehender Auswahlregel zu bearbeiten sind!

Anzahl der Druckseiten dieser Vorlage: 14

Zu den zwei Themenschwerpunkten A (Datenbanksysteme) und B(Betriebssysteme) ist jeweils entweder die Aufgabe 1 oder 2 zu wählen.

Auf der Vorderseite des Kopfbogens sind im Feld „gewähltes Thema: Nr.“ die Nummern der beiden gewählten Aufgaben anzugeben (z. B. A2, B1)!

Bitte wenden!

Themenschwerpunkt A (Datenbanksysteme)

Teilaufgabe 1:

1. Modellierung

Sie sollen ein System zur Verwaltung von Pferderennen entwerfen.
Gehen Sie dabei von folgendem Szenario aus:

- Unternehmen werden über ihre eindeutige Unternehmens-Id identifiziert. Sie haben eine Adresse und besitzen Rennställe.
- Der Name eines Rennstalls ist nur innerhalb eines Unternehmens eindeutig. Für jeden Rennstall wird das Gründungsdatum gespeichert.
- Pferde gehören immer zu einem Rennstall. Pferdenamen werden in einem Rennstall nur jeweils maximal einmal vergeben.
- Jockeys sind in einem Rennstall beschäftigt. Jeder Rennstall vergibt seine eigenen Personalnummern. Für jeden Jockey werden sein Vorname und Name gespeichert.
- Rennen haben ein Datum, ein Preisgeld und einen Namen, über den sie identifiziert werden.
- Unternehmen unterstützen Rennen finanziell mit einem bestimmten Betrag.
- Jockeys nehmen mit Pferden an Rennen teil. Im Rennen erreichen sie einen bestimmten Platz. Die Kombination aus Jockey und Pferd ist nicht fest, bei unterschiedlichen Rennen können Jockeys verschiedene Pferde reiten. Jockeys können auch mit Rennpferden von fremden Rennställen, die anderen Unternehmen gehören können, an Rennen teilnehmen.

- a) Entwerfen Sie für das beschriebene Szenario ein ER-Modell.
Bestimmen Sie hierzu:

- die Entity-Typen, die Relationship-Typen und jeweils deren Attribute,
- ein passendes ER-Diagramm,
- die Primärschlüssel der Entity-Typen, welche Sie anschließend in das ER-Diagramm eintragen, und
- die Funktionalitäten der Relationship-Typen, welche Sie ebenfalls in das ER-Diagramm eintragen.

- b) Überführen Sie das ER-Modell aus Aufgabe a) in ein verfeinertes relationales Modell. Geben Sie hierfür die verallgemeinerten Relationenschemata an. Achten Sie dabei insbesondere darauf, dass die Relationenschemata keine redundanten Attribute enthalten.

Fortsetzung nächste Seite!

2. Anfragen

Zu einer Website, auf der Besucher im Browser Online-Spiele spielen können, liegt das folgende relationale Schema einer Datenbank vor:

Team : {[TNr, Name, Teamfarbe]}
Spieler : {[SNr, Name, Icon, TNr, EMail]}
Minispiel : {[MNr, Name, Kategorie, Schwierigkeit]}
Wettkampf : {[WNr, Sieger, Geschlagener, MNr, Dauer]}

Auf der Website treten jeweils zwei Spieler gegeneinander in Minispielen an. In diesen ist es das Ziel, den gegnerischen Spieler in möglichst kurzer Zeit zu besiegen. Minispiele gibt es dabei in verschiedenen Schwierigkeitsstufen ('leicht', 'mittel', 'schwer', 'sehr schwer') und verschiedenen Kategorien ('Denkspiel', 'Geschicklichkeitsspiel', usw.). Die Attribute *Sieger* und *Geschlagener* sind jeweils Fremdschlüsselattribute, die auf das Attribut *SNr* der Relation *Spieler* verweisen. Beachten Sie, dass das *Dauer*-Attribut der Wettkampf-Relation die Dauer eines Wettkampfes in der Einheit Sekunden speichert.

