

TIETOJEN MALLINNUKSI

NORMALISOINTI

HOVI, HUOTARI, LAHDENMÄKI:

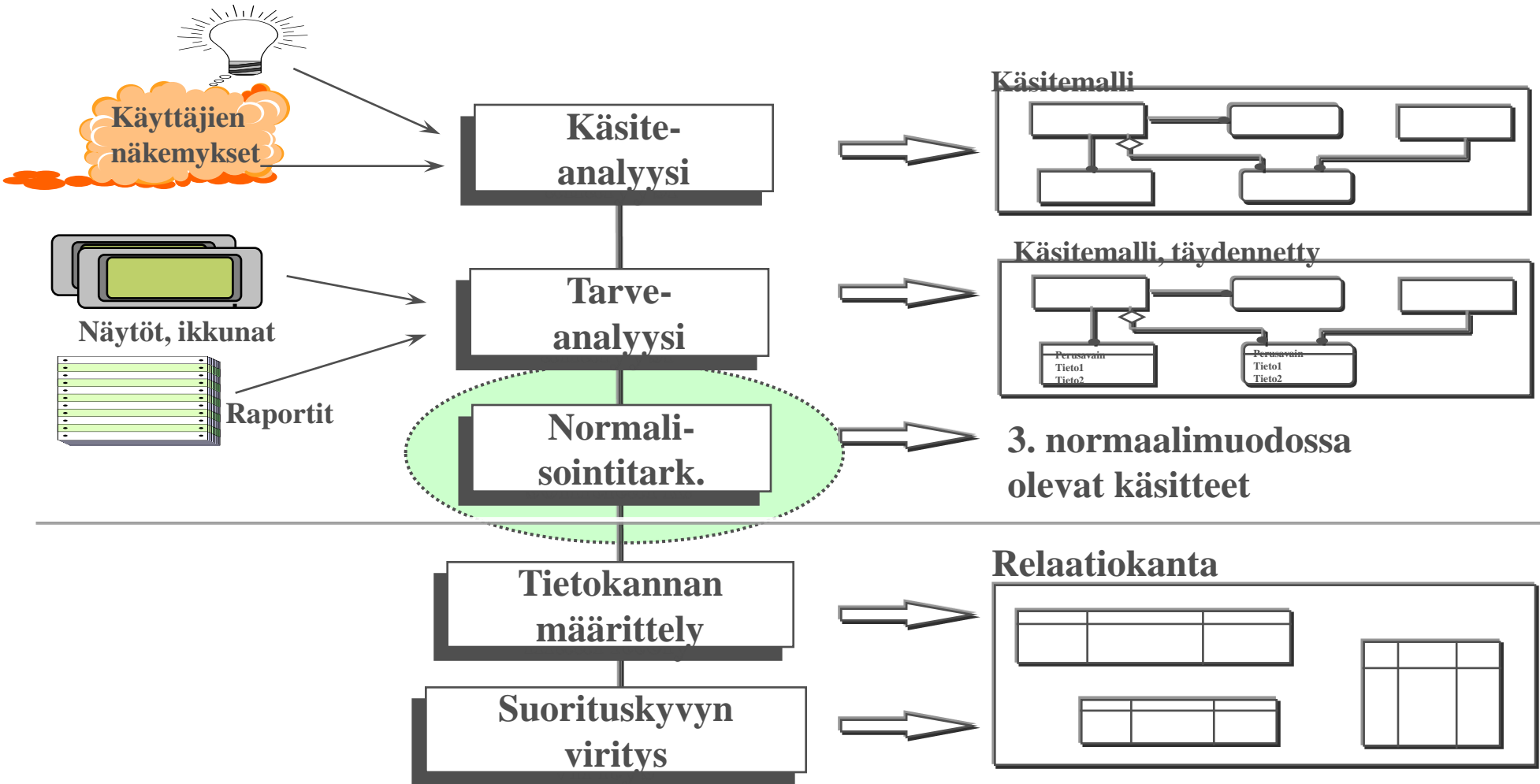
TIETOKANTOJEN SUUNNITTELU & INDEKSOINTI

DOCENDO (2003, 2005) LUKU 5

© JOUNI HUOTARI & ARI HOVI



SUUNNITTELUPUTKI



NORMALISOINNIN TARKOITUS

- Normalisointi (normalisation) on menetelmä, jonka avulla tietorakenteita voidaan jalostaa ”parempaan” tallennusmuotoon
- Parempi tarkoittaa tässä yhteydessä rakennetta,
 - jossa on tietojen toistaminen (redundanssi) minimoitu
 - joka on tehokas päivitysten kannalta
 - joka on helpompi pitää yhdenmukaisena, sillä tiedot tarvitsee päivittää vain yhteen paikkaan
 - joka on muutosjoustava

