

Helsingin yliopisto, Tietojenkäsittelytieteen laitos

Tietokantojen perusteet, 26.9.2006, H. Laine

Kirjoita kuhunkin erilliseen vastauspaperiin kurssin nimi, tentin päiväys, oma nimesi, syntymäaikasi ja nimikirjoituksesi.

Vastaukset sinisellä

Tarkastellaan elokuvaharrastajan tietokantaa. *Henkilö*-taulussa on tietoja näyttelijöistä, ohjaajista yms. elokuvan tekoon osallistuneista. Taulu *mukana* kytkee henkilöt elokuvaan. *Rooli* voi olla esimerkiksi 'ohjaaja' tai 'näyttelijä'. Taltio on esimerkiksi cd- tai dvd-levy. Tallenne on elokuvan kopio. Elokuvasta voi olla useita tallenteita, ja taltiolla voi olla useita tallenteita. Tallenteen *teko aika* ilmoittaa milloin kopio on tehty. *Teko aika* on timestamp-tyyppinen (päiväys ja kellonaika). *Kapasiteetti* ja *koko* ilmoitetaan megatavuina. Tallenteet numeroidaan juoksevasti. Taulussa *suurin_tallennetunnus* on kerrottu suurin käytössä oleva tallennetunnus. Taulussa *elokuva* voi olla tietoa elokuvista, joista ei vielä ole yhtään tallennetta.

```
henkilö(tunnus, nimi, tietoja) [200 riviä]
elokuva(elokuvatunnus, nimi, valmistumisvuosi, kuvailu) [100 riviä]
mukana(elokuvatunnus ->elokuva, rooli, tekijä -> henkilö) [1000 riviä]
taltio(taltiotunnus, tyyppi, kapasiteetti, säilytyspaikka) [150 riviä]
tallenne(tallennetunnus, tiedostonimi, elokuva->elokuva, taltio->taltio,
        teko aika, koko) [200 riviä]
suurin_tallennetunnus(nro) [1 rivi]
suurin_taltiotunnus(nro) [1 rivi]
```

1.

- a) Onko taulujen *henkilö* ja *elokuva* luonnollinen liitos mahdollinen, ja jos on, niin mitä sillä yritetään saada aikaan?

On mahdollinen. Luonnollinen liitos on aina mahdollinen. Liitosehto:
 $henkilö.nimi = elokuva.nimi$. Elokuva, jonka nimenä on jonkin henkilö-taulusta löytyvän elokuvantekijän nimi.

- b) Montako riviä sisältyy projektiioon $\pi_{elokuvatunnus, valmistumisvuosi}(elokuva)$?
100 koska avain mukana
- c) Miten suhtautuvat toisiinsa taulujen $\pi_{tunnus}(henkilö)$ ja $\pi_{tekijä}(mukana)$ rivimäärät?
Tekijä on tauluun henkilö viittaava viiteavain ja tunnus on henkilö-taulun avain, joten jokaisen tekijä-arvon pitää löytyä sarakkeesta tunnus. eli ensimmäinen on suurempi tai yhtä suuri kuin jälkimmäinen.
- d) Montako riviä on liitoksen $tallenne \bowtie_{elokuva=elokuvatunnus} elokuva$ tulostaulussa?
200, jokaiselle löytyy pari
- e) Taulusta *taltio* poistetaan rivi. Mitä on tehtävä muille tauluille viite-eheyden säilyttämiseksi?
(10p)
Poistettava tallenne-taulusta rivit, jotka viittaavat poistettuun riviin.

2. Anna yllä esiteltyihin tauluihin liittyen seuraavat SQL-kyselyt. Määrittele kyselyiden tuloksille tarkoituksenmukainen järjestys.

