

Tietokantojen perusteet, K 2004

Tietokantasuunnittelusta

- Tietokantasuunnittelun pääperiaatteena on tiedon toiston välttäminen.
- Tiedon toistumiseen liittyy monenlaisia ongelmia
 - toistuva tieto vie 'turhaa' tilaa
 - ylläpito muodostuu hankalaksi kaikki 'kopiot' päivitettävä
 - ylläpito-operaatioilla voi olla odottamattomia sivuvaikutuksia.
 - Toisaalta toistosta on etuakin – tiedon haku saattaa nopeutua

Tietokantojen perusteet, K 2004

Tietokantasuunnittelusta

- Esimerkki taulusta, joka ei käyttyä hyvin:
EMP_DEPT:

E_no	E_name	E_bdate	D_no	D_name	D_location
1	M.Seppä	1.3.59	3	Myynti	Helsinki
2	D.Leivo	4.10.40	3	Myynti	Helsinki
3	K.Koivu	30.1.66	4	Hallinto	Lahti
4	B.Oja	2.5.65	4	Hallinto	Lahti
5	O.Itä	10.2.55	6	Tuotanto	Helsinki

Avain: E_no
Jos O.Itä poistetaan, häviää tieto tuotanto-osastosta

Toistettava jokaisen osaston t:n kohdalla

Jos Hallinto muuttaa Espooseen om muutettava useita rivejä

Tietokantojen perusteet, K 2004

Ei tarpeetonta toistoa

E_no	E_name	E_bdate	D_no	D_no	D_name	D_location
1	M.Seppä	1.3.59	3	3	Myynti	Helsinki
2	D.Leivo	4.10.40	3	4	Hallinto	Lahti
3	K.Koivu	30.1.66	4	6	Tuotanto	Helsinki
4	B.Oja	2.5.65	4			
5	O.Itä	10.2.55	6			

Parempi ratkaisu

Tietokantojen perusteet, K 2004

Tietokantasuunnittelusta

- Tietokannan suunnittelun vaiheita ovat
 - tietosisällön kartoitus
 - luokkakaaviona (Johdatus sovellussuunnitteluun)
 - kartoittamalla attribuutit ja niiden väliset riippuvuudet
 - loogisten rakenteiden suunnittelu ja
 - teknisten rakenteiden suunnittelu.
- Loogisen rakenteen suunnittelun tavoite on sijoittaa yhteenkuuluvat tiedot samaan taulu

Tietokantojen perusteet, K 2004

Muunnos luokkakaaviosta relaatiokaavioon

- Järjestelmän määrittelyn yhteydessä järjestelmän tietosisältö voidaan määrittellä luokkakaavion avulla
- Luokkakaavio kuvaa tällöin kuvaa järjestelmän pysyväisluonteisia (persistent) tietoja - tietojen on säilyttävä ohjelman suorituskertojen välillä
- Luonteva paikka säilyttää olioiden tila on tietokanta:
 - oliorakenteinen tietokanta (oliotietokanta)
 - relaatiotietokanta

Tietokantojen perusteet, K 2004

Muunnos luokkakaaviosta relaatiokaavioon

- relaatiotietokanta
 - standardoitu SQL ja liittymät perinteisiin ohjelmointikieliin
 - ei suoraa tukea olioiden välisille yhteyksille
 - ei luokkahierarkiaa eikä periytymistä
 - vain data ei toimintoja
 - uusimmissa tkhj:ssä mukana oliopiirteitä
 - edellyttää muunnosta oliorakenteista relaatiorakenteiksi

Tietokantojen perusteet, K 2004

Muunnos luokkakaaviosta relaatiokaavioon

➤ Lähtökohtana normalisoitu luokkakaavio = **tiettyjä sääntöjä noudattava kaavio**

- Yhteydet näkyviä, ei attribuuteiksi piilotettuja
- Kukin asia esitetään vain kertaalleen
- Ei johdettavissa / pääteltävissä olevaa tietoja
- Yhteydet määritely oikeiden osapuolten välillä

Tietokantojen perusteet, K 2004

Yhteydet piilotettu

- Olioiden välinen kytkentä esitetään aina yhteytenä, sitä ei piiloteta ominaisuudeksi (attribuutiksi)

