

Helsingin yliopisto,
Tietojenkäsittelytieteen laitos
Tietokantojen perusteet, kurssikoe 29.2.2012
(vastauksia)

Litteenä on tiivistelmä SQL-syntaksista

Kirjoita jokaiseen erilliseen vastauspaperiin kurssin nimi, tenttipäivä, oma nimesi (selkeästi), opiskelijanumerosi ja nimikirjoituksesi

1.

Seuraavissa kaaviossa $X \rightarrow Y$ tarkoittaa, että X on tauluun Y viittaava viiteavain. $(X,Z) \rightarrow Y$ ilmaisee, että X ja Z yhdessä muodostavat viiteavaimeen.

- a) Selitä lyhyesti (parilla virkkeellä) mitä tarkoittaa käsite *viiteavain* (*foreign key*) relaatiomallissa.

Viiteavain X taulusta A tauluun B on sarake tai sarakkeyhdistelmää, jonka arvot esiintyvät taulussa B avainsarakkeen (tai yhdistelmän) K arvoina ja näin muodostavat kytkennän taulujen A ja B rivien välille.

Arvostelusta:

Jos kytkeminen tuotiin esiin jotakuinkin ymmärrettävällä tavalla, niin vähintään 1 piste.
Jos väitettiin, että viiteavain on avain, niin -1 p.

- b) Tarkastellaan relaatioita $A(\underline{B}, C, D, E)$ ja $F(\underline{G}, H, E, C)$. Tässä ei ole merkitty viiteavaimia. Voiko sarake G olla tauluun A viittaava viiteavain? Perustele vastauksesi lyhyesti.

Ei voi sillä viiteavaimen täytyy olla rakenteellisesti yhteensopiva kohdetaulun avaimen kanssa. Tässä ei ole: 1 sarake vs. 3 saraketta.

Virheellinen perustelu vei 1 pisteen, väärä päätelmä molemmat.

- c) Tarkastellaan relaatioita $J(\underline{K} \rightarrow P, L, M, N)$ ja $P(\underline{Q}, R, S)$. J:n monikkomäärä olkoon 2000 ja P:n 100. Relaatiot kytketään yhteen liitoksella, jossa liitosehtona on $J.K = P.Q$. Montako monikkoa sisältyy liitostulokseen?

2000, Viiteavain sisältyy taulun avaimen eikä siis voi olla tyhjä.

100-2000: 1 piste
100,200000: 0 pistettä
joku muu 0 pistettä

- d) Tarkastellaan kohdan c relaatiota P. Kumpi projektio on rivimäärältään suurempi $\pi_Q(P)$ vai $\pi_{Q,R}(P)$? Anna lyhyt perustelu.

Tulosrelaatioissa on sama määrä rivejä koska avain sisältyy kumpaankin projektiin.

jälkimmäinen vähintään yhtä suuri: 1p
ensimmäinen suurempi: 0p
Perustelu ontui tai väärin: -1p

(8p)

2.

Tarkastellaan lomamökkien vuokraukseen liittyvää tietokantaa:

```
asiakas(asiakasnumero, nimi, osoite, paikkakunta, puhelinnumero, sposti)
mökki(mtunnus, aluetunnus→alue, luokitus, osoite, paikkakunta, huoneiden_lkm,
      ajo_ohjeet)
varustelu(mtunnus→mökki, vtunnus→varuste, kuvailu, kuva)
viikkohinta(mtunnus→mökki, vuosi, viikko, hinta)
alue(aluetunnus, alueen_nimi, kuvailu)
varuste(vtunnus, nimi) -- esim. jääkaappi, pesukone, sauna, jne
varaus(vnro, asiakasnumero→asiakas, tehty_pvm)
varaussisältö(vnro→varaus, (mtunnus, vuosi, viikko)→viikkohinta)
```

Taulu viikkohinta sisältää rivin jokaiselle viikolle, jolle mökki on tai on ollut vuokrattavissa.

- a) Listaa Ivalossa sijaitsevat luokan 5 mökit (tunnus, osoite ja huoneiden lukumäärä), joita ei ole varattu viikolle 12 vuonna 2012.

```
select mtunnus, huoneiden_lkm, osoite
from mökki
where paikkakunta='Ivalo' and luokitus=5 and
mtunnus not in (select mtunnus from varaussisältö where vuosi=2012 and
viikko=12)
```

virheitä:

varattuna olon tarkistuksen puuttuminen -3p,

varattuna olon tarkistuksen virheellisyys: -1 - -3,

muita ongelmia -1..-2p

Ivalo on katsottu alueeksi: OK, mutta jos aluetietoa ei ole kytketty mökkiin -2p.

