

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.

1. (4 p) Kerro, mitä osia tietokannan hallintajärjestelmässä (engl. database management system) on ja mitkä ovat näiden osien tehtävät. Suositeltu vastauksen pituus on 0,5-1 sivua.
2. a) (8 p) Tee seuraavan kuvauksen pohjalta UML-kaavio tietokantaa varten, joka sisältää tietoa yrityksen laitteista, niiden varauksista ja huolloista sekä laitteista vastaavista työntekijöistä. Käytä kurssilla opetettua notaatiota ja merkitse avainattribuutit kurssilla opetetulla tavalla.
 Yritys omistaa joukon laitteita, joita se käyttää toiminnassaan. Jokaisella erilaisella laitteella on yksikäsitteinen laitetunnus. Sen lisäksi laitteesta tiedetään sen nimi, tyyppi, malli ja valmistajan nimi. Yrityksellä voi kuitenkin olla useita kappaleita samanlaisia laitteita. Saman laitteen eri kappaleet erotetaan toisistaan laitenumeron perusteella. Laitenumerot ovat kuitenkin yksikäsitteisiä vain samanlaisen laitteen eri kappaleiden välillä, ja kahden erilaisen laitteen kappaleilla voi olla sama laitenumero. Laitenumeron lisäksi jokaisesta kappaleesta on tiedossa sen käyttöönottopäivämäärä ja ostohinta. Saman laitteen eri kappaleet on voitu hankkia yritykseen eri aikaan, joten nämä tiedot voivat vaihdella eri kappaleiden välillä.
 Yrityksessä on kullekin laitteelle määrätty vastuutyöntekijä tai useita vastuutyöntekijöitä. Heidän velvollisuutenaan on seurata laitteen kuntoa, tilata tarvittaessa laitteelle huolto ja opastaa muita työntekijöitä laitteen käytössä. Kaikilla saman laitteen kappaleilla on sama vastuutyöntekijä tai samat vastuutyöntekijät. Jokaisella laitteella on vähintään yksi vastuutyöntekijä. Työntekijöistä tiedetään työntekijätunnus (yksikäsitteinen), syntymäaika, nimi ja sähköpostiosoite. Yhdellä työntekijällä voi olla vastuullaan useita erilaisia laitteita.
 Tietokantaan tallennetaan myös tieto eri laitteiden huolloista. Huoltotapahtumalla on yksikäsitteinen huoltotunnus. Kukin huoltotapahtuma liittyy aina yksittäiseen laitteen kappaleeseen. Sen lisäksi, mitä kappaletta on huollettu, tiedetään huoltokerrasta sen alkamisaika, kesto ja lyhyt kuvaus.
 Työntekijät voivat varata työtehtävissä tarvitsemiaan laitteita haluamilleen aikaväleille. Varaus kohdistuu aina yksittäiseen laitteen kappaleeseen ja varaus on aina täsmälleen yhdelle työntekijälle. Varauksilla on yksikäsitteinen varausnumero ja niistä tiedetään lisäksi varauksen alkamis- ja päättymisaika.
- b) (2 p) Muuta a-kohdan UML-kaavio relaatiomalliin. Esitä relaatiokaaviot ja alleviivaa niistä avainattribuutit.
3. Tehtävä liittyy alle annettuun tietokantakaavioon, joka kuvaa erään yrityksen työntekijöitä, toimipisteitä, asiakkaita, projekteja ja sitä, miten työntekijät ovat kirjanneet työtunteja eri projekteille.

Relaatio `Employees` sisältää tiedot yrityksen työntekijöistä. Jokaisesta työntekijästä tiedetään yksikäsitteinen tunnus (attribuutti `ID`), nimi, työntekijän tuntiveloitus (asiakkaalta perittävä maksu jokaisesta työntekijän tekemästä työtunnista) ja sen toimipisteen tunnus, jossa työntekijä työskentelee.

Relaatio `Branches` sisältää tiedot yrityksen toimipisteistä ja relaatio `Customers` yrityksen asiakkaista. Relaatio `Projects` sisältää tiedot yrityksen projekteista (sekä meneillään olevista että jo valmistuneista). Oletetaan, että jokaisen projektin on tilannut täsmälleen yksi asiakas, jonka tunnuksen kertoo attribuutti `customerID`. Attribuutti `budget` kertoo projektin suunnitellun budjetin arvon euroina.

Yrityksen työntekijöiden pitää päivittäin kirjata, kuinka monta tuntia he ovat käyttäneet aikaa mihinkin projektiin. Nämä kirjaukset tallennetaan relaatioon `Reports`. Attribuutti `employeeID` kertoo kirjauksen tehneen työntekijän tunnuksen, `projectnumber` projektin, jolle kirjaus tehdään, `date` päivämäärän, jota kirjaus koskee, sekä `hours` sen, kuinka monta tuntia työntekijä on kirjannut tehneensä töitä tälle projektille ko. päivänä.

