

Tentamen Informatica 6, 2IJ60, 18-03-2005

Dit tentamen bestaat uit 6 opgaven. Bij elke opgave staat aangegeven hoeveel punten te behalen zijn. Één punt ontvang je cadeau voor de aanwezigheid. Het eindresultaat wordt verkregen door de punten voor alle opgaven te sommeren en door 3 te delen met afronding naar boven of naar beneden.

Geef de oplossingen kort, bondig en duidelijk weer.

Dit is een open boek tentamen (je mag het Silberschatz boek en aantekeningen gebruiken, maar geen laptop).

Dit tentamen bestaat uit twee delen: een “pre”-tentamen en een “hoofd”-tentamen.

Het **pretentamen** bestaat uit drie vragen, waarvan **ten minste twee correct** beantwoord dienen te worden, voordat het hoofdtentamen wordt nagekeken. Als je geen of slechts één van de pretentamen-vragen correct beantwoordt, wordt je cijfer 1 onafhankelijk van de oplossingen die je voor de hoofdtentamen-vragen geeft. Wees dus zorgvuldig in het beantwoorden van het eerste deel. Daarna kun je verder gaan met de overige vragen.

In de query opgaven wordt een (deeltje van een) universiteitsbibliotheek-database gebruikt met de volgende tabellen en attributen:

De bibliotheekdatabase

Het onderstaande relationele schema stelt een (deel van een) sterk vereenvoudigde universiteitsbibliotheekdatabase voor. Het wordt gebruikt in alle onderstaande vragen.

In de query opgaven wordt een (deeltje van een) universiteitsbibliotheekdatabase gebruikt met de volgende tabellen en attributen:

boek : { ISBN, titel, uitgever, jaar }

auteur : { ISBN, voorletters, naam }

exemplaar : { barcode, ISBN, faculteit, exjaar, aanwezig }

reservering : { naam, faculteit, ISBN, datum, geannuleerd }

uitlening : { naam, faculteit, barcode, van, tot }

Korte beschrijving:

Elk boek heeft een uniek ISBN nummer. Het heeft een titel, uitgever, jaar van uitgifte en een aantal auteurs waarvan we de voorletters (samen als 1 string) en de achternaam bijhouden. Verschillende edities van eenzelfde boek hebben een verschillend ISBN nummer en zijn dus voor ons verschillende boeken. De database bevat alle auteurs van in de database aanwezige boeken, en geen auteurs die niet (met ISBN nummer) overeenkomen met een boek.

De universiteit heeft van elk boek 1 of meer exemplaren. Elk exemplaar is in een bepaald jaar (exjaar) aangekocht, heeft een sticker met een unieke barcode en is toegewezen aan een faculteit. (Boeken kunnen worden uitgeleend aan mensen van andere faculteiten.) Een exemplaar kan afwezig zijn omdat het nog niet geleverd is, wordt hersteld, of omdat het is uitgeleend.

Personen worden geïdentificeerd door hun achternaam en faculteit. Ze kunnen een boek reserveren op een bepaalde datum. (Het type van “datum” is zodanig dat data die niet NULL zijn met elkaar vergeleken kunnen worden, bijvoorbeeld met $<$ en \leq .) Personen kunnen een reservering op elk willekeurig moment weer annuleren. De datum blijft dan die van de

reservering, dus de datum waarop de reservering is geannuleerd wordt niet bijgehouden. (Het “geannuleerd” attribuut heeft altijd een waarde “ja” of “nee”.)

Ze kunnen een exemplaar van een boek lenen op een bepaalde “van” datum. Ze zijn dan “lener”. Een eventuele reservering van dat boek (op hun naam) wordt dan automatisch geannuleerd (de record van de reservering blijft in de database staan maar “geannuleerd” wordt “ja”). De “tot” datum blijft NULL tot het boek is teruggebracht. De “tot” datum is minstens 1 dag later dan de “van” datum. Wanneer een boek is teruggebracht wordt de “uitlening” in de database bewaard en wanneer een reservering wordt geannuleerd (bijvoorbeeld omdat het boek wordt geleend maar het kan ook om andere redenen zijn) wordt die reservering ook bewaard.