- b) Listaa Heinolassa sijaitsevat mökit (tunnus, osoite), joiden varustukseen sisältyy sekä sauna että kylpytynnyri.

```
select mtunnus, osoite
from mökki
where paikkakunta='Heinola' and
mtunnus in
(select mtunnus from varustelu, varuste
where varustelu.vtunnus=varuste.vtunnus and nimi='sauna') and
mtunnus in
(select mtunnus from varustelu, varuste
where varustelu.vtunnus=varuste.vtunnus and nimi='kylpytynnyri')
```

Ehto 'nimi='kylpytynnyri' and nimi='sauna' on aina epätosi koska rivillä on vain yksi arvo (-2p)

Ehto 'nimi='kylpytynnyriä or nimi='sauna' antaa väärän tuloksen mutta ei tyhjää -1p

Kytkenät puuttuvat -2p, Ristiriitaisia kytkentöjä, väärää tauluja -1 .. -3p

Muita ongelmia -1 .. -4p

sauna ja kylpytynnyri on hyväksytty myös varusteiden tunnuksiksi.

Heinola on tulkittu alueeksi OK, mutta kysely tulee hieman monimutkaisemmaksi.

- c) Laadi laskelma, josta käy ilmi kustakin Ivalossa sijaitsevasta mökistä vuonna 2011 maksettu kokonaishinta. Anna mökistä tunnus ja osoite.

```
select mökki.mtunnus, mökki.osoite, sum(hinta)
from mökki,viikkohinta,varaussisältö
where paikkakunta='Ivalo' and
      mökki.mtunnus=viikkohinta.mtunnus and
      viikkohinta.mtunnus=varaussisältö.mtunnus and
      viikkohinta.viikko=varaussisältö.viikko and
      viikkohinta.vuosi=varaussisältö.vuosi and
      viikkohinta.vuosi=2011
group by mökki.mtunnus, mökki.osoite
order by mökki.mtunnus, mökki.osoite
```

Arvostelusta:

Laskettu tuottoa myös viikoilta, jolloin mökki ei ole ollut varattuna: -2p
Ryhmittely puuttuu: -2
ryhmittelyssä ongelmia: -1..-2
Laskettu lukumäärä eikä summaa: -1
Liitosehdoissa ongelmia: 0..-2
syntaktisesti virheellinen kysely: -1..-3
tauluihin viitataan, mutta eivät ole mukana: -1..-2p

Tarpeettomien taulujen esim. varaus käytöstä ei ole sakotettu, paitsi jos kytkemättä

- d) Laadi luettelo vuoden 2011 huippuasiakkaista. Huippuasiakkaaksi lasketaan asiakkaat, jotka ovat vuoden aikana tehneet yli 3 varausta.

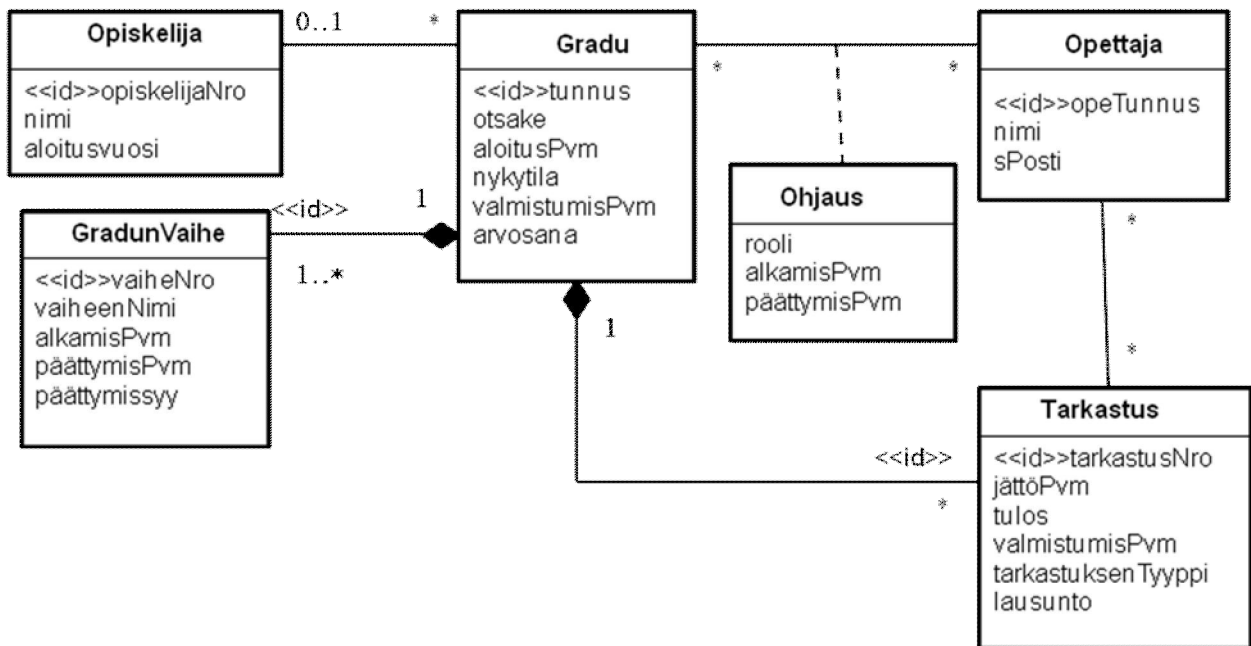
```
select nimi, asiakas.asiakasnumero, count(*) varauksia
from asiakas, varaus
where asiakas.asiakasnumero=varaus.asiakasnumero and
      tehty_pvm between date '2011-01-01' and date '2011-12-31'
group by nimi, asiakas.asiakasnumero
having count(*)>3
order by nimi, asiakas.asiakasnumero
```

Ryhmittely puuttuu: -2
ryhmittelyssä ongelmia: -1..-2
laskettu summa eikä lukumäärää: -1
Käytetty väärää tauluja: -3
Kytkentä puuttuu tai on virheellinen: -2p
päiväysten käsittelyssä ongelmia: 0..-1p

(16p)

Tehtävät 3-4 ovat toisella puolella.