Kurssi
Opettajan_nimi

Opettaja
Nimi

Kytkentä on piilotettuna kurssin attribuuttiin opettajan_nimi

Tietokantojen perusteet, K 2004

Yhteyden piilotus

- Olioiden välinen kytkentä esitetään aina yhteytenä, sitä ei piiloteta ominaisuudeksi (attribuutiksi)

~~Kurssi~~
~~Opettajan_nimi~~

Opettaja
 Nimi

opetus

Kurssi

Tietokantojen perusteet, K 2004

Asiat esitetään vain kertaalleen

- Samaa asiaa ei pidä esittää sekä yhteytenä että ominaisuutena

~~Tilaus~~
~~Tilaajan_nimi~~

0..n tilaaja

1

Asiakas
 Nimi

Tietokantojen perusteet, K 2004

Ei esitetä johdettavissa olevaa tietoa

- tulisi selvittää mitkä tiedot ovat perustietoja ja mitkä voidaan päätellä (laskea) muiden tietojen perusteella
- attribuutin arvo tai yhteyden olemassaolo voi olla pääteltävissä

Kurssi
 nimi
~~ilmoittautuneiden lkm~~

ilmoittautunut
 0..n 0..n

Opiskelija
 nimi

laskettavissa

Tietokantojen perusteet, K 2004

Ei esitetä johdettavissa olevaa tietoa

Kurssi

1..n

Harjoitusryhmä

0..n

~~Kurssilla~~

1..n

Opiskelija

Harjoitusryhmä

0..n ryhmässä

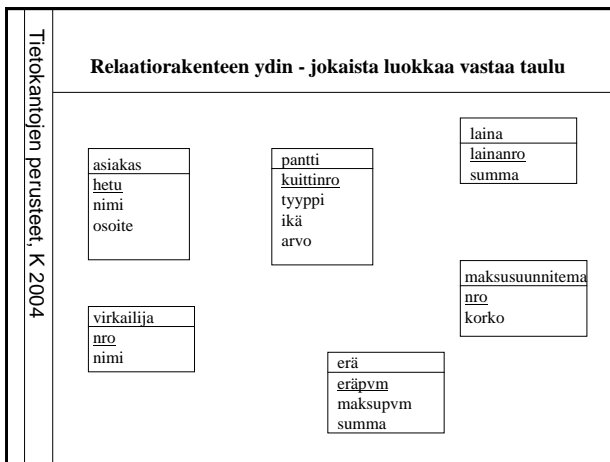
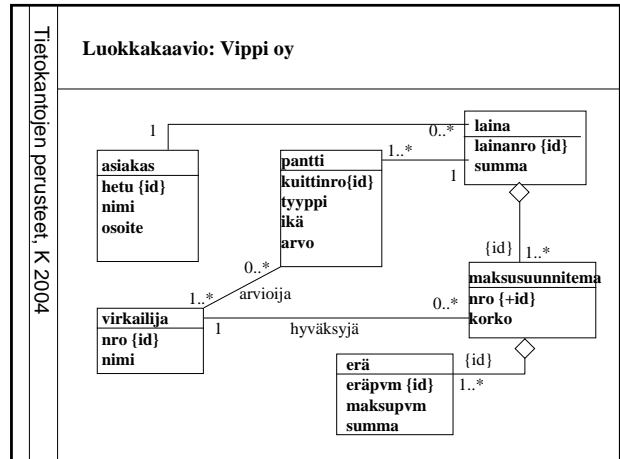
Opiskelija

Kurssilla-yhteys on pääteltävissä, jos jokaisen kurssillaolijan on oltava jossain ryhmässä

Tietokantojen perusteet, K 2004

Muunnoksen perussäännöt:

