

# Tentamen Databases voor **ica**

16 april 2015

17:00 - 20:00, Educatorium-beta

- **Lees eerst onderstaande aanwijzingen.**
- Vermeld op elk vel je naam en studentnummer. Indien één van deze zaken ontbreekt, wordt het vel niet nagekeken.
- Vul alle antwoorden in op deze tentamenbundel. Het bijgeleverde tentamenpapier is uitsluitend bedoeld als kladpapier.
- Bij elke meerkeuze vraag is één antwoord correct (of optimaal). Omcirkel de letter voor dit antwoord. Als je je keuze wilt corrigeren, kras de letter dan door en omcirkel alsnog de juiste letter. Als je veelvuldig je antwoord hebt gecorrigeerd, vermeld dan de letter van het juiste antwoord in de kantlijn links van de vraag en kras de andere letters door.
- Achter het nummer van elke vraag staat hoeveel punten deze vraag waard is.
- Toon bij het inleveren je collegekaart.
- Schrijf en formuleer duidelijk.
- Je mag een A4 met aantekeningen raadplegen.
- Het tentamen duurt 3 uur.
- Gebruikte afko's:
  - RDB = Relationale database
  - 2PL = Two-phase locking
  - 2PC = Two-phase commit
  - BCNF = Boyce-Codd normaalvorm
  - COORD = Coordinator
  - DP = dependency preserving
  - FD = functional dependency
  - RA = Relationale algebra
  - SQL = SQL (Structured Query Language)
- **Succes!**

Studentnr:	Achternaam:
------------	-------------

## Algemeen (16 punten)

*De vragen onder de noemer Algemeen worden bij een fout antwoord met een negatieve score (het aantal vermelde punten) beoordeeld. Bij ontbrekend antwoord is de score 0. Reden hiervoor is de verwachte scoringskans van 50% bij louter gokken.*

*De gezamenlijke score over de vragen onder de noemer Algemeen (dus vraag 1-8) kan niet negatief worden.*

### Vraag 1 [2]

Als een ER-diagram door verschillende competente ontwerpers afgebeeld wordt naar een RDB-schema, zal dit altijd tot hetzelfde RDB-schema leiden.

**A:** juist

**B:** onjuist

### Vraag 2 [2]

Als een minimal cover een FD:  $A \rightarrow BC$  bevat, dan is de decompositie op (AB) en (AC) een correct alternatief voor (ABC) ten behoeve van het genereren van een verliesvrije, DP 3NF-decompositie.

**A:** juist

**B:** onjuist

### Vraag 3 [2]

De herschrijfregel  $X \rightarrow YZ, Y \rightarrow UV \Rightarrow X \rightarrow V$  is:

**A:** juist

**B:** onjuist

### Vraag 4 [2]

Multivalued dependencies komen in de praktijk zelden voor.

**A:** juist

**B:** onjuist

### Vraag 5 [2]

Het terminatieprotocol van 2PC leidt altijd tot Abort.

**A:** juist

**B:** onjuist

### Vraag 6 [2]

Het terminatieprotocol van 2PC kan tot inconsistentie leiden bij een netwerkpartitie.

**A:** juist

**B:** onjuist

<b>Studentnr:</b>	<b>Achternaam:</b>
-------------------	--------------------

**Vraag 7** [2]

Twee verschillende SQL-queries kunnen (op hetzelfde DBMS en bij dezelfde inhoud van de database) leiden tot dezelfde methode voor de berekening van het resultaat van de query.

**A:** juist

**B:** onjuist

**Vraag 8** [2]

Twee dezelfde SQL-queries kunnen (op hetzelfde DBMS en bij verschillende inhoud van de database) leiden tot verschillende methodes voor de berekening van het resultaat van de query.