De bibliotheek bestaat al een hele tijd. Bijgevolg heeft elke faculteit wel wat (exemplaren van) boeken, heeft er uit elke faculteit wel eens iemand een boek gereserveerd en iemand een exemplaar geleend, en is er van elke faculteit wel eens een boek gereserveerd en wel eens een (exemplaar van) een boek geleend.

Let zeker op de volgende “valstrikken”:

- De attributen “naam” en “faculteit” worden in deze database in twee verschillende betekenissen gebruikt. Zorg dat je die niet per ongeluk door elkaar gebruikt. Let ook op voor de verschillen tussen “boeken” en “exemplaren”.
- In een vraag schrijven we misschien wel eens dat een “boek” wordt geleend, maar we bedoelen dan dat een “exemplaar” (eender welk!) van dat “boek” wordt geleend. Twee uitleningen van eenzelfde boek kunnen uitleningen van verschillende exemplaren zijn. Zorg dat je altijd goed in de gaten hebt of er een “boek” als generiek object wordt bedoeld of een “exemplaar” als concreet item dat kan worden geleend.

Pre-tentamendeel

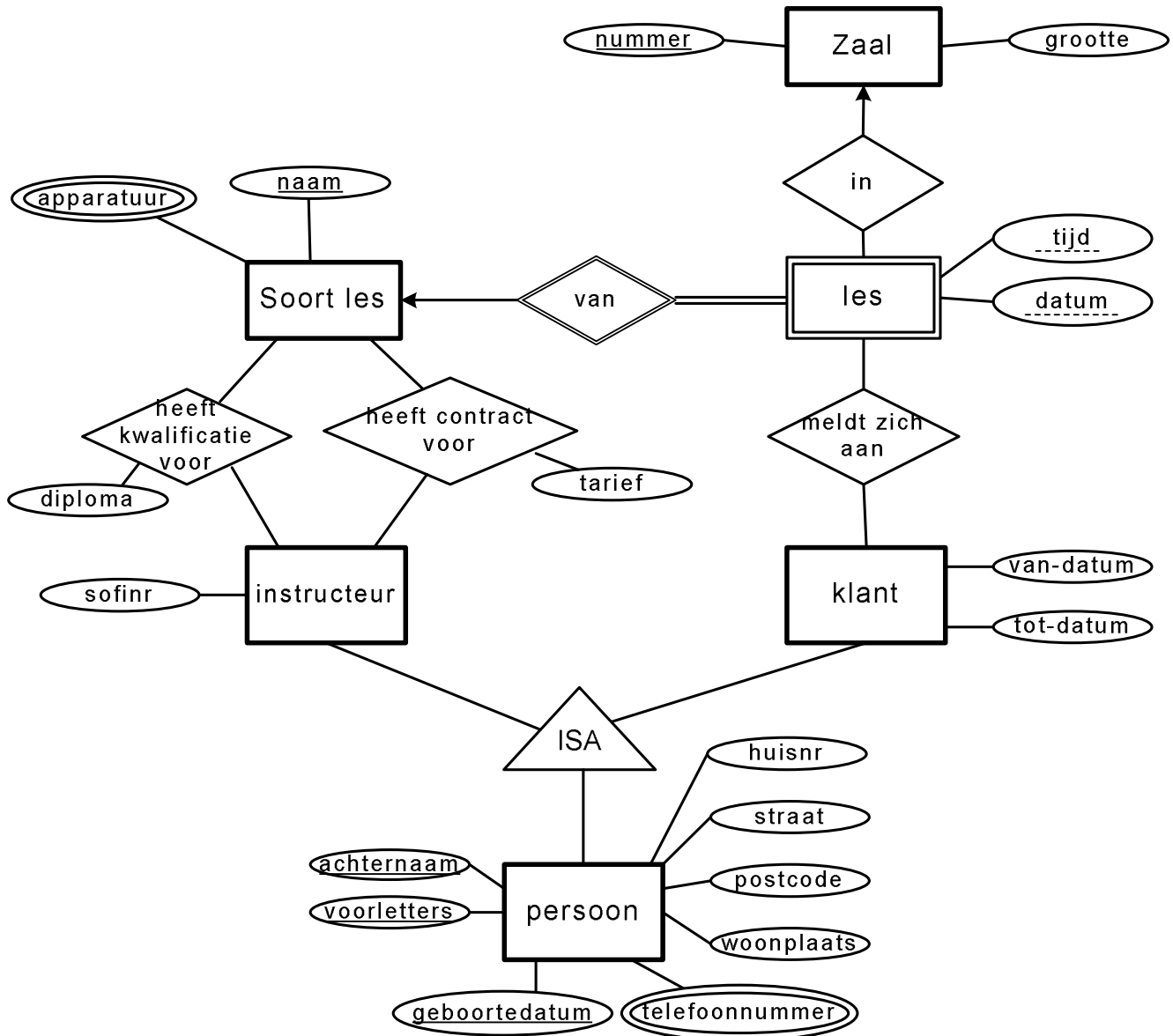
Deze vragen worden helemaal correct of helemaal fout gerekend.