- a) Formulieren Sie geeignete Anfragen in relationaler Algebra für die folgenden Teilaufgaben:
- Geben Sie die Namen der Minispiele zurück, die zur Kategorie 'Denkspiele' zählen.
 - Finden Sie die Spieler, die mindestens einmal in weniger als 5 Minuten besiegt wurden.
 - Geben Sie die Namen und E-Mail-Adressen aller Spieler zurück, die in einem Minispiel des Typs 'Geschicklichkeitsspiel' gewonnen haben.
- b) Formulieren Sie geeignete SQL-Anfragen für die folgenden Teilaufgaben. Beachten Sie dabei, dass Ihre Ergebnisrelationen keine Duplikate enthalten sollen.
- Geben Sie jede Spielekategorie aus, für die ein Minispiel der Schwierigkeitsstufe 'sehr schwer' vorhanden ist.
 - Geben Sie zu jedem Team seine *TNr* und die Anzahl der Mitglieder aus.
 - Geben Sie die Wettkämpfe aus, deren Dauer unter der durchschnittlichen Dauer der Wettkämpfe liegt.
 - Geben Sie für jeden Spieler seine *SNr*, seinen Namen, die Anzahl seiner Siege, die durchschnittliche Dauer seiner siegreichen Wettkämpfe und die Anzahl der Teams, aus denen er bereits mindestens einen Spieler besiegt hat, zurück.
 - Finden Sie die Teams, die bereits mindestens 10 Wettkämpfe gewonnen haben. Geben Sie zu diesen Teams die Teamnamen und die Anzahl der siegreichen Wettkämpfe zurück. Sortieren Sie die Ausgabe abfallend nach der Anzahl der siegreichen Wettkämpfe.
 - Aktualisieren Sie den Wettkampf mit der *WNr* 42. Setzen Sie den Sieger auf den Spieler mit der *SNr* 5037 und den Verlierer auf den Spieler mit der *SNr* 13. Erhöhen Sie zusätzlich die Dauer dieses Wettkampfs um 2 Minuten.

Fortsetzung nächste Seite!

3. Entwurfstheorie

Gegeben ist der folgende Ausschnitt einer relationalen Datenbank:

Vereine							
UserId	ClubId	Vorname	Name	Login	ClubName	Funktion	Vorstand
1	1	Max	Muster	maxm	IT Kicker	Trainer	false
1	2	Max	Muster	maxm	CC Nerds	2. Vorsitzende(r)	true
2	2	Marta	Maier	maier1	CC Nerds	1. Vorsitzende(r)	true
3	1	Tim	Peters	timpeters	IT Kicker	1. Vorsitzende(r)	true
4	1	Peter	Huber	timpeters	IT Kicker	Spieler(in)	false
5	3	Tina	Müller	tinam	IT Club	Kassenprüfer(in)	true
...

In der Datenbank sollen folgende funktionale Abhängigkeiten gelten:

$UserId \rightarrow Vorname, Name, Login$

$Login \rightarrow Vorname, Name, UserId$

$UserId, ClubId, Name \rightarrow Funktion, ClubName$

$Login, ClubId, Name \rightarrow Funktion, ClubName$

$ClubId \rightarrow ClubName$

$Funktion \rightarrow Vorstand$

- Geben Sie sämtliche Schlüsselkandidaten an.
- Welcher Normalform genügt die Relation *Vereine*? Begründen Sie Ihre Antwort ausführlich.
- Zeigen Sie anhand von Beispielen die drei möglichen Anomalien auf, die bei der Verwendung der Relation *Vereine* auftreten können!
- Verwenden Sie den Synthesalgorithmus, um die Relation in die dritte Normalform zu überführen.
- Befinden sich die neuen Relationen in Boyce-Codd-Normalform? Begründen Sie Ihre Antwort.

Fortsetzung nächste Seite!

