

# Data & Informatie Toets 2: Databaseontwerp en SQL

Sportcentrum, 16 mei 2014, 15.45 – 17.30 uur

## Antwoorden

De toets bevat 3 vragen op 3 pagina's die gezamenlijk 100 punten (en 5 bonuspunten) opleveren. Bij het tentamen mogen geen boeken, aantekeningen of elektronische apparaten worden gebruikt.

### Antwoord 1

Beoordeling: 16 punten voor de FDs (2 punt aftrek voor ieder fout antwoord W/O en voor iedere foute motivatie); 4 punt voor de MVDs.

Vraag	FD/MVD	W/O	Motivatie
a)	$BW \rightarrow A$	W	Volgt uit gegeven 2.
b)	$WA \rightarrow B$	O	Verschillende boeken kunnen in gelijke hoeveelheid in één W liggen.
c)	$K \rightarrow N$	W	Volgt uit aanname c. (Volgt niet uit 5: “N is een naam van K”).
d)	$K \rightarrow W$	O	Dit is gegeven a:
e)	$BM \rightarrow V$	W	Volgt uit gegeven 6.
f)	$B \rightarrow V$	O	In verschillende magazijnen kan de V voor B verschillend zijn.
g)	$N \rightarrow K$	O	Verschillende klanten kunnen gelijke namen hebben.
h)	$W \rightarrow K$	W	Dit is aanname b.
i)	$B \twoheadrightarrow AWK$	O	B is geen sleutel van BAWK en geen sleutel van BNVM. Bovendien is AWK niet onafhankelijk van NVM gegeven B (want c: $K \rightarrow N$ )
j)	$B \twoheadrightarrow VM$	W	Gegeven B is VM onafhankelijk van AWKN (de voorraad van boeken in het magazijn, is onafhankelijk van het aantal boeken in de winkelwagen, en de klant en zijn/haar naam.)

### Antwoord 2

Beoordelingsrichtlijn. 6 punt voor a), 9 punt voor b), 12 punt voor c), 3 punt voor d).

**Antwoord a)** Schema R heeft precies vier sleutels, namelijk: ABE, ACE, BDE, CDE. Elke FD in  $\mathcal{F}$  is een schending van de BCNF-conditie, omdat elke FD niet-triviaal is en het linkerlid ervan geen sleutel bevat (is geen superkey).

**Antwoord b)** Neem een FD uit  $\mathcal{F}$  die de BCNF-conditie schendt; bijvoorbeeld  $AB \rightarrow D$ . Transformeer R in  $R_1 = ABD$  (alle attributen van de FD) en  $R_2 = ABCE$  (alles zonder de attributen uit het rechterlid van de FD). Bij  $R_1$  en  $R_2$  horen respectievelijk de verzamelingen van functionele afhankelijkheden  $\mathcal{F}_1$  en  $\mathcal{F}_2$ , waarbij  $\mathcal{F}_1 = \{AB \rightarrow D, BD \rightarrow A\}$  en  $\mathcal{F}_2 = \{B \rightarrow C, C \rightarrow B\}$

*Toelichting:*

In het antwoord hierboven is  $\mathcal{F}_i$  een basis voor de verzameling van FDs uit  $\mathcal{F}^+$  waarin uitsluitend de attributen uit  $R_i$  voorkomen.

NB1:  $BD \rightarrow A$  zit niet in  $\mathcal{F}$  maar wel in  $\mathcal{F}^+$  (zie regel 12 in de tabel) en komt dus in  $\mathcal{F}_1$ .

NB2.  $CD \rightarrow A$  zit weliswaar niet in  $\mathcal{F}_1 \cup \mathcal{F}_2$  maar wel in  $(\mathcal{F}_1 \cup \mathcal{F}_2)^+$  en gaat dus niet verloren.