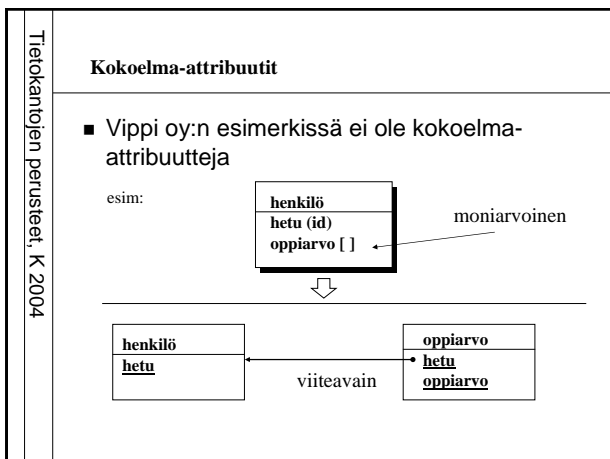
- Kutakin olioluokkaa vastaa samanniminen taulu
- Kullakin luokan yksiarvoisella attribuutilla on samanniminen vastinsarake luokkaa vastaavassa taulussa
 - Luokan tunnistavia attribuutteja vastaavat sarakkeet kuuluvat luokkaa vastaavan taulun avaimen



Tietokantojen perusteet, K 2004

Kokoelma-attribuutit

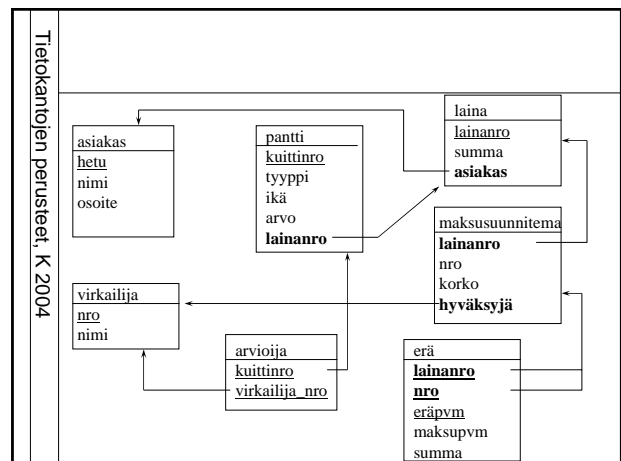
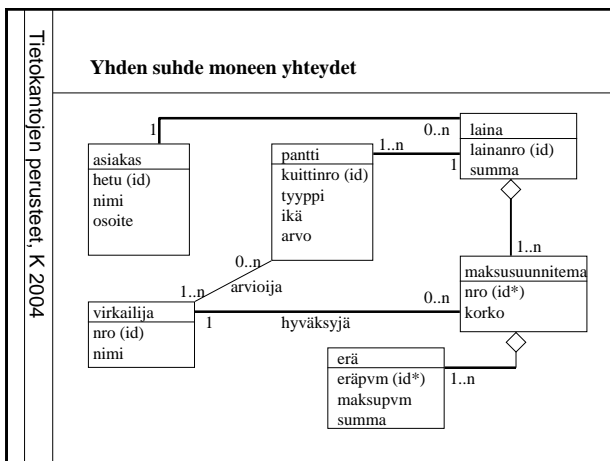
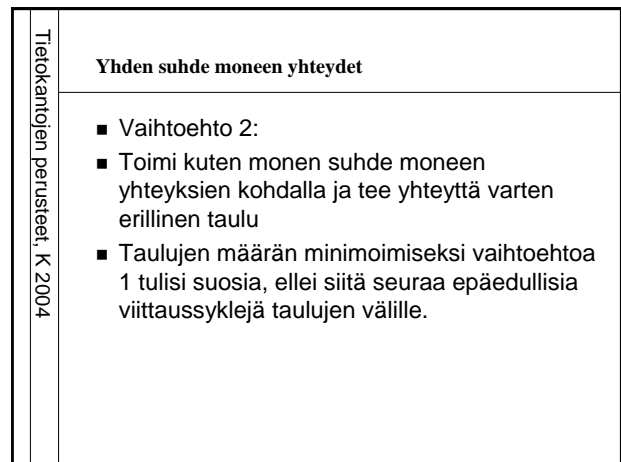
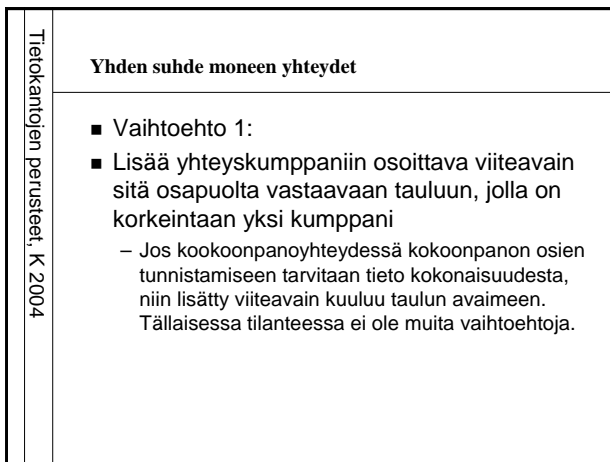
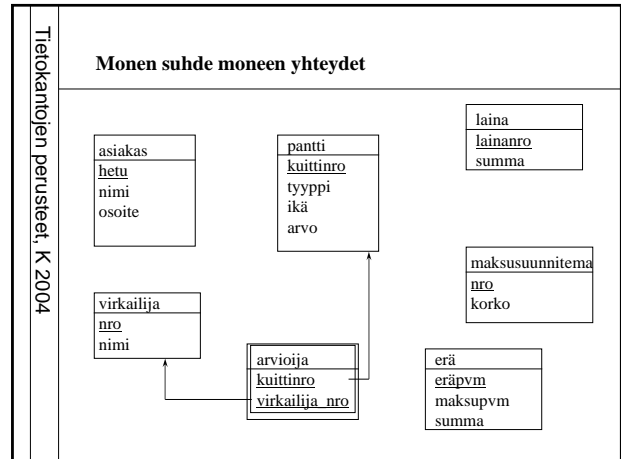
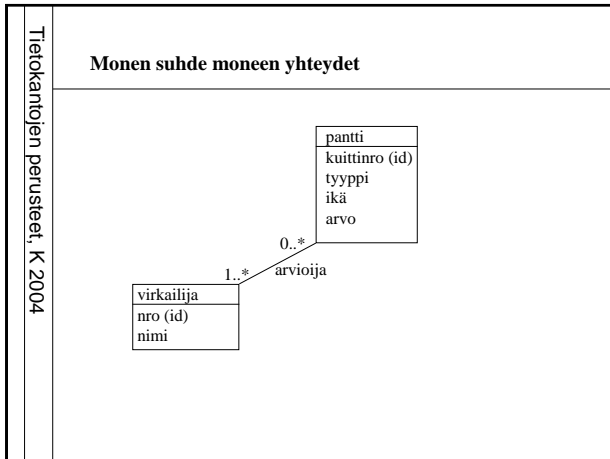
- Kutakin kokoelma-attribuuttia (moniarvoista) vastaa taulu, jonka sarakkeina ovat
 - viiteavain kokoelmatyyppisen attribuutin sisältävää luokkaa vastaavaan tauluun,
 - sarake attribuutin arvoa varten.
- Taulun kaikki sarakkeet kuuluvat avaimen