**A:** juist

**B:** onjuist

## Dependencies (16 punten)

*De vragen over Dependencies worden bij een fout antwoord met een negatieve score beoordeeld. Bij ontbrekend antwoord is de score 0. De gezamenlijke score over de vragen onder de noemer Dependencies kan niet negatief worden.*

Een zonovertogen universiteit in het midden des lands gebruikt als studievoortgangssysteem het onvolprezen SIRISO. Elke gelijkenis met OSIRIS berust niet op toeval. U wordt geacht uw domeinkennis betreffende OSIRIS te gebruiken. Vakken kennen in het algemeen meerdere jaargangen. We onderscheiden ten behoeve van de opslag van studieresultaten onder meer de volgende attributen:

- *studentnummer*: het studentnummer.
- *studentnaam*: de achternaam en voorletters van de student.
- *gebdat*: de geboortedatum van de student.
- *vakcode*: de welbekende SIRISO vakcode, vergelijkbaar met onze OSIRIS vakcode.
- *vaknaam*: de naam van een vak.
- *collegejaar*: het collegejaar waarin een jaargang van een vak gegeven wordt (het vak INFODB dat u volgt heeft collegejaar 2014).
- *periode*: de periode waarin een vak gegeven wordt: 1, 2, 3, of 4.
- *toetsdatum*: de datum waarop een toets afgenomen wordt.
- *toetsuitslag*: de bijbehorende toetsuitslag.
- *docentcode*: de unieke identificatiecode van een docent.

**Vraag 9** [1]

De FD *studentnummer*  $\rightarrow$  *studentnaam*, *gebdat*

**A:** geldt

**B:** geldt niet

<b>Studentnr:</b>	<b>Achternaam:</b>
-------------------	--------------------

**Vraag 10** [1]

De FD *studentnummer*  $\rightarrow$  *studentnummer, studentnaam, gebdat*

**A:** geldt

**B:** geldt niet

**Vraag 11** [1]

De FD *studentnaam, gebdat*  $\rightarrow$  *studentnummer*

**A:** geldt

**B:** geldt niet

**Vraag 12** [1]

De FD *studentnummer, studentnaam, gebdat*  $\rightarrow$  *studentnummer*

**A:** geldt

**B:** geldt niet

**Vraag 13** [1]

De FD *vaknaam*  $\rightarrow$  *vakcode*

**A:** geldt

**B:** geldt niet

**Vraag 14** [1]

De FD *studentnummer, vakcode*  $\rightarrow$  *toetsuitslag*

**A:** geldt

**B:** geldt niet

**Vraag 15** [1]

De FD *studentnummer, vakcode, collegejaar*  $\rightarrow$  *toetsuitslag*

**A:** geldt

**B:** geldt niet

**Vraag 16** [1]

De FD *studentnummer, vakcode, collegejaar, periode*  $\rightarrow$  *toetsuitslag*

**A:** geldt

**B:** geldt niet

**Vraag 17** [1]

De FD *studentnummer, vakcode, collegejaar, periode, toetsdatum*  $\rightarrow$  *toetsuitslag*

**A:** geldt

**B:** geldt niet

<b>Studentnr:</b>	<b>Achternaam:</b>
-------------------	--------------------

**Vraag 18** [1]De FD *studentnummer, vakcode, collegejaar, toetsdatum*  $\rightarrow$  *toetsuitslag***A:** geldt**B:** geldt niet**Vraag 19** [1]De FD *studentnummer, vakcode, toetsdatum*  $\rightarrow$  *toetsuitslag***A:** geldt**B:** geldt niet**Vraag 20** [1]De FD *studentnummer, docentcode, toetsdatum*  $\rightarrow$  *toetsuitslag***A:** geldt**B:** geldt niet**Vraag 21** [2]We beperken ons nu even tot de projectie op (*studentnummer, vakcode, toetsuitslag*).**A:** De MVD *studentnummer*  $\rightarrow$  *vakcode* geldt**B:** De MVD *studentnummer*  $\rightarrow$  *toetsuitslag* geldt**C:** De MVD's *studentnummer*  $\rightarrow$  *vakcode* en *studentnummer*  $\rightarrow$  *toetsuitslag* gelden beide**D:** De MVD's *studentnummer*  $\rightarrow$  *vakcode* en *studentnummer*  $\rightarrow$  *toetsuitslag* gelden geen van beide

Studentnr:

Achternaam:

## Minimal covers en 3NF (16 punten)

We hebben een schema  $R(ABCDEFG)$  en een FDset  $\{AB \rightarrow D, B \rightarrow C, BD \rightarrow E, C \rightarrow EF, BF \rightarrow D\}$ .

