



Institut für Informatik - Lehrstuhl Datenbanken

Datenbanken I

Wintersemester 11/12 - Dr. Alexander Hinneburg

Übung 3: Relationales Modell, Integritätsbedingungen
(Abgabe bis 7.11.2011, 13:45 Uhr)

12 Punkte

Aufgabe 3.1:

6 Punkte

Schreiben sie ein Programm in einer beliebigen Programmiersprache, das eine CSV-Datei lesen kann und für eine gegebene Menge von Attributen prüft, ob jede auftretende Kombination der Attributwerte nur einmal vorkommt (Schlüsselbedingung siehe Folien 2-46 bis 2-48). Geben Sie falls, die Schlüsselbedingung nicht gilt, alle Belegungen der Attribute aus, welche die Bedingung verletzen, das heisst mehr als einmal vorkommen. Wenn mehr als 5 verschiedene Belegungen gefunden worden, geben Sie die weiteren nicht aus, sondern nur die Anzahl der verschiedenen Belegungen, die mehr als einmal vorkommen. Die Eingaben für das Programm – Dateiname und zu prüfende Attributmeng – können als Konstanten im Programmcode stehen. Beschreiben sie ihr Programm und erklären sie den Programmcode. Testen sie ihr Programm mit `airports.csv` und der Attributmeng `{continent, iso_country, iso_region, municipality}` und zeigen sie die Ausgabe in der zusätzlich abzugebenden PDF-Datei.

Wenn Sie zuviel Zeit zum Schreiben von funktionsfähigen Programmen brauchen (> 2h), geben Sie Pseudo-Kode ab. Dadurch bekommen Sie nicht die volle Punktzahl, aber es wird auch gewertet.

Fiktives Beispiel für die Ausgabe:

Die Attribut-Kombination `{continent, iso_country, iso_region, municipality}` ist kein Schlüssel für `airports.csv`.

Folgende Attributwertkombinationen verletzen die Schlüsselbedingung:

AF BJ BJ-AQ Cotonou

AF BJ BJ-ZO Bohicon

AF BJ BJ-AL Kandi

AF BJ BJ-AK Natitingou

AF BJ BJ-AK Porga

...

Insgesamt verletzen 5000 Attributwertkombinationen die Schlüsselbedingung.

Aufgabe 3.2:

6 Punkte

Schreiben sie ein Programm in einer beliebigen Programmiersprache, das zwei CSV-Dateien lesen kann und für jede CSV-Datei jeweils eine Menge von Attributen als Eingabe hat. Das Programm soll prüfen, ob zwischen den beiden Attributmengen und CSV-Dateien die Fremdschlüsselbedingung gilt. Geben Sie falls die Bedingung nicht gilt alle Tupel aus, welche die Bedingung verletzen. Wenn mehr als 5 verschiedene Tupel gefunden worden, geben Sie die weiteren nicht aus, sondern nur die Anzahl der Tupel, die die Fremdschlüsselbedingung verletzen. Die Eingaben für das Programm – Dateinamen und zu prüfende Attributmengen – können als Konstanten im Programmcode stehen. Beschreiben sie ihr Programm und erklären bzw. kommentieren sie den Programmcode. Testen sie ihr Programm mit folgenden Eingaben:

- Datei `airports.csv` mit Attributmenge `{iso_country}` und der Datei `country.csv` mit Attributmenge `{code}`
- Datei `airports.csv` mit Attributmenge `{iso_country}` und der Datei `country_onerowmissing.csv` mit Attributmenge `{code}`

Zeigen sie die Ausgabe in der zusätzlich abzugebenden PDF-Datei.

Wenn Sie zuviel Zeit zum Schreiben von funktionsfähigen Programmen brauchen (> 2h), geben Sie Pseudo-Kode ab. Dadurch bekommen Sie nicht die volle Punktzahl, aber es wird auch gewertet.

Fiktives Beispiel für die Ausgabe:

Die Attribut-Kombination `{iso_country}` in `airports.csv` ist kein Fremdschlüssel für `{code}` in `country.csv`. Für folgende Tupel aus `airports.csv` gibt es kein Gegenstück in `country.csv`:

```
2090 DGAA large_airport Kotoka International Airport 5.60518980026245 -0.166786000132561
205 AF GH GH-AA Accra yes DGAA ACC http://en.wikipedia.org/wiki/Kotoka_International_Airp
```

```
2091 DGLE medium_airport Tamale Airport 9.55718994140625 -0.863214015960693 553
AF GH GH-NP Tamale yes DGLE TML http://en.wikipedia.org/wiki/Tamale_Airport
```

```
30910 DGLN small_airport Navrongo Airport 10.8999996185303 -1.10000002384186
2221 AF GH GH-UE Navrongo no DGLN http://en.wikipedia.org/wiki/Navrongo_Airport
```

```
2092 DGLW small_airport Wa Airport 10.0826997756958 -2.50768995285034 1060 AF
GH GH-UW Wa no DGLW http://en.wikipedia.org/wiki/Wa_Airport
```

```
30911 DGLY small_airport Yendi Airport 9.42500019073486 -0.00472200009971857
2408 AF GH GH-NP Yendi no DGLY http://en.wikipedia.org/wiki/Yendi_Airport
```

...

Insgesamt verletzen 5000 Tupel die Fremdschlüsselbedingung.