

DBIS PROBEKLAUSUR

SOMMERSEMESTER 2011

1 Konzeptuelle Modellierung (12 Punkte)

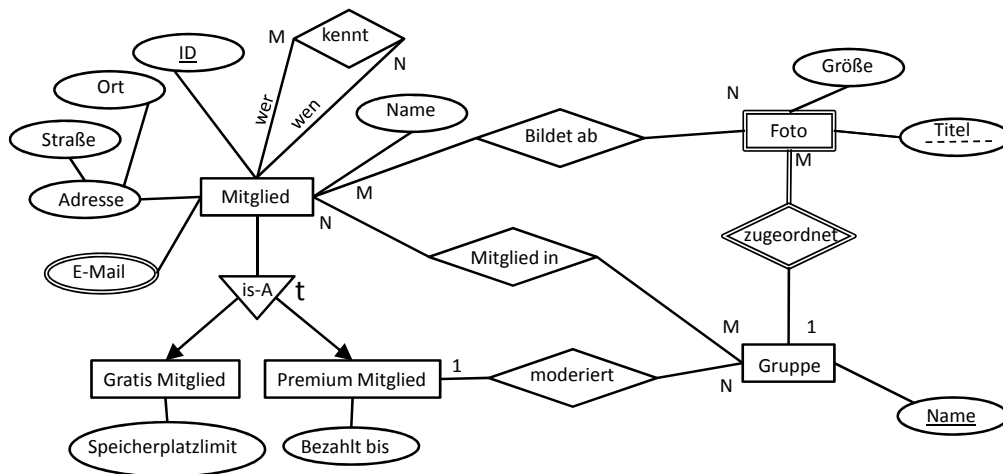
Eine Pizzeria möchte eine Datenbank zur Verwaltung ihres Lieferdienstes erstellen. Hierfür sollen die folgenden Anforderungen gelten (*Anmerkungen, die nur der Erklärung dienen, aber nicht modelliert werden sollen, sind kursiv gedruckt.*):

- Die Pizzeria hat verschiedene Kundentypen. Sie unterscheidet ausschließlich zwischen Stammkunden, Onlinekunden sowie Telefonkunden. Kunden können natürlich in mehrere dieser Kategorien fallen (*z.B. ein Stammkunde, der mal telefonisch und mal online bestellt hat*). Kunden werden erst gespeichert, wenn sie mindestens einmal etwas bestellt haben.
- Jeder Kunde ist eindeutig über seine Kundennummer identifizierbar und hat einen Namen, sowie eine Adresse bestehend aus der Straße, PLZ und Ort. Bei Stammkunden wird zudem der Geburtstag gespeichert (*so kann die Pizzeria eine Geburtstagskarte schicken und die Kundenbindung pflegen*). Für jeden Onlinekunden wird eine E-Mail-Adresse gespeichert. Bei Telefonkunden wird die Telefonnummer (*für eventuelle Rückfragen*) gespeichert.
- Kunden geben Bestellungen auf, wobei jede Bestellung von genau einem Kunden stammt. Zu jeder Bestellung wird eine eindeutige BestellID gespeichert.
- Eine Bestellung besteht aus mindestens zwei Gerichten (*sonst lohnt sich für die Pizzeria die Auslieferung nicht*).
- Die Pizzeria unterscheidet zwischen genau vier Arten von Gerichten: Pizza, Pasta, Salat und Eis. Alle Gerichte haben einen Preis und eine eindeutige Nummer. Zu der Pizza wird die Größe gespeichert und die Beläge (in der Regel mehrere!). Zu der Pasta werden die Nudelsorte, sowie die Soße und bei Salaten das Dressing abgespeichert.
- Jedem Stammkunden kann genau ein Eis als Lieblingsnachspeise zugeordnet werden.
- Eine Liefer-Tour besteht aus maximal 5 Bestellungen (*sonst wird das Essen kalt*). Zu jeder Liefer-Tour wird das Datum und die Uhrzeit gespeichert, zudem wird zu jeder Liefer-Tour das Kennzeichen des zugehörigen Fahrzeugs gespeichert.
- Jeder Liefer-Tour ist genau ein Lieferant zugeordnet. Die Liefer-Tour ist eindeutig anhand des Lieferanten und dem Datum und der Uhrzeit zu identifizieren.
- Lieferanten selbst werden durch ihre Personalnummer identifiziert, zudem werden der Name und (ggf. mehrere) Telefonnummern gespeichert. Auch neue Lieferanten, die noch nichts ausgeliefert haben werden direkt in die Datenbank aufgenommen.