Teilaufgabe 2:**1. Allgemeinwissen**

- 1.1 Bewerten Sie die folgenden Aussagen. Geben Sie für die folgenden Aussagen an, ob sie richtig oder falsch sind. Begründen Sie ihre Antworten kurz!
- a. B-Bäume verlieren ihre Vorteile gegenüber binären Bäumen in Hauptspeicherdatenbanken.
 - b. Eine Relation kann mehrere Primärschlüsselattribute besitzen.
 - c. Heutzutage werden zumeist hierarchische Datenbanksysteme eingesetzt.
 - d. Fremdschlüsselwerte dürfen nicht NULL sein.
 - e. Alle Metadaten werden in Tabellen gespeichert.
- 1.2 Isolation von Transaktionen: Erläutern Sie kurz die Anomalien und geben Sie jeweils ein Beispiel an!
- a. Lost upgrade
 - b. Dirty read
 - c. Non-repeatable read
 - d. Phantom

2. Relationale Algebra

- 2.1 Nennen Sie die fünf Grundoperationen der relationalen Algebra.
- 2.2 Warum wird in gewöhnlichen Datenbanken die relationale Algebra nicht strikt eingehalten? Geben Sie ein Beispiel an und erklären Sie, wie man die relationale Eigenschaft erzwingen kann.

Fortsetzung nächste Seite!

3. ER-Modellierung und Relationenmodell

- 3.1 Nennen und erläutern Sie kurz vier Beispiele für Integritätsbedingungen.
- 3.2 Sie sollen die Datenbank eines Auktionshauses modellieren. Erstellen Sie ein ER-Diagramm bestehend aus Entitäts- und Beziehungstypen sowie Attributen. Geben Sie auch die Kardinalitäten mit an.

Verwenden Sie bei Entitäten und Attributen ausschließlich die beschriebenen, soweit dies möglich ist. Verwenden Sie auch keine künstlichen Schlüssel, es sei denn, diese sind der Aufgabenstellung eindeutig zu entnehmen. Die Kardinalitätseinschränkungen können Sie entweder in der (min,max)-Notation oder der einfachen Notation nach Chen (1:1, 1:N, M:N) angeben.

In der Formulierung der Beziehungen sind die Kardinalitätseinschränkungen genau zu beachten. Im Allgemeinen gilt: Alles was nicht explizit eingeschränkt ist, muss im Modell auch uneingeschränkt bleiben (keine Einschränkungen in eine Beschreibung hineininterpretieren). Alles was explizit eingeschränkt ist, darf im Modell nicht uneingeschränkt bleiben (alle formulierten Einschränkungen sollen erkennbar sein).

1. Im Zentrum des Systems stehen die Benutzer. Sie werden durch eine ID identifiziert und haben einen Benutzernamen, sowie eine Adresse, die sich aus Strasse, PLZ und Ort zusammensetzt.
 2. Benutzer initiieren Auktionen. Jede Auktion ist genau einem Benutzer zugeordnet und wird ebenfalls über eine ID identifiziert. Daneben hat sie einen Namen, sowie Beginn- und Endezeitpunkt.
 3. Benutzer geben Gebote für Auktionen ab. Zu jedem Gebot werden Betrag und Zeitpunkt des Gebotes gespeichert. Natürlich soll ein Benutzer auch mehrfach auf eine Auktion bieten können.
 4. Zu guter Letzt existiert ein Bewertungssystem. Eine Bewertung wird von einem Benutzer zu einer bestimmten Auktion abgegeben. Zu ihr sollen die Bewertung selbst (ein numerischer Wert) und der Zeitpunkt der Bewertung gespeichert werden. Pro Benutzer und Auktion kann nur eine Bewertung abgegeben werden.
- 3.3 Erstellen Sie zu dem ER-Diagramm aus Aufgabe 3.2 ein Relationenschema. Seien Sie dabei so vollständig wie möglich – bspw. Berücksichtigung totaler Partizipationen. Vermeiden Sie unnötiges Ausprägen von Relationen bei allen Beziehungen.

Beispiel für die Notation:

Relationenname (Primärschlüssel, Attribut1, Attribut2, ...,
Fremdschlüssel[AndereRelation],
(FKAttrA, FKAttrB)[DritteRelation])
Attribut2 NOT NULL

Fortsetzung nächste Seite!