NB3.  $R_1$  staat in BCNF (de keys zijn AB en BD) en  $R_2$  bevat precies twee schendingen van de BCNF-conditie; samen dus minder dan de drie schendingen in  $R$ . (Dat het aantal schendingen in  $R_1$  en  $R_2$  samen minder is dan in  $R$  volgt ook uit het feit dat hier één iteratiestap van het BCNF-algoritme is toegepast.)

**Antwoord c)** We passen het BCNF-algoritme toe. De eerste stap is in de vorige vragen gedaan en levert bovengenoemde decompositie  $R_1 = ABD$  en  $R_2 = ABCE$  van  $R$  op. Schema  $R_1$  staat in BCNF. We bekijken nu  $R_2 = ABCE$ . In  $R_2$  zijn ABE en ACE de sleutels, dus  $B \rightarrow C$  en  $C \rightarrow B$  uit  $\mathcal{F}_2$  zijn beide een schending van de BCNF-conditie. We kiezen  $B \rightarrow C$  ter eliminatie. Dus transformeren we  $R_2$  in  $R_{2a} = BC$  en  $R_{2b} = ABE$ . Bij  $R_{2a}$  en  $R_{2b}$  horen respectievelijk de functionele afhankelijkheden  $\mathcal{F}_{2a}$  en  $\mathcal{F}_{2b}$ , waarbij  $\mathcal{F}_{2a} = \{B \rightarrow C, C \rightarrow B\}$  en  $\mathcal{F}_{2b} = \{ \}$ . Dus  $R_{2a}$  en  $R_{2b}$  staan beide in BCNF.

*Toelichting:*

Dus  $\{R_1, R_{2a}, R_{2b}\}$  is een gevraagde decompositie van  $R$  in BCNF. Omdat deze decompositie lossless is, geldt:  $R = R_1 \bowtie R_{2a} \bowtie R_{2b}$ , ofwel:  $ABCDE = ABD \bowtie BC \bowtie ABE$ . Het BCNF-algoritme levert *altijd* een lossless decompositie.

**Antwoord d)** Ja; alle functionele afhankelijkheden zijn behouden (zie ook NB2 hierboven).

*Alternatief antwoord:*

Wanneer je met een andere schending begint, kun je mogelijk een andere decompositie bereiken, en een andere uitslag voor het behoud van de functionele afhankelijkheden. Een andere compositie wordt verkregen door in het begin, bijvoorbeeld,  $B \rightarrow C$  te kiezen ter eliminatie: Transformeer  $R$  in  $R_1 = BC$  (alle attributen van de FD) en  $R_2 = ABDE$  (alles zonder de attributen uit het rechter lid van de FD), waarbij  $\mathcal{F}_1 = \{B \rightarrow C, C \rightarrow B\}$  een basis is voor de verzameling van FDs uit  $\mathcal{F}^+$  waarin attributen uit  $R_1$  voorkomen; en  $\mathcal{F}_2 = \{AB \rightarrow D, BD \rightarrow A\}$  een basis is voor de verzameling van FDs uit  $\mathcal{F}^+$  waarin attributen uit  $R_2$  voorkomen.

NB1.  $BD \rightarrow A$  zit niet in  $\mathcal{F}$  maar wel in  $\mathcal{F}^+$  (zie bijv. regel 12 in de tabel) en komt dus in  $\mathcal{F}_2$ .

NB2.  $CD \rightarrow A$  zit weliswaar niet in  $\mathcal{F}_1 \cup \mathcal{F}_2$  maar wel in  $(\mathcal{F}_1 \cup \mathcal{F}_2)^+$  en gaat dus niet verloren.

NB3.  $R_1$  staat in BCNF (de keys zijn B en C) en  $R_2$  bevat precies twee schendingen van de BCNF-conditie; samen dus minder dan de drie schendingen in  $R$ .