1. (2 punten) Bestudeer grondig de beschrijving van de bibliotheek-database. Beschrijf dan de volgende vraag in de relationele algebra:
“Geef de naam en voorletters van de auteurs van boeken waarvan minstens één exemplaar ooit gereserveerd was (ook kan het op dit moment gereserveerd zijn).”
2. (2 punten) Beschrijf de volgende vraag over de bibliotheek-database in SQL:
“Geef de naam en faculteit van personen die ooit een exemplaar van een boek geleend hebben bij hun eigen faculteit”.
3. (2 punten) Beschouw een relationeel schema R met attributen $\{A, B, C, D\}$, elk met een groot domein zoals integers of strings, en de verzameling functionele afhankelijkheden $F = \{AB \rightarrow C, C \rightarrow ABD\}$. Is dit schema in BCNF? Leg uit waarom dit wel of niet zo is. (Zonder uitleg wordt deze vraag fout gerekend.)

Einde pre-tentamendeel. Wanneer je niet minstens twee van deze drie vragen goed hebt wordt de rest van het tentamen niet nagekeken. Je kunt voor jezelf bepalen of je nog verder wilt gaan!

Hoofd-tentamendeel

Je hoeft de vragen niet te beantwoorden in de onderstaande volgorde. Begin met wat je het beste kan!



1. (2 punten)

Geef een vertaling van het bovenstaande E-R model naar een relationeel model, d.w.z. geef aan hoe deze informatie met behulp van relaties (tabellen) kan worden opgeslagen. Vermeld per relatie de attributen en sleutels.

2. (5 punten)

Geef een ER-model (ER-diagram) voor de data in het informatiesysteem. Geef aan welke voorwaarden/beperkingen uit deze opgave niet in het model zichtbaar zijn en waarom niet.