4. Normalformen

- 4.1 Die Normalisierung von Relationenschemata dient der Vermeidung von Redundanzen und dadurch bedingter Anomalien. Diskutieren Sie anhand des folgenden Beispiels die Einfüge-, Änderungs- und Löschanomalien.

AbtNr	AbtName	AbtOrt	MitarbeiterNr	MitarbeiterName
1	Entwicklung	München	1	Schmidt, H.
2	Forschung	München	2	Müller, G.
1	Entwicklung	München	3	Kohl, H.
3	Fertigung	Nürnberg	4	Weizenbaum, L.
4	Logistik	Passau	5	Nienhof, R.
6	Vertrieb	Würzburg	6	Baier, V.

- 4.2 In welcher Normalform ist das Beispiel aus Aufgabe 4.1? Welche Bedingung der nächsthöheren Normalform ist verletzt? Berücksichtigen Sie, dass eine Relation nur dann in n-ter Normalform ist, wenn sie die Bedingungen aller m-ten Normalformen mit $m \leq n$ erfüllt.

Fortsetzung nächste Seite!

Themenschwerpunkt B
(Betriebssysteme)**Teilaufgabe 1:****1. Scheduling, Round-Robin, FCFS, SJF**

- a) Beschreiben Sie die Prozess-Scheduling-Strategien „Round-Robin“, „First Come First Served“ und „Shortest Job First“ knapp (in jeweils 1–2 Sätzen). Welche der Strategien sind präemptiv, welche nicht?
- b) „Round-Robin“-Scheduler benutzen üblicherweise eine Liste aller derzeit ausführbarer Prozesse, wobei jeder Prozess maximal einmal in der Liste vorkommt.

Was würde passieren, wenn ein Prozess zweimal in der Liste vorkäme? Was für einen Grund könnte es geben, das zu erlauben?

- c) Es sollen fünf Prozesse mit verschiedenen Lauf- und Ankunftszeiten ausgeführt werden:

Prozess	Laufzeit	Ankunftszeit
p_1	5	2
p_2	4	3
p_3	4	0
p_4	2	9
p_5	3	6

- Geben Sie an, in welcher Reihenfolge ein Scheduler mit der Strategie „First Come First Served“ die Prozesse ausführen würde.
 - In welcher Reihenfolge würde ein Scheduler mit der Strategie „Shortest Job First“ die Prozesse ausführen?
 - Bestimmen Sie zu beiden Schedules die mittlere Durchlaufzeit.
- d) Welche der drei Strategien ist beweisbar optimal? Unter welchen Voraussetzungen? (Die Angabe des Beweises ist nicht notwendig.)

Fortsetzung nächste Seite!

2. Verdrängungsstrategien

- a) Paging-Verdrängungsstrategien treffen ihre Entscheidungen basierend auf Kennzahlen wie Anzahl der Zugriffe, Zeit des ersten Zugriffs, Zeit des nächsten Zugriffs oder Zeit des letzten Zugriffs. Ordnen Sie die vier oben genannten Kennzahlen den folgenden vier Strategien zu:
- First-In, First-Out (FIFO);
 - (Hypothetische) optimale Strategie (OPT);
 - Least Frequently Used (LFU);
 - Least Recently Used (LRU).
- b) Es stehen drei Seitenrahmen zur Verfügung, die anfangs leer sind. Wenden Sie die vier Strategien FIFO, LFU, LRU und OPT auf die Folge von Seitenanforderungen an:

2; 4; 2; 1; 1; 5; 1; 3; 1; 2.

Geben Sie für jede der Strategien die Belegung der drei Seitenrahmen nach jeder Seitenanforderung an. Notieren Sie jeweils in Klammern den Wert der relevanten Kennzahl aus dem Aufgabenteil a). Geben Sie für jede Strategie die Anzahl der vorgenommenen Verdrängungen an.