Tietokantojen perusteet, K 2004

Monen suhde moneen yhteydet

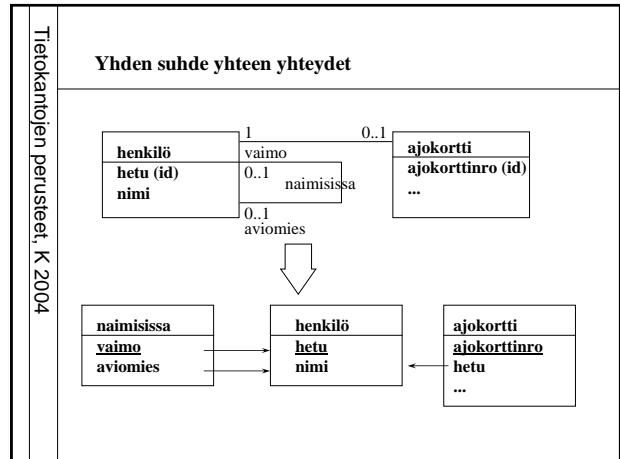
- Kutakin monen suhde moneen yhteyttä vastaa taulu
 - Taulun nimi = yhteyden nimi
 - Taulun sarakkeina ovat yhteyden osapuoliin osoittavat viiteavaimet.
 - Taulun kaikki sarakkeet kuuluvat taulun pääavaimen
 - monen suhde moneen yhteys on yhteys, jossa kummankin osapuolen maksimiosallistumisrajoite on suurempi kuin 1.