Modellieren Sie ein ER-Diagramm, das den obigen Anforderungen entspricht. Verwenden Sie zur Repräsentation der Kardinalitäten die Min-Max-Notation.
(12 Punkte)

2 Relationale Datenbanken (29 Punkte)

Folgendes ER-Diagramm zeigt ein soziales Netzwerk:

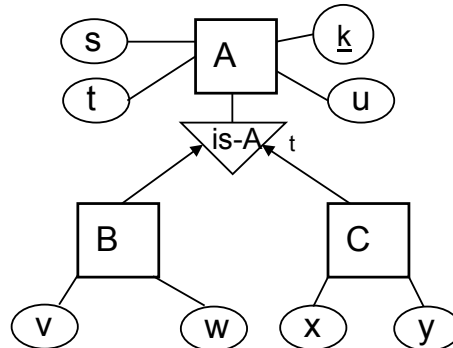


Hinweis: Die Relation *kennt* gibt an, wer wen kennt. Sie können davon ausgehen, dass Alice genau dann Bob kennt wenn Bob auch Alice kennt. Anders ausgedrückt: Wie in einem realen Sozialen Netzwerk beruht *kennen* immer auf Gegenseitigkeit (zudem *kennt* niemand sich selbst). Dies ist für das Datenbankschema unerheblich, vereinfacht aber ggf. die Anfragen.

- Bilden Sie das obige ER-Diagramm in ein relationales Datenbankschema ab. Verwenden Sie die formale Notation, wie sie in der Vorlesung und Übung vorgestellt wurde, einschließlich intra- und interrelationaler Abhängigkeiten.
 (8 Punkte)
- Drücken Sie das folgende Informationsbedürfnis in SQL aus: Die Namen aller "Premium Mitglieder", die auf keinem Foto gemeinsam mit dem Mitglied "Hans Werner" abgebildet sind.
Hinweis: Mitglieder, die auf keinem Foto sind, sind auch auf keinem Foto mit "Hans Werner"
 (6 Punkte)
- Drücken Sie das folgende Informationsbedürfnis durch einen Ausdruck der relationalen Algebra aus: Alle Namen von Mitgliedern, die den Moderator mindestens einer der Gruppen kennen, in der sie selbst Mitglied sind.
 (4 Punkte)
- Bearbeiten Sie diese Aufgabe nicht, wenn Sie die 90-minütige Klausur schreiben!**
 Drücken Sie das folgende Informationsbedürfnis durch einen Ausdruck in SQL aus: Zu jeder Gruppe soll die Anzahl der Mitglieder und der Name des Moderators ausgegeben werden. Ausgeschlossen werden sollen die Gruppen, deren Moderator höchstens bis zum 1.9.2010 gezahlt hat. Es sollen nur Gruppen beachtet werden, die mindestens ein "Gratis Mitglied" und insgesamt mindestens 50 Mitglieder haben.
 (6 Punkte)
- Bearbeiten Sie diese Aufgabe nicht, wenn Sie die 90-minütige Klausur schreiben!**
 Drücken Sie das folgende Informationsbedürfnis durch einen Ausdruck im Relationen-Domänenkalkül aus: Die Namen aller "Gratis Mitglieder", die in den Gruppen in denen sie Mitglied sind keine anderen Gruppenmitglieder kennen. Dies schließt auch die "Gratis Mitglieder" ein, die in keiner Gruppe Mitglied sind.
 (5 Punkte)

3 Relationale Datenbanken (6 Punkte)

Betrachten Sie das folgende ER-Diagramm:



- Sie kennen aus der Vorlesung vier Möglichkeiten eine is-A Beziehung in ein relationales Datenbankschema zu überführen. Überlegen Sie welche der Möglichkeiten hier umgesetzt werden können und beschreiben Sie diese mit eigenen Worten.
- Nennen Sie Vor- bzw. Nachteile der von Ihnen verwendeten Möglichkeiten.

(6 Punkte)

4 Entscheidungsfragen (7 Punkte)

Entscheiden Sie für jede der folgenden Aussagen, ob sie wahr oder falsch ist. Eine Begründung Ihrer Entscheidung brauchen Sie nicht anzugeben. Es gilt folgendes Bewertungsschema:

Für jede Aussage: korrekte Antwort: 1 Pkt., falsche Antwort: 0 Pkt., keine Antwort: 0 Pkt.