- a) Laadi luettelo elokuvista, joiden valmistumisvuotta ei ole kirjattu.
`select nimi, elokuvatunnus`

```
from elokuva
where valmistumisvuosi is null
order by nimi
```

- b) Listaa elokuvat, jotka Clint Eastwood on ohjannut, mutta joissa hän ei ole näytellyt.

```
select elokuva.nimi, elokuvatunnus from elokuva
where elokuvatunnus in
(select elokuvatunnus from mukana,henkilö
where tekijä=henkilö.tunnus and where rooli='ohjaaja' and
henkilö.nimi ='Clint Eastwood') and
elokuvatunnus not in
(select elokuvatunnus from mukana,henkilö
where tekijä=henkilö.tunnus and where rooli='näyttelijä' and
henkilö.nimi ='Clint Eastwood')
order by elokuva.nimi
```

- c) Selvitä eri ohjaajien ohjaamien elokuvien määrä keräilijän kokoelmissa. Järjestä tulos laskevaan järjestykseen määrän perusteella.

```
select henkilö.nimi, henkilö.tunnus, count(distinct elokuvatunnus))
from henkilö, mukana,tallenne
where henkilö.tunnus=mukana.tekijä and
mukana.elokuvatunnus=tallenne.elokuva
group by henkilö.nimi, henkilö.tunnus
```

- d) Mistä elokuvasta on eniten tallenteita? (16)

```
select elokuvatunnus, nimi, count(*)
from tallenne,elokuva
where tallenne.elokuva=elokuva.elokuvatunnus
group elokuvatunnus, nimi
having count(*) >=
all (select count(*) from tallenne group by elokuva)
```

3. Tarkastellaan yllä esiteltyä tietokantaa. Taltio '123' (dvd-levy) sisältää yhden tallenteen. Keräilijä tekee taltiosta kopion, jonka sijoittaa säilytyspaikkaan 'kaappi'. Mitä tietokannan ylläpito-operaatioita tarvitaan tilanteen kirjaamiseksi tietokantaan. Anna operaatiot myös SQL:llä. (6p)

```
Aloitetaan transaktio
commit;
```

```
Kasvatetaan taltiolaskuria
update suurin_taltiotunnus
set nro=nro+1;
```

```
Kasvatetaan tallennelaskuria
update suurin_tallennetunnus
set nro=nro+1;
```

```
Luodaan rivi uutta taltiota varten
insert into taltio
select nro, tyyppi, kapasiteetti,'kaappi'
```

```
from taltio, suurin_taltiotunnus
where taltiotunnus=123;
```

Luodaan rivi uutta tallennetta varten

```
insert into tallenne
select suurin_tallennetunnus.nro, tiedostonimi, elokuva, suurin_taltiotunnus.nro,
current_date, koko
from tallenne, suurin_tallennetunnus, suurin_taltiotunnus
where taltio=123;
```

```
Lopetetaan transaktio
commit;
```

4. Seuraava kaavio kuvaa järjestelmän tietosisältöä käsitetasolla. Miten tämä tietosisältö esitettäisiin relaatiotietokannan tauluina. Esitä ratkaisu tehtäväpaperin alussa käytetyllä kaaviotekniikalla. (6p)



kurssi(tunnus,nimi, muuta)
oppikirja(kirjatunnus,nimi)
tekijä(kirjatunnus->oppikirja,tekijä)
kurssi_oppikirja(tunnus->kurssi, kirjatunnus->oppikirja)

5. Tarkastellaan yrityksen tilaukset sisältävää relaatiota

tilaus (henkilönumero, henkilönnimi, henkilöosoite, tilausnumero, toimitusosoite, tilausrivinnumero, tuotetunnus, tilattumäärä).

- Mitä tarkoittaa riippuvuus toimitusosoite -> tilausnumero
Samaan osoitteeseen voi olla vain yksi tilaus
- Millaisia olisivat yrityksen tilaukset, jos tilausnumero olisi tämän taulun avain?
Kussakin tilauksessa voisi tilata vain yhtä tuotetta
- Miten esität riippuvuutena säännön: tuote esiintyy vain kertaalleen samassa tilauksessa.
tilausnumero, tuotetunnus -> tilausrivinnumero
-

Relaatiossa ovat voimassa riippuvuudet

- henkilönumero -> henkilönnimi,
- henkilönumero -> henkilöosoite
- tilausnumero -> henkilönumero,
- tilausnumero -> toimitusosoite
- tilausnumero, tilausrivinnumero -> tuotetunnus
- tilausnumero, tilausrivinnumero -> tilattumäärä

Minkä vuoksi relaatio ei ole Boyce-Codd -normaalimuodossa? (10p)

Boyce-Codd normaalimuodossa kaikissa riippuvuuksissa määrääjän pitää sisältää relaation avain. Tämän relaation avain sisältää ainakin sarakkeet tilausnumero ja tilausrivinnumero. Täten 4 ensimmäistä riippuvuutta rikkovat BC-vaatimuksen.