We bekijken nu  $R_2 = ABDE$ . In  $R_2$  zijn  $AB \rightarrow D$  en  $BD \rightarrow A$  beide een schending van de BCNF-conditie. We kiezen  $AB \rightarrow D$  ter eliminatie. Dus transformeren we  $R_2$  in  $R_{2a} = ABD$  en  $R_{2b} = ABE$ . Bij  $R_{2a}$  en  $R_{2b}$  horen respectievelijk de verzameling van functionele afhankelijkheden  $\mathcal{F}_{2a}$  en  $\mathcal{F}_{2b}$ , waarbij  $\mathcal{F}_{2a} = \{AB \rightarrow D, BD \rightarrow A\}$  en  $\mathcal{F}_{2b} = \{ \}$ .  $R_{2a}$  en  $R_{2b}$  staan beide in BCNF. Alle FDs blijven behouden.

## Antwoord 3

### *Antwoord a)*

```
SELECT s.name FROM Ship s, Class c
WHERE s.classname = c.name
AND c.type = 'bc'
AND s.launched < '1700-01-01';
```

### *Antwoord b)*

```
SELECT c.country, MIN(b.date), MAX(b.date)
FROM Class c, Ship s, Outcome o, Battle b
WHERE c.name = s.classname
AND s.name = o.shipname
AND o.battlename = b.name
GROUP BY c.country
HAVING COUNT(DISTINCT b.name) >= 10;
```

### *Antwoord c)*

Stap 3: vul de ontbrekende delen ('s is uit NL' en 's is betrokken bij b')

```
SELECT b.name FROM Battle b
WHERE EXISTS (
    SELECT * FROM Ship s
    WHERE EXISTS (
        SELECT * FROM Class c
        WHERE c.name = s.classname
        AND c.country = 'NL'
    )
    AND EXISTS (
        SELECT * FROM Outcome o
        WHERE o.shipname = s.name
        AND o.battlename = b.name
    )
);
```

Stap 4a: shunting once (4a, 4b, and 4c may be done in one go)

```
SELECT b.name FROM Battle b
WHERE EXISTS (
    SELECT * FROM Ship s, Class c
    WHERE c.name = s.classname
    AND c.country = 'NL'
    AND EXISTS (
        SELECT * FROM Outcome o
        WHERE o.shipname = s.name
        AND o.battlename = b.name
    )
);
```

Stap 4b: shunting twice

```
SELECT b.name FROM Battle b
WHERE EXISTS (
    SELECT * FROM Ship s, Class c, Outcome o
    WHERE c.name = s.classname
    AND c.country = 'NL'
    AND o.shipname = s.name
    AND o.battlename = b.name
);
```

Stap 4c: shunting sold!

```
SELECT DISTINCT b.name FROM Battle b, Ship s, Class c, Outcome o
WHERE c.name = s.classname
AND c.country = 'NL'
AND o.shipname = s.name
AND o.battlename = b.name;
```

Stap 5: o.battlename = b.name

```
SELECT DISTINCT o.battlename FROM Battle b, Ship s, Class c, Outcome o
WHERE c.name = s.classname
AND c.country = 'NL'
AND o.shipname = s.name
AND o.battlename = b.name;
```

Stap 6: shunting Battle back

```
SELECT DISTINCT o.battlename FROM Ship s, Class c, Outcome o
WHERE c.name = s.classname
AND c.country = 'NL'
AND o.shipname = s.name
AND EXISTS (
    SELECT * FROM Battle b
    WHERE o.battlename = b.name
);
```

Stap7: Foreign key rule

```
SELECT DISTINCT o.battlename FROM Ship s, Class c, Outcome o
WHERE c.name = s.classname
AND c.country = 'NL'
AND o.shipname = s.name;
```

**Antwoord d)** (bonus question)

```
SELECT DISTINCT c.name FROM Class c, Battle b
WHERE NOT EXISTS (
    SELECT * FROM Ship s, Outcome o
    WHERE s.classname = c.name
    AND o.shipname = s.name
    AND o.battlename = b.name
    AND o.result = 'sunk'
    AND NOT EXISTS (
        SELECT * FROM Outcome o2
        WHERE o2.shipname = s.name
        AND o2.result <> 'sunk'
    )
);
```