

Universität Dortmund SoSe 2004
Übungen zu *Wissensentdeckung in Datenbanken*

Blatt 10. Abgabe bis Montag, den 5.7.2004

Zur Hilfestellung beim Umgang mit der Datenbank und SQL werden sich Michael Wurst und Timm Euler **am Dienstag, den 29.6.04 in der Zeit von 13 bis 15 Uhr** wieder in den Rechnerpools im GB V aufhalten.

Auf den Rechnern des Fachbereichs Informatik steht Ihnen ein Datenbank-Benutzeraccount zur Verfügung; der Benutzername ist derselbe wie für die Rechner selbst (also wid0XX), das Passwort teilen wir aus oder Sie können es per Email erfragen. Am Terminal geben Sie bitte ein:

```
module add oracle
```

und dann

```
sqlplus wid0XX@fbi
```

(Der Name der Datenbank ist `fbi`.) Nach Eingabe des Passwortes können Sie nun SQL-Befehle eingeben. Sie beenden mit `exit`. Um die Eingabe der Befehle zu vereinfachen, können Sie die Befehle auch zeilenweise in eine Datei schreiben und die Datei dann am SQL-Prompt mit `start <dateiname>` aufrufen. Bei langen Befehlen ist dies sogar notwendig, da die Eingabelänge am Prompt beschränkt ist. Jeder Befehl muss mit einem Semikolon enden. Achtung: wenn Sie Tabellen angelegt, geändert oder Inhalte modifiziert haben, müssen Sie den Befehl `commit`; eingeben, um den aktuellen Stand über die aktuelle Sitzung hinaus zu speichern.

Weitere SQL-Hinweise: Das Passwort lässt sich ändern mit dem Befehl `Alter User wid0XX Identified By <neuesPassw>`. Um den Inhalt einer Tabelle zu löschen, steht der Befehl `delete from <tabellenname>` zur Verfügung, der auch mit einer `Where`-Bedingung modifiziert werden kann; die Tabelle selbst wird mit `drop table <tabellenname>` entfernt. Der Durchschnittswert einer Spalte wird mit `Select AVG(<Spaltenname>) FROM...` ausgegeben; analog funktionieren `MIN`, `MAX`, `SUM`, `COUNT`. Die verschiedenen vorkommenden Werte einer Spalte liefert `Select DISTINCT(<Spaltenname>)`

FROM. . . . Sie finden außerdem viele gute SQL-Tutorials im Internet.

Aufgabe 1 In dieser Aufgabe werden einige Datenbanktabellen angelegt, die in den folgenden Aufgaben verwendet werden.

Wir verwenden einen Auszug aus einer Datenbank einer Telekommunikationsfirma, die jeden Anruf jedes Kunden mit Zeit und Datum, Länge usw. in einer Tabelle speichert. Eine zweite Tabelle enthält Informationen über die Kunden, die mit ihrer Telefonnummer identifiziert werden.

Legen Sie eine Tabelle `CallDetails` an mit den numerischen Attributen `Caller`, `Called`, `Calllength`, `Callunits`, dem Textattribut `Callclass` und dem Datumsattribut `Calltime` (SQL-Datentypen `NUMBER`, `VARCHAR2` und `DATE`). Legen Sie eine zweite Tabelle `Customers` an mit dem numerischen Attribut `Caller`, den Textattributen `Gender`, `TariffType`, `PayMethod` und dem Datumsattribut `Birthday`. (Das Attribut `Caller` dient als Schlüsselattribut.)

Auf dem Server finden Sie die beiden Dateien `importCalldetails.sql` und `importCustomers.sql`. Rufen Sie sie am SQL-Prompt mit `start` auf, um Ihre beiden Tabellen mit Daten zu füllen. (Die Daten wurden künstlich erzeugt.)

Geben Sie bitte die SQL-Anweisungen ab, mit denen Sie die Tabellen erzeugt haben.

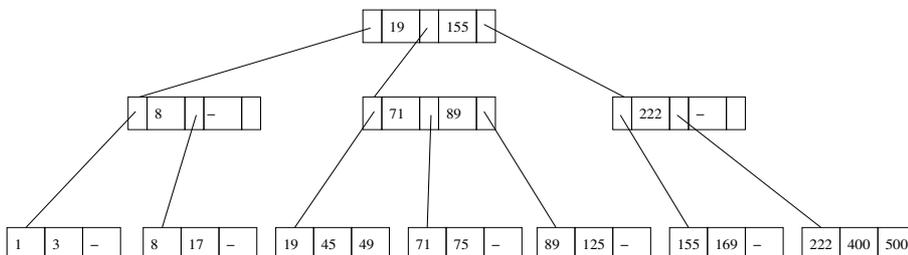
Aufgabe 2 1. Für die folgenden Fragen geben Sie bitte die jeweilige SQL-Abfrage ab, die sie beantwortet (nicht die Antwort auf die Frage abgeben). Versuchen Sie, möglichst wenige Abfragen pro Frage zu verwenden; notwendig ist jeweils nur eine.

- (a) Wieviele Kunden (also Anrufer, `Caller`) bezahlen nicht per "Pre Pay"?
- (b) In welchen Klassen (Attribut `callclass`) kommen Anrufe der Länge 0 vor?
- (c) Wie lang war der längste Anruf eines Mannes?
- (d) Hat jemand sich selbst angerufen?
- (e) Welcher Anrufer führte das längste Gespräch?

2. Erzeugen Sie eine Sicht, die genau einen Eintrag enthält für jede Frau, die sowohl Anrufe in Klasse 12 als auch in Klasse 19 (Attribut `callclass`) getätigt hat. Versuchen Sie wieder, möglichst

wenige SQL-Befehle zu verwenden; notwendig ist nur einer. (Der resultierende Befehl ist vermutlich zu lang für eine Eingabe am Prompt, siehe oben.)

Aufgabe 3 In dieser Aufgabe geht es um eine Speicherstruktur namens B-Tree. Beim Einfügen und Löschen in solchen Bäumen müssen einige Dinge beachtet werden. Im folgenden sollen Sie auf dem B-Tree in der Abbildung unten mehrere Einfüge- und Löschoperationen durchführen. Geben Sie dabei jeweils den Baum an, der nach jeder Operation entsteht, sowie für jede Operation alle Schritte, die ausgeführt werden müssen. Die folgenden Operationen sollen durchgeführt werden:



1. Einfügen von 18
2. Einfügen von 33
3. Löschen von 8
4. Löschen von 71

Hinweis: Das Löschen von Knoten wurde zwar nicht explizit in der Vorlesung besprochen, aber es sollte Ihnen keine Probleme bereiten, sich den entsprechenden Algorithmus aus der Grundstruktur von B-Trees herzuleiten. Alternativ können Sie auch Gebrauch von entsprechender Literatur machen. Geben Sie aber in jedem Fall jeden einzelnen Schritt an, der jeweils beim Löschen durchgeführt werden muss.