3. Seuraavassa käsitekaaviossa on kuvattu graduhallintajärjestelmän tietosisältöä.



Määrittele kaavion perusteella tietosisällön esittämiseen tarvittavien relaatiotietokannan taulujen rakenne. Esitä taulurakenne tehtävissä 1 ja 2 käytetyllä tekniikalla. (6p)

```

opiskelija(opiskelijanumero,nimi,aloitusvuosi)
gradu(tunnus,otsake,aloitusPvm,nykytila,valmistumisPvm,arvosana,opiskelijaNro→Opiskelija)
opettaja(opeTunnus,nimi,sPosti)
gradunvaihe(tunnus→gradu,vaiheNro,alkamisPvm,paattymisPvm,paattymissy)
ohjaus(graduTunnus→gradu,opeTunnus→opettaja,rooli,alkamisPvm,paattymisPvm)
tarkastus(graduTunnus→gradu,tarkastusNro,jattoPvm,valmistumisPvm,tarkastuksenTyyppi,lausunto)
tarkastaja((graduTunnus,tarkastusNro)→tarkastus,opeTunnus→opettaja)
  
```

Virheitä:

Yhteyttä ei ole toteutettu -1p, kuitenkin max -4p

Yhteys toteutettu väärin -1p, kuitenkin max -4p

Viiteavaimet merkitty väärin, mutta idea selviää jotenkuten -2..-3p

Avaimet on merkitty väärin: -1 ..-2p

Ohjaus kohteen toteutus puuttuu kokonaan -1p

Avaimen ja viiteavaimen yhteensopivuusongelmia -1p

Hyväksytty myös ratkaisu, jossa tarkastusNro on yksinään Tarkastus-taulun avain.

4. Kalastuskerholla on noin 10 kilpailua vuodessa. Kerhon kilpailutoiminnan tulosten kirjausta varten on suunniteltu seuraava kalastajan tuloksia kuvaava taulu:
saalis (kilpailun_tunnus, kilpailupaikka, kilpailupäivä, kalastajatunnus,
kalastajan_nimi, kalalaji, kalastajan_syntymävuosi,
kalastajan_saaliin_kokonaispaino,
kalojen_kappalemäärä_kalastajan_saaliissa,
kalastajan_sijoitus_kisassa)

a) Mitä tarkoittaisi käytännössä riippuvuus *kilpailun_tunnus* → *kalastajan_syntymävuosi*?

kilpailut ovat ikäluokkakohtaisia, eli jokaisella osallistujalla on sama syntymävuosi.
virheessä voi olla aste-eroja selityksessä :- 1 tai 2p

b) Esitä funktionaalisen riippuvuutena sääntö 'samalla paikalla voidaan järjestää vain yksi kilpailu samana päivänä'.

kilpailupaikka, kilpailupäivä → kilpailutunnus

(oikeat sarakkeet mutta riippuvuus väärin 1p)

c) Olkoon taulussa voimassa riippuvuudet
kilpailun_tunnus → kilpailupaikka,
kilpailun_tunnus → kilpailupäivä,
kalastajatunnus → kalastajan nimi,
kalastajatunnus → kalastajan_syntymävuosi,
kilpailun_tunnus, kalastajatunnus → kalastajan_saaliin_kokonaispaino,
kilpailun_tunnus, kalastajatunnus →
kalojen_kappalemäärä_kalastajan_saaliissa,
kilpailun_tunnus, kalastajatunnus → kalastajan_sijoitus_kisassa,

Mikä sarake tai sarakeyhdistelmä on taulun avain. Onko taulu Boyce-Codd normaalimuodossa? Perustele vastauksesi lyhyesti. (6p)

Sarakkeet kilpailutunnus, kalastajatunnus ja kalalaji eivät esiinny missään riippuvuuksissa oikealla puolella, joten ne kuuluvat avaimen.

Kaikki riippuvuudet rikkovat BCNF:n vaatimusta siitä, että riippuvuudet olisivat avainperustaisia.

Avain väärin -1p.

Voitaisiin kuitenkin ajatella, että kilpailutunnus ehkä sittenkin määräisi kalalajin, joten myös avain (kilpailutunnus, kalastajatunnus) hyväksyttiin oikeaksi (jopa ilman perustelua).

Päätelmä tai sen perustelu väärin. -1p.

Jos avaimena on kilpailutunnus, kalastajatunnus niin 4 ensimmäistä rikkoo vaatimuksen. Siis ei ole BCNF tässäkään tapauksessa.

Tehtävät 1 ja 2 ovat toisella puolella.