We gaan om te beginnen een minimal cover berekenen.

### Vraag 22 [6]

Vul in onderstaande ruimte in welke FDs veranderd worden door de linkerkantreductie en hoe deze er daarna uitzien.

### Vraag 23 [6]

Vul in onderstaande ruimte in welke FDs verwijderd worden in verband met overtolligheid.

### Vraag 24 [4]

Vul in onderstaande ruimte in welke verliesvrije 3NF, DP decompositie uit de minimal cover gegenereerd kan worden. Geef daartoe de relatieschema's en voor elk schema de candidate keys.

Studentnr:	Achternaam:
------------	-------------

## BCNF (12 punten)

*De vragen onder de noemer BCNF worden bij een fout antwoord met een negatieve score beoordeeld. Bij ontbrekend antwoord is de score 0. De gezamenlijke score over de vragen onder de noemer BCNF kan niet negatief worden.*

We hebben een schema  $R(ABCDE)$  en een FDset  $\{E \rightarrow D, D \rightarrow BC, A \rightarrow C\}$ .

### Vraag 25 [2]

Is de decompositie (EC), (DBC), (AD) verliesvrij?

**A:** ja

**B:** nee

### Vraag 26 [2]

Is de decompositie (EC), (DBC), (AD) in BCNF?

**A:** ja

**B:** nee

### Vraag 27 [2]

Is de decompositie (EC), (DBC), (AD) DP?

**A:** ja

**B:** nee

### Vraag 28 [2]

Is de decompositie (ED), (DBC), (AE) verliesvrij?

**A:** ja

**B:** nee

### Vraag 29 [2]

Is de decompositie (ED), (DBC), (AE) in BCNF?

**A:** ja

**B:** nee

### Vraag 30 [2]

Is de decompositie (ED), (DBC), (AE) DP?

**A:** ja

**B:** nee

Studentnr:	Achternaam:
------------	-------------

## Queries (16 punten)

We kunnen maar geen genoeg krijgen van IMDB. We hebben het volgende database schema.

```
Film ( filmid, titel, maatschappij, jaar, regisseur )
Acteur ( actid, naam, gebdatum, geslacht )
Speeltin ( filmid, actid )
```

Geslacht kent de waarden M en V; NULL-values komen niet voor.

Hieronder zie je vier queries en een groot aantal expressies in de RA en SQL. We korten de tabelnamen af tot hun eerste letter. De relationele division duiden we aan met een %.

Q1: In welke films (filmid volstaat) spelen uitsluitend vrouwelijke acteurs mee

Q2: Geef de titels van de films waarvoor geldt dat er in één jaar minstens twee verschillende films met deze titel zijn verschenen

Q3: Geef de acteurs (actid) die in alle films die regisseur Scorsese in de jaren 1990 en 1991 gemaakt heeft meespelen

Q4: Welke acteur (actid) heeft in de meeste films gespeeld

E1:  $\pi_{filmid}(\sigma_{geslacht='V'}(A) \bowtie S)$

E2:  $\pi_{filmid}(\sigma_{geslacht \neq 'M'}(A) \bowtie S)$

E3:  $\pi_{filmid}(F) - \pi_{filmid}(\sigma_{geslacht='M'}(A) \bowtie S)$

E4:  $\pi_{filmid,actid}(S) \% \pi_{actid}(\sigma_{geslacht='V'}(A))$

E5:  $\pi_{titel}(F \bowtie F)$

E6:  $\pi_{actid}(S \% \pi_{filmid}(\sigma_{regisseur='Scorsese' \wedge jaar \geq 1990 \wedge jaar \leq 1991}(F)))$

