

CS–A1150 Tietokannat

Tentti 22.5.2019

Ei apuvälineitä eikä laskimia

Opiskelijat, joiden äidinkieli ei ole suomi, saavat halutessaan käyttää sanakirjaa, jos siinä ei ole merkintöjä. Nämä henkilöt voivat halutessaan saada myös tehtäväpaperin sekä suomeksi että englanniksi tai ruotsiksi.

Kirjoita jokaisen palauttamasi paperin yläreunaan selvästi kurssin koodi ja nimi sekä tentin päivämäärä, nimesi, opiskelijanumerosi ja tutkinto-ohjelmasi sekä palauttamiesi paperien kokonaismäärä.

1. a) (8 p) Tee seuraavan kuvauksen pohjalta UML-kaavio elokuvateatterin tietokantaa varten. Käytä kurssilla opetettua notaatiota ja merkitse avainattribuutit kurssilla opetetulla tavalla.
Elokuvateatteriin kuuluu useita saleja, joilla on yksikäsitteiden nimi. Saleista tiedetään myös niiden paikkojen lukumäärä. Saleissa järjestetään näytöksiä. Näytöksistä tiedetään näytöksen sali, päivämäärä, alkamisaika ja mikä elokuva näytöksessä esitetään. Elokuvilla on yksikäsitteinen tunnus, ja lisäksi elokuvasta tiedetään ainakin sen nimi, ohjaajan nimi ja elokuvan kesto minuutteina. Samaa elokuvaa voidaan esittää samaan aikaan useassa eri näytöksessä eri saleissa.
Elokuvateatterin tietokantaan on tallennettu asiakkaita, jotka voivat ostaa lippuja teatterin verkkokaupan kautta. Tehtävän yksinkertaistamiseksi oletetaan, että lipussa ei ole paikkanumeroita. Asiakkailla on yksikäsitteinen käyttäjätunnus ja lisäksi heistä tallennetaan nimi, sähköpostiosoite ja puhelinnumero. Asiakkaiden tekemistä ostoista tallennetaan yksikäsitteinen tunnus, mihin näytökseen liput on ostettu, osto-aika, kuinka monta lippua näytökseen on ostettu (yksi asiakas voi ostaa kerralla useita lippuja samaan näytökseen itseään ja seuralaisiaan varten – tällöin seuralaisten tietoja ei tallenneta minnekään), mitä lisäpalveluja ostoon kuuluu ja oston kokonaishinta. Osa lipuista myydään suoraan elokuvateatterin kassalta. Myös nämä ostot tallennetaan tietokantaan muuten samalla tavalla, mutta niistä ei tallenneta lainkaan asiakkaan tietoja. Mahdollisia ostoja sisältäviä lisäpalveluja ovat esimerkiksi juomat ja ruoat. Yhteen ostoon voi sisältyä esimerkiksi neljä lippua samaan näytökseen, kolme isoa virvoitusjuomaa, yksi pieni pop corn -annos ja kaksi jäätelöä. Lisäpalveluista tallennetaan niiden yksikäsitteisen tunnuksen lisäksi lisäpalvelun nimi ja koko.
 - b) (2 p) Muuta a-kohdan UML-kaavio relaatiomalliin. Esitä relaatiokaaviot ja alleviivaa niistä avainattribuutit.
2. Tehtävä liittyy seuraavaan tietokantakaavioon, joka kuvaa erään yrityksen osastoja, työntekijöitä, asiakkaita ja sen saamia tilauksia. Yrityksen työntekijöiden tehtävänä on hankkia sille tilauksia asiakasyrityksiltä.

Tietokantakaavio:

```
Employees(number, name, salary, departmentId)
Departments(id, name, directorNumber)
Customers(customerId, firmName, address)
Orders(orderNo, EmployeeNumber, customerId, date, value, description)
```

Relaatio Employees sisältää tiedot yrityksen työntekijöistä. Työntekijöistä on tallennettu numero, nimi, kuukausipalkka euroina ja sen osaston tunnus, jolla työntekijä työskentelee. Relaatio Departments sisältää tiedot yrityksen osastoista. Osastosta tiedetään sen tunnus, nimi ja osastoa johtavan työntekijän numero. Relaatio Customers sisältää tiedot yrityksen asiakasyrityksistä, joista tallennetaan asiakastunnus, nimi ja osoite. Relaatio Orders sisältää tiedot jokaisesta tilauksesta, jonka yrityksen työntekijät ovat hankkineet. Se kertoo tilausnumeron, tilauksen hankkineen työntekijän numeron, asiakasyrityksen asiakastunnuksen, tilauksen päivämäärän, arvon euroina ja lyhyen kuvauksen tehdystä tilauksesta.