Het organisatiecomité van een conferentie wil met behulp van een database alle informatie betreffende de conferentie registreren. De informatie betreft de leden van het organisatiecomité, de leden van het programmacomité, auteurs van ingezonden artikelen en overige conferentiedeelnemers. Van ieder van die personen zijn de namen, voornamen, titels, bedrijf, werkadres, telefoonnummer en faxnummer bekend. Van elk lid van het programmacomité en het organisatiecomité is bovendien bekend of dat een gewoon lid of een chairman is. Ingezonden artikelen worden in de database bewaard. Van elk artikel is een titel, resumé, auteurs en een aantal sleutelwoorden bekend. Ook programmacomitéleden mogen artikelen voor de conferentie inzenden. Elk artikel wordt aan drie programmacomitéleden toegewezen voor een recensie. Niemand mag zijn eigen artikel recenseren. De recensie van elk programmacomitélid bevat een cijfer voor het artikel, een confidentieniveau van de recensent (een cijfer van 1 t/m 3) en een commentaar. Op basis van de recensies beslist het programmacomité welke artikelen geaccepteerd worden en welke niet, hetgeen ook in de database wordt opgenomen. Het programmacomité maakt verder een rooster van de conferentie. In het rooster staan er voor elk van de geaccepteerde artikelen de datum, het tijdstip en de presentatornaam. De presentator moet een van de auteurs van het artikel zijn.

3. (2 punten per query)

Druk de volgende vragen over de bibliotheek uit in één van de bestudeerde querytalen (relationele algebra, tupel calculus of SQL) naar keuze. Hint: sommige vragen zijn in sommige talen gemakkelijker dan in andere. Je mag het jezelf dus gemakkelijk maken door een verstandige keuze van querytalen. Je mag bij elke vraag een andere querytaal gebruiken als je wilt, maar je mag ze ook allemaal in eenzelfde taal uitwerken.

- a) "Geef de uitgevers die in de database bekend zijn en van wie de bibliotheek na 1994 geen enkel exemplaar van zijn boeken heeft gekocht."
- b) "Geef per (elke bekende in de database) uitgever het aantal uitleningen van zijn boeken in 2004."
- c) "Geef alle uitgevers wiens boeken bij elke faculteitsbibliotheek aanwezig zijn."

4. (2 punten per query)

Leg in normaal klinkend Nederlands uit wat er in de volgende queries gevraagd wordt:

- a)

```
SELECT u.naam, u.faculteit
FROM uitlening AS u, uitlening AS uu, exemplar AS e, exemplar AS ee
WHERE u.barcode = e.barcode AND uu.barcode = ee.barcode AND
e.ISBN = ee.ISBN AND
u.naam = uu.naam AND u.faculteit = uu.faculteit AND
(u.van <> uu.van OR u.barcode <> uu.barcode)
```

- b) $\pi_{\text{naam, voorletters(auteur)}} - \pi_{\text{naam, voorletters(auteur)} \bowtie \pi_{\text{barcode(uitlening)} \bowtie \pi_{\text{barcode, ISBN (exemplaar)}}$

5. (6 punten)

Veronderstel dat we een database willen construeren uit een verzameling data items $\{A, B, C, D, E, G\}$ (die later attributen van tabellen zullen worden) en een verzameling F van functionele afhankelijkheden: $F = \{ABC \rightarrow DE, AD \rightarrow E, AB \rightarrow C\}$.

- a) Bereken een canonieke afdekking (canonical cover) F_C voor deze verzameling afhankelijkheden. (als je overbodige attributen schrapt, beargumenteer met het gebruik van Armstrong's axioma's en union, decompositie en pseudotransitivity regels dat deze attributen inderdaad overbodig zijn. Zonder argumentatie wordt het antwoord fout gerekend.)
- b) Begin met een tabel R met als attributen $ABCDEG$ en decomposeer deze met behoud van afhankelijkheden en verliesloos (dependency-preserving en lossless-join) tot een database in 3NF.
- c) Is deze decompositie ook in BCNF? Beargumenteer je antwoord. (Zonder uitleg wordt het antwoord fout gerekend.)
- d) Er zijn twee verschillende decomposities van R gegeven. Decompositie 1: in tabellen (A,B,C,D) en (B,E,G) , en decompositie 2: in tabellen (A,B,C,D,E) en (A,B,G) . Geef voor elke decompositie aan of deze decompositie lossless-join is en beargumenteer je antwoord. (Zonder argumentatie wordt het antwoord fout gerekend.)