- Für die Definition von SQL Schemata verwendet man SQL-DDL.
- In einem ER-Diagramm zeigen bei einer disjunkten is-A-Beziehung die Pfeile auf die erbenden Entitätstypen.
- Wenn $X \rightarrow A$ und $Y \rightarrow B$ funktionale Abhängigkeiten für eine Relation $r \in Rel(Z)$ sind (mit $X \subseteq Z, Y \subseteq Z, A, B \in Z$), so gilt auch $XY \rightarrow AB$.
- Zwei Operationen $p \in op(t)$ und $q \in op(t')$ zweier Transaktionen t und t' ($t \neq t'$) stehen genau dann in einem Schedule s in Konflikt, wenn sie auf demselben Objekt x arbeiten und sowohl p als auch q Schreiboperationen sind.
- In einem B^+ -Baum mit $k = 2$ und $k^* = 2$ haben innere Knoten mindestens 3, aber höchstens 5 Söhne.
- Für einen nicht leeren B-Baum der Ordnung k gilt: Ein innerer Knoten mit b Schlüsseln hat genau b Söhne.
- Sei A eine Attributmenge und $a \in Rel(A)$, dann gilt: $a \times b = \Pi_A(a \times b)$.

5 Funktionale Abhängigkeiten und Normalformen (14 Punkte)

Betrachten Sie das folgende Relationenschema:

$$R = (U, F) = (\{ABCDEG\}, \{A \rightarrow BE, AE \rightarrow BD, G \rightarrow CD, CD \rightarrow BEG, CG \rightarrow B\})$$

- Identifizieren Sie alle Schlüsselkandidaten von R , zeigen Sie, dass es sich um Schlüsselkandidaten handelt und beweisen Sie, dass es keine anderen Schlüsselkandidaten gibt. (4 Punkte)
- Bearbeiten Sie diese Aufgabe nicht, wenn Sie die 90-minütige Klausur schreiben!**
Was ist die höchste Normalform von R ? Begründen Sie. (2 Punkte)
- Bearbeiten Sie diese Aufgabe nicht, wenn Sie die 90-minütige Klausur schreiben!**
Wenden Sie den Synthese-Algorithmus zur Überführung des Schemas in die 3. Normalform an (8 Punkte)

6 Serialisierbarkeit und Nebenläufigkeit (9 Punkte)

Betrachten Sie die folgenden Schedules:

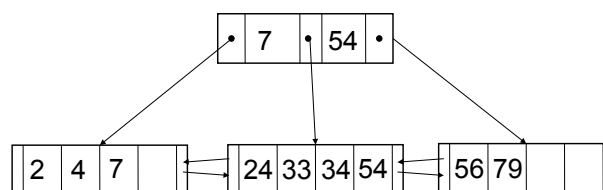
$$s_1 = w_1(x) r_1(y) r_2(y) w_2(y) c_2 r_3(y) w_3(x) w_1(y) c_3 r_1(x) c_1$$

$$s_2 = r_1(y) w_1(y) r_2(y) w_1(x) r_1(x) c_1 w_2(y) w_3(x) c_2 r_3(y) c_3$$

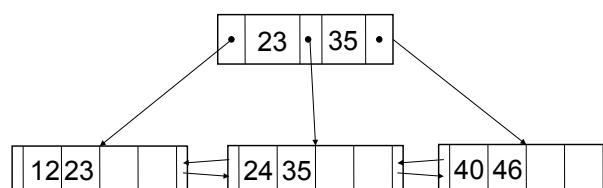
- Zeigen oder widerlegen Sie, dass diese beiden Schedules konfliktäquivalent sind. (3 Punkte)
- Sind die Schedules konfliktserialisierbar? Begründen Sie Ihre Antwort. (2 Punkte)
- Bearbeiten Sie diese Aufgabe nicht, wenn Sie die 90-minütige Klausur schreiben!**
Entscheiden Sie für die gegebenen Schedules, und jede der Klassen RC, ACA und ST, ob sich der Schedule in der jeweiligen Klasse befindet. Begründen Sie ihre Entscheidung indem Sie für die relevanten Konflikte angeben, warum diese der Definition der jeweiligen Klasse (nicht) widersprechen. (4 Punkte)

7 B⁺-Bäume (4 Punkte)

- Fügen Sie in folgenden B^+ -Baum ($k = 1, k^* = 2$) das Element 10 ein.



- Löschen Sie aus folgendem B^+ -Baum ($k = 1, k^* = 2$) das Element 23.