3. Deadlocks, Bankier-Algorithmus

- a) Nennen Sie die vier Eigenschaften, die ein Betriebssystem besitzen muss, so dass ein Deadlock auftreten kann. Beschreiben Sie knapp die Eigenschaften. (Eine Begründung, warum diese Eigenschaften zu einem Deadlock führen können, ist nicht notwendig.)
- b) Beim Bankieralgorithmus wird der Begriff des „sicheren Zustands“ verwendet. Wie lautet die Definition eines sicheren Zustands?
- c) Führt jeder unsichere Zustand unweigerlich zu einem Deadlock? Begründen Sie Ihre Antwort kurz.

Fortsetzung nächste Seite!

4. Dateisysteme

- a) Skizzieren Sie kurz (jeweils 1-2 Sätze) die Vor- und Nachteile der folgenden drei Konzepte zur Realisierung von Dateien in einem Dateisystem:

- Zusammenhängende Belegung
- Verkettete Listen
- I-Nodes

Betrachten Sie hierbei insbesondere die Effizienz beim wahlfreien Zugriff, den Verbrauch an Hauptspeicher und die Probleme beim Löschen und Vergrößern von Dateien.

- b) Wie läuft ein wahlfreier Zugriff auf das Byte Nr. 50000 einer Datei beim I-Nodes Konzept ab?

Der entsprechende I-Node sei schon im Hauptspeicher vorhanden. Die Blockgröße sei 4 KB und die Zeigergröße sei 4 Byte. Die Anzahl der direkten Zeiger ist 10. Die Nummerierung der Bytes fängt mit der Nummer 0 an.

Bitte geben Sie an, welche Zeiger daran beteiligt sind, an welcher Position in den Blöcken diese zu finden sind und wohin sie zeigen.

Fortsetzung nächste Seite!

Teilaufgabe 2:**1 Systempuffer**

- 1.1 Erläutern Sie die Schritte, die die Pufferverwaltung durchführen muss, wenn eine Seite angefordert wird.
- 1.2 Für einen Puffer stehen drei Frames im Hauptspeicher zur Verfügung. Zu Beginn seien alle Frames leer. Geben Sie an, welche Seiten zu welchen Zeitpunkten in den Frames des Prozesses eingelagert sind. Notieren Sie auch die Kontrollzustände für jede Belegung. Geben Sie auch die Anzahl der nötigen Seiteneinlagerungen an.
Verwenden Sie die Ersetzungsstrategie LRU.

Seitenreferenz		1	2	3	4	5	6	7	8	9	10	11	12
Referenzfolge		1	3	2	5	3	1	4	4	2	4	1	5
Hauptspeicher	Kachel 1												
	Kachel 2												
	Kachel 3												
Kontrollzustände (Zeit des letzten Zugriffs)	Kachel 1												
	Kachel 2												
	Kachel 3												

- 1.3 Nennen Sie zwei weitere Ersetzungsstrategien und charakterisieren Sie diese kurz!

Fortsetzung nächste Seite!

2. Programme, Prozesse, Threads

- a) Was versteht man unter einem Programm, einem Prozess und einem Thread?
- b) Ein Programm kann von mehreren Prozessen ausgeführt werden.
Wie kann diese Situation entstehen?
Was sind die Gemeinsamkeiten, was sind die Unterschiede zwischen den Prozessen?
- c) Wie werden Prozesse vom Betriebssystem verwaltet?
Welche Eigenschaften, Attribute etc. werden für die Verwaltung benutzt?
- d) Welche strategischen Entscheidungen muss das Betriebssystem bezüglich der Zuteilung von Ressourcen an Prozesse treffen?
Geben Sie Beispiele für Vorgehensweisen und konkrete Strategien bei den Ressourcen CPU (drei Beispiele) und Speicher (zwei Beispiele) an und erläutern Sie diese.
- e) Beschreiben Sie die unterschiedlichen Möglichkeiten zur Realisierung von Threads.
Was sind jeweils die Vor- und Nachteile?