Tuntiveloitukset, tuntimäärät ja budjetit ovat desimaalilukuja. Muiden attribuuttien arvot ovat merkkijonoja. Voit olettaa, että attribuuteilla ei ole NULL-arvoja.

Tietokantakaavio:

```
Employee(ID, name, hourlyfee, branchID)
Branch(BID, name, address)
Project(number, name, customerID, budget)
Customer(customerID, name, address)
Report(employeeID, projectnumber, date, hours)
```

Kirjoita seuraavat kyselyt SQL-kielellä:

- a) (2 p) Niiden asiakkaiden tunnukset ja nimet, jotka ovat tilanneet yritykseltä vähintään yhden projektin, jonka budjetti on yli 50 000 euroa.
- b) (2 p) Niiden projektille 'PA-123' (projektin numero) tunteja kirjanneiden työntekijöiden tunnukset ja nimet, joiden tuntiveloitus on yli 50 euroa ja joiden toimipisteen osoitteeseen sisältyy sana 'Helsinki'.
- c) (2 p) Niiden työntekijöiden tunnukset, nimet ja tuntiveloitukset, jotka ovat kirjanneet tunteja vain projektille 'PA-123' (projektin numero), mutta eivät millekään muulle projektille.
- d) (2 p) Hae ne toimipisteet, joiden työntekijät ovat kirjanneet 'PA-123' (projektin numero) yhteensä vähintään 500 tuntia (tässä lasketaan siis yhteen kunkin toimipisteen kaikkien työntekijöiden ko. projektille kirjaamien tuntien määrä). Kyselyn pitää tuottaa näistä toimipisteistä tunnus, nimi ja niiden työntekijöiden lukumäärä, jotka ovat tässä toimipisteessä kirjanneet tunteja tälle projektille.

Kirjoita seuraavat kyselyt relaatioalgebran lausekkeina:

- e) (2 p) Niiden työntekijöiden tunnukset, jotka ovat kirjanneet tunteja jollekin 'Huippufirma'-nimisen yrityksen tilaamalle projektille.
 - f) (2 p) Niiden asiakkaiden tunnukset ja nimet, jotka ovat tilanneet yritykseltä vähintään kaksi eri projektia, mutta joiden tilaamista projekteista yhdenkään budjetti ei ole yli 50 000 euroa.
4. Tarkastellaan relaatiota $R(A, B, C, D, E)$, jossa on voimassa funktionaaliset riippuvuudet $A \rightarrow B C D$, $B \rightarrow C$ ja $E \rightarrow A$.
- 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. Vastaa muutamalla lauseella seuraaviin kysymyksiin.

- a) (2 p) Tarkastellaan kahta transaktiota (engl. transactions). Transaktiossa T_1 asiakas A varaa paikan 4.4.2020 lähtevälle lennolle AY001 ja transaktiossa T_2 asiakas B varaa paikan 5.4.2020 lähtevälle lennolle AY001. Oletetaan, että transaktioihin kuuluvat operaatiot tapahtuvat seuraavassa järjestyksessä ja että samaan aikaan ei suoriteta muita transaktioita, joilla olisi tapauksen kannalta merkitystä.

1. Asiakas A etsii vapaan paikan 4.4.2020 lähtevältä lennolta AY001 ja saa tiedon, että paikka 16A on vapaa. (Operaatio kuuluu transaktioon T_1 .)
2. Asiakas B etsii vapaan paikan 5.4.2020 lähtevältä lennolta AY001 ja saa tiedon, että paikka 16A on vapaa. (Operaatio kuuluu transaktioon T_2 .)
3. Asiakas A varaa paikan 16A 4.4.2020 lähtevältä lennolta AY001. (Operaatio kuuluu transaktioon T_1 .)
4. Asiakas B varaa paikan 16A 5.4.2020 lähtevältä lennolta AY001. (Operaatio kuuluu transaktioon T_2 .)

Onko tämä transaktioiden T_1 ja T_2 suoritus sarjallistuva (engl. serializable)? Perustele vastauksesi lyhyesti.

- b) (2 p) Jos a-kohdan suoritus ei ole sarjallistuva, miten sitä pitäisi muuttaa, jotta se olisi sarjallistuva? Jos a-kohdan suoritus on sarjallistuva, minkä siinä pitäisi olla eri tavalla, jotta se ei olisi sarjallistuva? Vastaa joko ensimmäiseen tai toiseen kysymykseen sen mukaan, mitä vastasit a-kohtaan.
- c) (3 p) Mitä eroja on transaktioiden eristyvyystasoilla (engl. isolation levels) READ COMMITTED, REPEATABLE READ ja SERIALIZABLE? Miksi aina ei vaadita täyttä sarjallistuvuutta?