Tietokantojen perusteet, K 2004

Yhden suhde yhteen yhteydet

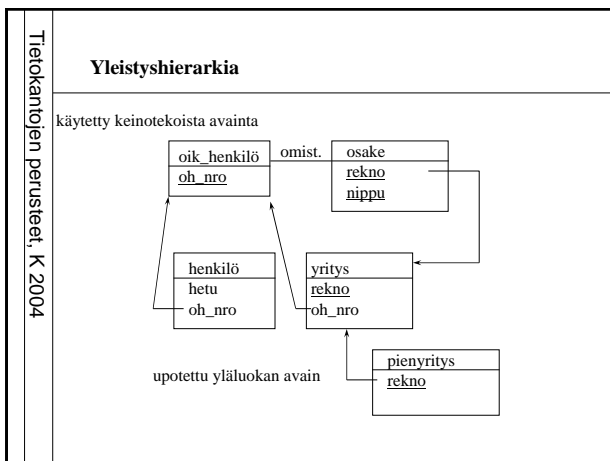
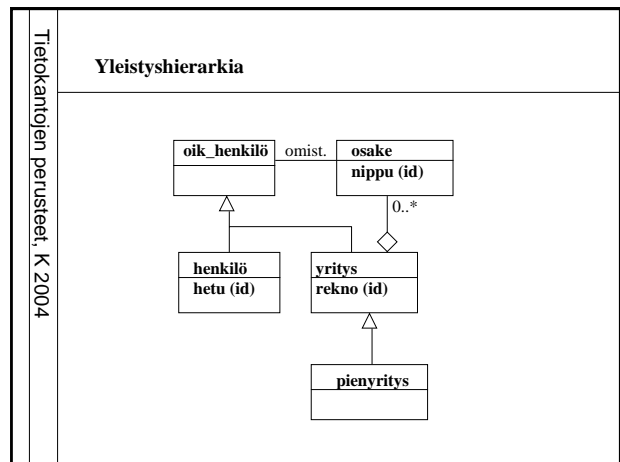
- Vaihtoehtoja:
- Lisää yhteyskumppaniin osoittava viiteavain yhteyden toista osapuolta vastaavaan tauluun
 - ensisijaisesti sille, jolle yhteys on pakollinen
 - jos yhteys on pakollinen molemmille, valitse kumpi tahansa
- Tee erillinen taulu, kuten monen suhde moneen tapauksessa.
 - jos yhteys on molemmille osapuolille valinnainen



Tietokantojen perusteet, K 2004

Yleistyshierarkia

- Vaihtoehto 1:
- Taulu sekä ala- että yläluokalle. Alaluokan tauluun sijoitetaan yläluokan tauluun osoittava viiteavain.
 - Jos yläluokalla on käyttäjän määrittelemiä tunnistavia attribuutteja otetaan viiteavain alaluokan taulun avaimeksi.
 - Jos yläluokan taululla ei ole käyttäjän määrittämää tunnistavaa attribuuttia, joudutaan sille luomaan keinotekoinen avain. Tätä ei oteta alaluokan taulun avaimeksi, jos alaluokalle on määritelty tunnistavia attribuutteja.



Tietokantojen perusteet, K 2004

Yleistyshierarkia

- Vaihtoehto 2: Alaluokalle ei tehdä omaa taulua, vaan alaluokan tauluun loogisti kuuluvat sarakkeet upotetaan yläluokan tauluun.

