

## Schriftelijke Opgaven Bestandsorganisatie en Databases

---

- Motiveer duidelijk en bondig alle antwoorden
  - Er zijn vier opgaven.
- 

### 1. Queries

De volgende relaties hebben betrekking op vlucht-informatie van een luchtvaartmaatschappij:

Flights (flno: integer, from: string, to: string, distance: integer, departs: time, arrives: time, tariff: integer)

Aircraft (aid: integer, aname: string, cruising range: integer)

Certified (eid: integer, aid: integer)

Employees (eid: integer, ename: string, salary: integer)

De Employees relatie bevat niet alleen informatie over piloten, maar ook over andere medewerkers van de luchtvaartmaatschappij. Alleen piloten echter zijn bevoegd om vluchten uit te voeren; iedere piloot is slechts bevoegd om op bepaalde toestellen te vliegen: dit is aangegeven in de relatie Certified.

Formuleer de volgende queries in SQL:

- a. Geef de *eid*'s van medewerkers met het hoogste salaris.
- b. Geef de *eid*'s van medewerkers met het op een na hoogste salaris.
- c. Geef de *aname*'s van alle toestellen (aircraft) waarvoor geldt dat piloten die daarop mogen vliegen meer dan 100 000 aan salaris ontvangen.
- d. Geef voor iedere piloot die op meer dan drie toestellen (aircraft) mag vliegen, de *eid* van die piloot en de maximum *cruising range* van de toestellen waarvoor hij/zij gecertificeerd is.
- e. Geef voor alle toestellen (aircraft) met een *cruising range* groter 1 000 de naam van dat toestel en het gemiddelde *salaris* van alle piloten die op dat toestel mogen vliegen.
- f. Geef de *aid*'s van alle toestellen (aircraft) waarvan de *cruising range* groter is dan de *distance* van Los Angeles naar Chicago.
- g. Bepaal het verschil van het gemiddelde *salaris* van de piloten en het gemiddelde *salaris* van alle medewerkers (inclusief de piloten).

## 2. Functionele afhankelijkheden en normaliseren

- a. Wat verstaan we onder een decompositie van een database relatie? Wanneer is het een lossless join-preserving decompositie? Geef een zo eenvoudig mogelijk voorbeeld van een decompositie die niet join-preserving is.
- b. Wat is het voordeel en wat is het nadeel van BCNF (Boyce Codd Normaal Vorm) in vergelijking met 3NF (Derde Normaal Vorm)?
- c. Gegeven een database relatie  $R$  met vier attributen,  $A, B, C, D$ . Hieronder staan vier sets FDs (Functional Dependencies). Doe voor elk van die vier sets het volgende: geef de kandidaat sleutels en als  $R$  niet in BCNF is geef dan een dependency preserving decompositie in een verzameling BCNF relaties (ga ervan uit dat er geen andere afhankelijkheden dan die in de betreffende set zijn):

- 1)  $C \twoheadrightarrow D, C \twoheadrightarrow A, B \twoheadrightarrow C$
- 2)  $B \twoheadrightarrow C, D \twoheadrightarrow A$
- 3)  $ABC \twoheadrightarrow D, D \twoheadrightarrow A$
- 4)  $A \twoheadrightarrow B, BC \twoheadrightarrow D, A \twoheadrightarrow C$

## 3. Query evaluaties

Gegeven twee relaties  $R$  en  $S$ , die als volgt zijn opgeslagen in een extern geheugen:

$R$  bestaat uit tupels van 300 bytes en beslaat 10 pagina's.

$S$  bestaat uit tupels van 500 bytes en beslaat 100 pagina's.

Iedere pagina (page) bestaat uit 1024 bytes.

Verder is gegeven dat  $R$  bestaat uit de attributen  $A, B$  en andere;  $S$  bestaat uit  $C, D$  en andere. De gecombineerde omvang van de attributen  $A$  en  $B$  is 200 bytes, die van  $C$  en  $D$  is 250 bytes.  $A$  is sleutel voor  $R$ ;  $C$  is sleutel voor  $S$ .

De volgende query moet worden uitgerekend:

$$\rho_{A,B,C,D}(R \bowtie_{A=C} S)$$

Gegeven is dat in de join iedere tupel van  $S$  met precies één tupel van  $R$  geassocieerd wordt.

Neem aan dat de projectieroutine gebaseerd is op sorteren en dat ongewenste attributen en eventuele duplicaten on-the-fly worden verwijderd. Neem verder aan dat er drie buffer pagina's beschikbaar zijn en dat de join routine van het type block nested loops is.

- a. Wat zijn de kosten als eerst de projectie wordt uitgevoerd en daarna de join?
- b. Wat zijn de kosten als eerst de join wordt uitgevoerd en daarna de projectie?
- c. Wat zijn de kosten als eerst de join wordt uitgevoerd en dan de projectie on-the-fly?
- d. Zou je andere antwoorden krijgen als het aantal beschikbare buffers verhoogd wordt naar 11?

#### 4. Uitdrukingskracht

Bestaan er zinvolle queries voor een relationele database die niet uit te drukken zijn in SQL, en ook niet in relationele algebra en niet in relationele calculus?

---

#### **BELANGRIJK!**

Alle opgaven graag in het *engels* inleveren. Het e-mail adres om de opgaven naartoe te sturen is: [nld@liacs.nl](mailto:nld@liacs.nl)