3. Speicher und Adressräume

- a) Eine wesentliche Systemfunktion bei Mehrprogrammbetrieb ist Adressraumschutz. Erläutern Sie, was man darunter versteht und warum Adressraumschutz essentiell ist. Vor welchen Gefahren schützt er?
- b) Erläutern Sie in Stichworten und jeweils anhand einer Skizze die Funktionsweise von "Adressraumschutz durch Eingrenzung" sowie von "Adressraumschutz durch Segmentierung". Erläutern Sie Vor- und Nachteile und geben Sie dabei auch an, zu welchem Zeitpunkt jeweils die Bindung von Programmadressen an Arbeitsspeicheradressen erfolgt, wer die Bindung besorgt und wie der Zugriffsschutz konkret funktioniert.

Fortsetzung nächste Seite!

4 Synchronisation

Ein Datensicherungsprogramm sei so realisiert, dass es in einer ersten Phase eine Reihe von m Dateikatalogen parallel auf unterschiedliche Bandlaufwerke sichert. Es stehen n Bandlaufwerke zur Verfügung, für jedes Laufwerk gibt es eine Datenstruktur, die alle für den Zugriff erforderlichen Daten enthält. Die Datenstrukturen der freien Geräte sind in einem Ringpuffer verwaltet.

Ein "Haupt-Thread" arbeitet die m Dateikataloge der Reihe nach ab. So lange freie Laufwerke verfügbar sind, erzeugt er für jeden Dateikatalog einen eigenen "Arbeiter-Thread", entnimmt die Datenstruktur eines freien Laufwerks aus dem Ringpuffer und übergibt sie zusammen mit der Information über den zu bearbeitenden Dateikatalog dem Arbeiter-Thread. Sind keine Laufwerke mehr verfügbar, wartet er, bis wieder ein Laufwerk frei wird, und erzeugt dann den nächsten Arbeiter-Thread.

Sobald ein Arbeiter-Thread mit der Sicherung seines Dateikatalogs fertig ist, trägt er die Datenstruktur seines Laufwerks wieder in den Ringpuffer ein und terminiert.

Der Haupt-Thread wartet schließlich, bis alle Arbeiter-Threads fertig sind, und erzeugt dann eine Abschlussmeldung.

- a) An welchen Stellen ist zwischen den beteiligten Threads Synchronisation erforderlich?
- b) Skizzieren Sie in einer Programmiersprache Ihrer Wahl die wesentlichen Datenstrukturen und Schritte im Programmablauf sowie die Synchronisation mit Hilfe von Semaphoren. Erläutern Sie kurz die Bedeutung der von Ihnen eingesetzten Semaphore und welche Werte sie initial haben müssen.

Fortsetzung nächste Seite!

5 SQL

Gegeben ist folgende Datenbank:

SQL> select * from angestellte;

PERS_NR	ABT_NR	NAME	GRUNDGEHALT	BONUS
1	1	Schmidt	40.000	4.000
2	1	Maier	45.000	6.000
3	3	Ewing	38.000	
...

SQL> select * from abteilung;

ABT_NR	NAME	TYP	UMSATZ
1	V Region Nord	VERTRIEB	3.000.000
2	V Region Süd	VERTRIEB	4.500.000
3	F&E	ENTWICKLUNG	
...

In angestellte sind PERS_NR Primärschlüssel und ABT_NR Fremdschlüssel.

In abteilung ist ABT_NR Primärschlüssel.

SQL-Anfragen

Schreiben Sie SQL-Statements für die folgenden Fragestellungen. Wenn Sie möchten, können Sie zur Vereinfachung Sichten anlegen.

- Welche Abteilungen haben mehr als 10 Angestellte?
- Welche Mitarbeiter verdienen mindestens 25% mehr als der Durchschnitt ihrer Abteilung? Maßgeblich ist die Summe aus Grundgehalt und Bonus.
- Welches sind die drei Abteilungen vom Typ VERTRIEB, die das schlechteste Verhältnis von UMSATZ zu Personalkosten haben? Die Personalkosten bestimmen sich dabei aus der Summe der Gehälter der Angestellten der jeweiligen Abteilung (wiederum Grundgehalt plus Bonus). Geben Sie neben dem Abteilungsnamen auch das Verhältnis von Umsatz und Personalkosten aus.