Tietokantojen perusteet, K 2004	<p>Tietokantasuunnittelusta</p> <ul style="list-style-type: none"> Loogisen rakenteen suunnittelun tavoite on sijoittaa yhteenkuuluvat tiedot samaan tauluun Relaatiomalliin liittyvässä suunnitteluteoriassa yhteenkuuluvuus määritellään riippuvuuksien avulla Keskeinen riippuvuus: funktionaalinen riippuvuus Attribuutti B on funktionaalisesti riippuva attribuutista A (A määrää funktionaalisesti B:n), jos ja vain jos kaikissa relaatiokaavion R ilmentymissä kuvaus A:n arvojoukolta B:n arvojoukolle on funktionaalinen.
---------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Tietokantojen perusteet, K 2004	<p>Tietokantasuunnittelusta</p> <ul style="list-style-type: none"> Kuvaus $f:v(A) \rightarrow v(B)$ on funktionaalinen, jos kaikissa relaatiokaavion R ilmentymissä jokainen A:n arvo kuvautuu yhdelle B:n arvolla eli, <ul style="list-style-type: none"> jos riveillä r ja s attribuutilla A on sama arvo ($r.A=s.A$), niin näillä riveillä täytyy myös B-attribuuteilla olla keskenään sama arvo ($r.B=s.B$). Funktionaalinen riippuvuus tarkoittaa sitä, että attribuutin B arvo on <u>yksikäsitteisesti selvitettävissä</u> kun tiedetään attribuutin A arvo. Selvittäminen voisi tapahtua kyselyllä select distinct B from R where A=a; (tuloksena olisi enintään yksi rivi)
---------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Tietokantojen perusteet, K 2004	<p>Tietokantasuunnittelusta</p> <ul style="list-style-type: none"> Funktionaalista riippuvuutta, jossa A määrää B:n merkitään $A \rightarrow B$. Attribuuttia A kutsutaan <u>määrääjäksi</u>. Yksittäisen attribuutin A tilalla voi olla myös attribuuttiyhdistelmä. Tavoitteena on kuitenkin löytää yhdistelmät, joissa on minimaalinen määrä attribuutteja, Jos $A \rightarrow B$, voidaan määrääjään lisätä mikä tahansa attribuutti x ja pätee $Ax \rightarrow B$. Avain määrää funktionaalisesti kaikki relaatiokaavan attribuutit Riippuvuudet ilmaisevat jonkin kuvattavaan todellisuuteen liittyvän säännön
---------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Tietokantojen perusteet, K 2004	<p>Tietokantasuunnittelusta</p> <ul style="list-style-type: none"> Tarkastellaan relaatiokaaviota Kurssilainen(<ul style="list-style-type: none"> Kurssikoodi, Hetu, OpiskelijaNimi, KurssiNimi, TehtavaLkm). Oletetaan että relaatio sisältää tietoja useista opiskelijoista ja useista kursseista ja kuvaa opiskelijoiden osallistumista kursseille.
---------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Tietokantojen perusteet, K 2004	<p>Tietokantasuunnittelusta</p> <ul style="list-style-type: none"> Hetu-->OpiskelijaNimi Yhtä henkilötunnusta kohden on vain yksi Opiskelijanimi Henkilötunnuksen perusteella on selvitettävissä opiskelijan nimi Kurssikoodi->KurssiNimi Yhtä kurssikoodia kohti on vain yksi kurssinimi Hetu, Kurssikoodi --> TehtäväLkm <ul style="list-style-type: none"> Yhtä henkilötunnus – kurssikoodi yhdistelmää kohti on vain yksi tehtävälukumäärä = Henkilön tekemien tehtävien lukumäärä – tehtävämäärä – on kurssikohtainen
---------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Tietokantojen perusteet, K 2004	<p>Tietokantasuunnittelusta</p> <ul style="list-style-type: none"> Mitkä sarakkeet samaan relaatiokaavioon? Erilaisia kriteerejä Yhteenkuuluvuussääntönä, ns. Boyce-Codd normaalimuodon sääntö: <ul style="list-style-type: none"> relaatiokaavion R attribuutit kuuluvat yhteen, jos ja vain jos relaatiokaavioon R ei liity yhtään sellaista funktionaalista riippuvuutta, jossa määrääjä ei sisältäisi relaation avainta. Kurssilainen relaation avain on pari Kurssikoodi, Hetu => Kaikki attribuutit eivät BC-säännön mukaan kuulu yhteen koska <ul style="list-style-type: none"> Hetu-->OpiskelijaNimi ja Kurssikoodi->KurssiNimi rikkovat BC-sääntöä