E7:  $\pi_{actid,filmid}(S) \% \pi_{filmid}(\sigma_{regisseur='Scorsese' \wedge jaar \geq 1990 \wedge jaar \leq 1991}(F))$

E8:

```
SELECT filmid FROM F
```

```
WHERE NOT EXISTS (
```

```
    SELECT * FROM S, A
```

```
    WHERE geslacht = 'M' AND S.filmid = F.filmid
```

```
)
```



E9:

```
SELECT filmid FROM F
WHERE filmid NOT IN (
    SELECT filmid FROM S, A
    WHERE geslacht = 'M' AND S.actid = A.actid
)
```

E10:

```
SELECT F1.titel FROM F F1, F F2
WHERE F1.titel = F2.titel
AND F1.jaar = F2.jaar
AND F1.filmid <> F2.filmid
```

E11:

```
SELECT naam FROM A
WHERE NOT EXISTS (
    SELECT * FROM F
    WHERE regisseur = 'Scorsese'
    AND jaar >= 1990 AND jaar <= 1991
    AND NOT EXISTS (
        SELECT * FROM S
        WHERE S.filmid = F.filmid
        AND S.actid = A.actid
    ))
```

E12:

```
SELECT naam FROM A
WHERE NOT EXISTS (
    SELECT * FROM S
    WHERE S.actid = A.actid
    AND NOT EXISTS (
        SELECT * FROM F
        WHERE regisseur = 'Scorsese'
        AND jaar >= 1990 AND jaar <= 1991
        WHERE S.filmid = F.filmid
        AND S.actid = A.actid
    ))
```

Studentnr:

Achternaam:

E13:

```
SELECT actid FROM S
GROUP BY actid
HAVING MAX(COUNT(*))
```

E14:

```
SELECT actid FROM S
GROUP BY actid
HAVING SELECT MAX(COUNT(*))
      FROM S
      GROUP BY actid
)
```

E15:

```
SELECT actid FROM S
GROUP BY actid
HAVING COUNT(*) >= ALL (
      SELECT SUM(*) FROM S
      GROUP BY actid
)
```

E16:

```
SELECT actid FROM S
GROUP BY actid
HAVING COUNT(*) >= ALL (
      SELECT COUNT(*) FROM S
      GROUP BY actid
)
```

**Vraag 31** [4]

Geef hieronder aan welke expressies corresponderen met Q1.

--

Studentnr:	Achternaam:
------------	-------------

**Vraag 32** [4]

Geef hieronder aan welke expressies corresponderen met Q2.

**Vraag 33** [4]

Geef hieronder aan welke expressies corresponderen met Q3.

**Vraag 34** [4]

Geef hieronder aan welke expressies corresponderen met Q4.

## Concurrency (14 punten)

We beschouwen de volgende twee schedules.

$S_1$					$S_2$				
T1	T2	T3	T4	T5	T1	T2	T3	T4	T5
w(x)					w(y)				
	w(z)		w(y)			r(y)		w(x)	
	w(y)					w(z)			w(y)
r(z)		w(y)						r(z)	
				r(y)	r(x)		w(x)		

**Vraag 35** [2]

Geef een opsomming van *alle* pijlen van de precedentiegraaf  $G(S_1)$ ). Gebruik voor een pijl dit format:  $T_i \rightarrow T_j$ . Sorteer de lijst oplopend voor  $T_i$ .

<b>Studentnr:</b>	<b>Achternaam:</b>
-------------------	--------------------

**Vraag 36** [2]

Is  $S_1$  serializeerbaar? Zo ja, geef een equivalente seriële schedule van  $S_1$ .

**Vraag 37** [2]

Geef een opsomming van *alle* pijlen van de precedentiegraaf  $G(S_2)$ , wederom gesorteerd.