8 XPath und XQuery (19 Punkte)

Betrachten Sie das folgende XML-Dokument als ein Beispiel eines (wahrscheinlich größeren) XML-Dokuments gleicher Struktur, das eine Datenbank von Auftraggebern und Aufträgen einer Werbeagentur darstellt:

```
<Auftragsdatenbank>
  <Kunde>
    <Name>Gabis Imbiss</Name>
    <ID KundenID="4327"/>
    <Adresse>
      <Strasse>Hamsterweg 7</Strasse>
      <PLZ>91282</PLZ>
      <Ort>Hunger</Ort>
    </Adresse>
    <Zahlungsmoral Bewertung="gut"/>
  </Kunde>

  <Kunde>
    <Name>Konditor Opper</Name>
    <ID KundenID="7342"/>
    <Adresse>
      <Strasse>Brunnenstrasse 51</Strasse>
      <PLZ>73329</PLZ>
      <Ort>Kuchen</Ort>
    </Adresse>
    <Zahlungsmoral Bewertung="schlecht"/>
  </Kunde>

  <Kunde>
    <Name>Fahrrad Fritz</Name>
    <ID KundenID="9458"/>
    <Adresse>
      <Strasse>Birkenweg 23</Strasse>
      <PLZ>21449</PLZ>
      <Ort>Radbruch</Ort>
    </Adresse>
    <Zahlungsmoral Bewertung="gut"/>
  </Kunde>

  ...

  <Auftrag Art="Printwerbung">
    <Kosten PreisInEuro="1000" Anzahlung="500"/>
    <Kunde>9458</Kunde>
    <Mitarbeiter>Marietta Weyand</Mitarbeiter>
    <Mitarbeiter>Lisa Paulus</Mitarbeiter>
    <Optionen Farbe="SchwarzWeiss" Schmuckfarbe="keine"/>
  </Auftrag>
```

```

<Auftrag Art="Printwerbung">
  <Kosten PreisInEuro="750" Anzahlung="250"/>
  <Kunde>7342</Kunde>
  <Mitarbeiter>Anton Wetzels</Mitarbeiter>
  <Optionen Farbe="4C" Schmuckfarbe="Gold"/>
</Auftrag>

<Auftrag Art="Radio-Werbung">
  <Kosten PreisInEuro="5000" Anzahlung="1000"/>
  <Kunde>9458</Kunde>
  <Länge>0:23</Länge>
  <Mitarbeiter>Anton Wetzels</Mitarbeiter>
  <Regie>Simone Bratenstein</Regie>
</Auftrag>

<Auftrag Art="Radio-Werbung">
  <Kosten PreisInEuro="3000" Anzahlung="500"/>
  <Kunde>4327</Kunde>
  <Länge>0:15</Länge>
  <Mitarbeiter>Christina Gräther</Mitarbeiter>
  <Regie>Rudolf Dhein</Regie>
</Auftrag>

```

...

</Auftragsdatenbank>

- a) Adressieren Sie durch Angabe entsprechender XPath-Ausdrücke die wie folgt beschriebenen XML-Knoten
- (a) Alle Printwerbungen, deren Kosten unter 1000 Euro liegen.
(3 Punkte)
 - (b) Alle KundenIDs von Kunden, die vom Mitarbeiter "Anton Wetzels" betreut werden.
(3 Punkte)
 - (c) Alle Aufträge für Printwerbungen, bei denen weder Farbe (somit Farboption "SchwarzWeiss") noch Schmuckfarben verwendet werden.
(3 Punkte)

- b) Formulieren Sie eine XQuery-Anfrage, die alle Projektkosten und Anzahlungen von Projekten aller Kunden auflistet, die eine gute Zahlungsmoral haben. Das Ergebnis soll in folgender Form angegeben werden: (10 Punkte)

```

<Projektliste>
  <Kunde>
    <Firma Name="Gabis Imbiss"/>
    <Projekt>
      <Radio-Werbung>
        <Kosten PreisInEuro="3000" Anzahlung="500"/>
      </Radio-Werbung>
    </Projekt>
  </Kunde>
  <Kunde>
    <Firma Name="Fahrrad Fritz"/>
    <Projekt>
      <Printwerbung>
        <Kosten PreisInEuro="1000" Anzahlung="500"/>
      </Printwerbung>
    </Projekt>
    <Projekt>
      <Radio-Werbung>
        <Kosten PreisInEuro="5000" Anzahlung="1000"/>
      </Radio-Werbung>
    </Projekt>
  </Kunde>
</Projektliste>

```