---------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Tietokantojen perusteet, K 2004	<p>Tietokantasuunnittelusta</p> <ul style="list-style-type: none">■ Attribuutit on uudelleenjärjesteltävä relaatiokaavioiksi, joissa yhteenkuuluvuussäännöt ovat voimassa.■ päädytään relaatiokaavioihin<ul style="list-style-type: none">• Opiskelija(Hetu, OpiskelijaNimi)• Kurssi(Kurssikoodi, KurssiNimi)• Osallistuminen(Hetu, Kurssikoodi, TehtäväLkm).
---------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Tietokantojen perusteet, K 2004	<p>Jako Boyce-Codd normaalimuotoon</p> <ul style="list-style-type: none">■ Miten järjestellään:<ul style="list-style-type: none">- Karsi riippuvuusjoukko minimaaliseksi – poista johdettavissa olevat riippuvuudet (materiaalissa sääntöjä)- Ryhmitä riippuvuudet yhteisen määrääjän (vasen puoli) perusteella.- Muodosta jokaista ryhmää kohti oma relaatio- Jos alkuperäisen kaavion yksikään avain ei sisälly muodostuneisiin relaatiokaavioihin tee sille oma kaavio- Anna relaatiokaavioille nimet
---------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Tietokantojen perusteet, K 2004	<p>Tietokantasuunnittelusta</p> <ul style="list-style-type: none">■ Tilauslomaketta analysoitaessa löydettiin seuraavat attribuutit:<ul style="list-style-type: none">■ lomakenumero,■ tilaajan tunnus,■ tilaajan nimi,■ tilaajan osoite,■ tilaajan puhelinnumero,■ toimitusosoite,■ rivinumero,■ tavarankoodi,■ tavaranimi,■ tilattu määrä, ja■ tilauspäivä.
---------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Tietokantojen perusteet, K 2004	<p>Tietokantasuunnittelusta</p> <ul style="list-style-type: none">• lomakenumero → tilaajan tunnus (lomakkeella voidaan ilmoittaa vain yksi tilaaja)• tilaajan tunnus → tilaajan nimi (tunnus identifioi tilaajan, joten sen perusteella saamme selville kaikki tilaajaan liittyvät tiedot)• tilaajan tunnus → tilaajan osoite,• tilaajan tunnus → tilaajan puhelinnumero,• lomakenumero → toimitusosoite (lomakkeella voidaan tilata tavaroita vain yhteen paikkaan)
---------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Tietokantojen perusteet, K 2004	<p>Tietokantasuunnittelusta</p> <ul style="list-style-type: none">• tavarankoodi → tavaranimi (tavarankoodi on tavarantunniste, jonka kautta päästään kaikkiin tavarantietoihin)• lomakenumero, rivinumero → tavarankoodi (lomakkeen rivillä voi ilmoittaa yhden tilattavan tavarankoodin)• lomakenumero, rivinumero → tilattu määrä (lomakkeen rivillä voi ilmoittaa yhden tilattavan tavarankoodin määrän)• lomakenumero → tilauspäivä (lomakkeeseen liittyy yksikäsitteinen tilauspäivä)
---------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Tietokantojen perusteet, K 2004	<p>Tietokantasuunnittelusta</p> <ul style="list-style-type: none">■ Jako relaatioihin yhteisen määrääjän perusteella:<ul style="list-style-type: none">■ X(tilaajan tunnus, tilaajan nimi, tilaajan osoite, tilaajan puhelinnumero)■ Y(tavarankoodi, tavaranimi)■ Z(lomakenumero, tilauspäivä, tilaajan tunnus, toimitusosoite)■ T(lomakenumero, rivinumero, tavarankoodi, tilattu määrä)
---------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Tietokantojen perusteet, K 2004

Tietokantasuunnittelusta

- Jos kaavioille löytyy kuvaava nimi jako on onnistunut:
 - tilaaja(tilaajan tunnus, tilaajan nimi, tilaajan osoite, tilaajan puhelinnumero)
 - tavara(tavaran koodi, tavaran nimi)
 - tilaus(lomakenumero, tilauspäivä, tilaajan_tunnus, toimitusosoite)
 - tilausrivi(lomakenumero, rivinumero, tavaran koodi, tilattu määrä)