**Vraag 38** [2]

Is  $S_2$  serializeerbaar? Zo ja, geef een equivalente seriële schedule van  $S_2$ .

**Vraag 39** [6]

Eén van de transacties hierboven vertoont geen 2PL-gedrag. Welke is dat?

**A:**  $T_1$  van  $S_1$

**B:**  $T_2$  van  $S_1$

**C:**  $T_3$  van  $S_1$

**D:**  $T_4$  van  $S_1$

**E:**  $T_5$  van  $S_1$

**F:**  $T_1$  van  $S_2$

**G:**  $T_2$  van  $S_2$

**H:**  $T_3$  van  $S_2$

**I:**  $T_4$  van  $S_2$

**J:**  $T_5$  van  $S_2$

Studentnr:	Achternaam:
------------	-------------

## Recovery (12 punten)

We beschouwen nonquiescent recovery met gecombineerde **UNDO** en **REDO** logging. Hieronder vind je een log met before en after images.

```
<START T1>
<T1, B, 10, 11>
<COMMIT T1>
<START T3>
<START T2>
<T2, A, 5, 6>
<T3, C, 10, 11>
<START CKPT (T2, T3)>
<T2, D, 15, 16>
<START T4>
<START T5>
<T4, F, 25, 26>
<T5, H, 12, 13>
<COMMIT T2>
<COMMIT T3>
<END CKPT>
<COMMIT T4>
<T5, G, 30, 31>
```

Stel dat een crash optreedt direct na `<COMMIT T2>` . (Het resterende gedeelte van de log wordt dan niet geschreven.)

### Vraag 40 [2]

Welk gedeelte van de log file wordt gescand?

### Vraag 41 [2]

Welke transacties worden undone?

<b>Studentnr:</b>	<b>Achternaam:</b>
-------------------	--------------------

**Vraag 42** [2]

Welke transacties worden redone?

Stel nu dat een crash optreedt direct na  $\langle T5, G, 30, 31 \rangle$  .

**Vraag 43** [2]

Welk gedeelte van de log file wordt gescand?

**Vraag 44** [2]

Welke transacties worden undone?

**Vraag 45** [2]

Welke transacties worden redone?

Studentnr:	Achternaam:
------------	-------------

## Query processing (10 punten)

*De vragen onder de noemer Query processing worden bij een fout antwoord met een negatieve score beoordeeld. Bij ontbrekend antwoord is de score 0. De gezamenlijke score over de vragen onder de noemer Query processing kan niet negatief worden.*

We gaan uit van twee relatieschema's  $R$  en  $S$  die één attribuut  $A$  gemeenschappelijk hebben. We definiëren een algebraïsche operator  $\bowtie$  (semi-join) als volgt:

$R \bowtie S$  bevat de tuples  $r$  in  $R$  waarvoor geldt dat er een tuple  $s$  in  $S$  bestaat met  $r.A = s.A$ .

Met andere woorden, de semi-join verwijdert uit  $R$  de tuples die de natural join met  $S$  niet zullen "overleven".

### Vraag 46 [2]

Is de algebraïsche equivalentie  $R \bowtie S = S \bowtie R$  geldig?

A: ja

B: nee

### Vraag 47 [2]

Is de algebraïsche equivalentie  $R \bowtie S = (R \bowtie S) \bowtie S$  geldig?

A: ja

B: nee

### Vraag 48 [2]

Is de algebraïsche equivalentie  $R \bowtie S = (R \bowtie (\pi_A S)) \bowtie S$  geldig?

A: ja

B: nee

### Vraag 49 [2]

Is de algebraïsche equivalentie  $R \bowtie S = (R \bowtie S) \bowtie (S \bowtie R)$  geldig?

A: ja

B: nee

### Vraag 50 [2]

Is de algebraïsche equivalentie  $R \bowtie S = (R \bowtie (S \bowtie R)) \bowtie S$  geldig?

A: ja

B: nee

**Einde**