Työntekijöiden palkat ja tilausten arvot ovat desimaalilukuja. Muiden attribuuttien arvot ovat merkkijonoja. Voit olettaa, että relaatioiden monikoiden attribuuteilla ei ole NULL-arvoja.

Kirjoita seuraavat kyselyt SQL-kielillä:

- a) (2 p) Niiden asiakkaiden asiakastunnukset ja nimet, jotka ovat tehneet yli 1000 euron arvoisen tilauksen.
- b) (2 p) Niiden työntekijöiden numerot ja nimet, joiden palkka on alle 3000 euroa, mutta jotka ovat hankkineet tilauksen yritykseltä, jonka nimi on Top Games.
- c) (2 p) Niiden työntekijöiden numerot ja nimet, joiden palkka on yli 5000 euroa, mutta jotka eivät ole hankkineet yhtään tilausta Top Games -nimiseltä yritykseltä.
- d) (2 p) Haetaan osastot, joiden työntekijät ovat hankkineet yhteensä vähintään 30 kappaletta yli 5000 euron arvoisia tilauksia (tässä siis lasketaan kunkin osaston kaikkien työntekijöiden tarpeeksi suurten tilausten kokonaismäärä). Kyselyn pitää tuottaa jokaisesta tällaisesta osastosta tunnus, nimi ja yli 5000 euron arvoisten tilausten lukumäärä.

Kirjoita seuraavat kyselyt relaatioalgebran lausekkeina:

- e) (2 p) Niiden osastojen nimet ja tunnukset, joiden johtajan palkka on alle 5500 euroa.
- f) (2 p) Pekka Puupää -nimisen työntekijän johtaman osaston kaikkien työntekijöiden numerot, nimet ja palkat. Jos yrityksessä on useita Pekka Puupää -nimisten työntekijöiden johtamia osastoja, kyselyn pitää tuottaa niiden kaikkien työntekijät.
3. a) (1 p) Tee kurssin SQL-lopputesti ja -kysely A+-järjestelmässä. Testi ja kysely ovat auki 29.5. asti.
- b) (3 p) Oletetaan, että tehtävän 2 tietokanta on niin suuri, että se ei mahdu kokonaan keskusmuistiin. Oletetaan edelleen, että yrityksessä on noin 1000 työntekijää, jotka työskentelevät 10 eri osastolla ja että jokaisella osastolla on suunnilleen yhtä monta työntekijää. Taulu `Employees` on tallennettu 100 levysivulle niin, että työntekijät ovat siinä täysin satunnaisessa järjestyksessä. Tietokannassa tehdään hyvin usein esimerkiksi seuraavia kyselyjä:
- Haetaan työntekijöitä, joilla on määrätty nimi, esimerkiksi

```
SELECT *  
FROM Employees  
WHERE name = 'Tiina Teekkari';
```
 - Haetaan työntekijöitä, jotka työskentelevät määrättyllä osastolla, esimerkiksi

```
SELECT *  
FROM Employees  
WHERE departmentID = 'D14';
```
- Haetut nimet ja osastot vaihtelevat, eivätkä ne ole joka kerralla samat kuin esimerkikykyselyssä. Kannattaako taululle `Employees` tehdä hakemisto (engl. index) attribuutin `name` perusteella? Entä attribuutin `departmentID` perusteella? Perustele vastauksesi muutamalla lauseella.
4. Tarkastellaan relaatiota $R(A, B, C, D, E)$, jossa on voimassa funktionaaliset riippuvuudet $A \rightarrow B$, $B \rightarrow C$ ja $B \rightarrow D$.
- a) (1 p) Perustele, miksi relaatio ei ole Boyce-Codd-normaali muodossa (BCNF:ssä).
- b) (6 p) Osita relaatio Boyce-Codd-normaali muotoon käyttämällä kurssilla (ja oppikirjassa) esitettyä algoritmia. Perustele lyhyesti jokainen muodostamasi uusi relaatio. Jatka osittamista niin pitkälle, että jäljellä on vain BCNF:ssä olevia relaatioita. Perustele, miksi lopulliset relaatiot ovat BCNF:ssä.
5. (7 p) Mitkä ovat transaktioilta (tapahtumilta, engl. transaction) vaadittavat ominaisuudet (neljä kappaletta) tietokannan hallintajärjestelmässä? Kuvaile jokaista ominaisuutta ja sen merkitystä muutamalla lauseella. Pelkät ominaisuuksien nimet ja 1-2 lauseen selitykset eivät riitä täysiin pisteisiin.

Muista vastata kurssin palautekyselyyn viimeistään 4.6.2019. Kyselyyn vastaamisesta saa 20 harjoitustehtävapistettä. Linkki kyselyyn on lähetetty kurssille ilmoittautuneille 15.5. sähköpostitse. Huomaa, että tämä on eri kysely kuin tehtävässä 3a mainittu lopputesti ja -kysely.