NORMALISOINTIVAIHEET

1. Normaalimuoto

Erota toistuvat ryhmät ja moniarvoiset sarakkeet omaksi käsitteekseen

2. Normaalimuoto

Jokaisen ei-avaintiedon tulee olla riippuvainen koko perusavaimesta

3. Normaalimuoto

Poista sisäiset (ei-avaimeen kohdistuvat) riippuvuudet

1. NORMAALIMUOTO

HENKILÖ

<u>Htun</u>	Nimi	Palkka	Pvm_1	Palkka_1	Pvm_2	Palkka_2

1. Normaalimuoto :

Poista toistuvat ryhmät ja moniarvoiset sarakkeet

HENKILÖ

<u>Htun</u>	Nimi	Palkka

PALKKAHISTORIA

<u>Htun</u>	<u>Pvm</u>	Palkka

1. NORMAALIMUOTO (VERSIO 2)

HENKILÖ

<u>Htun</u>	Nimi	Palkat	Pvmt

1. Normaalimuoto :

Poista toistuvat ryhmät ja moniarvoiset sarakkeet

HENKILÖ

<u>Htun</u>	Nimi	Muut tiedot

PALKKAHISTORIA

<u>Htun</u>	<u>Pvm</u>	Palkka

FUNKTIONAALINEN RIIPPUVUUS



Sarake B on funktionaalisesti riippuvainen A:sta, jos A:ta kohti on korkeintaan yksi B:n arvo kunakin ajanhetkenä.

HENKILÖ (Htun, HenkNimi, Osoite)



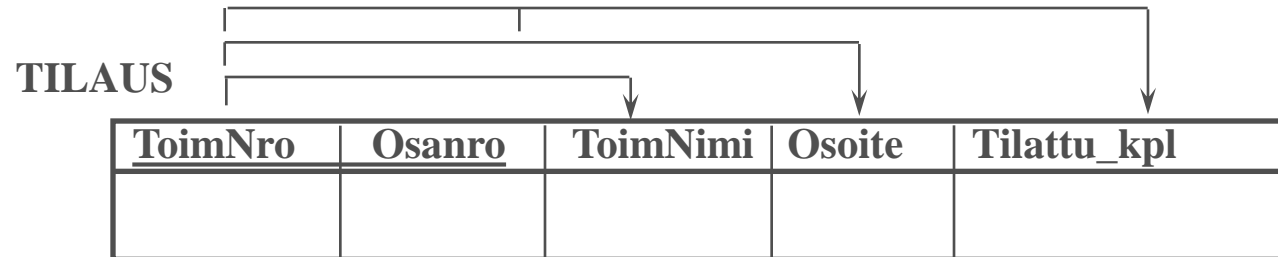
?

HENKILÖ (Htun, HenkNimi, Osoite)



?

2. NORMAALIMUOTO



2. Normaalimuoto:

Jos avain koostuu useasta sarakkeesta, on jokaisen muun sarakkeen oltava riippuvainen *koko* avaimesta.

TOIMITTAJA

<u>ToimNro</u>	Nimi	Osoite

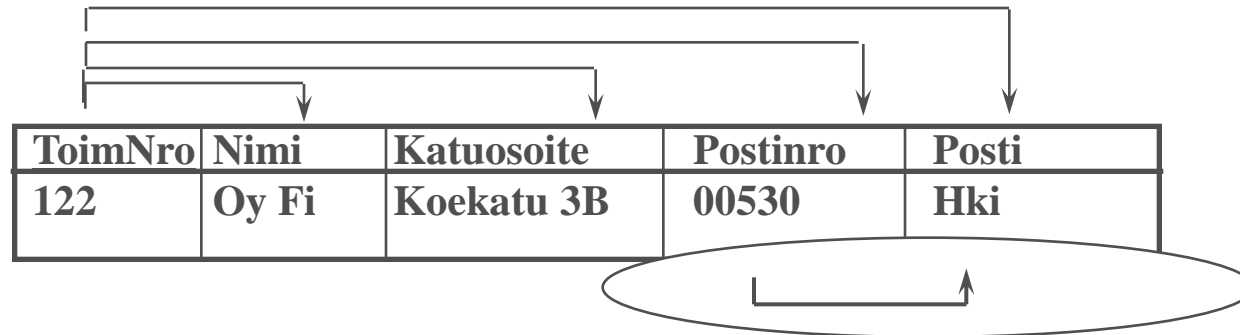
TILAUS

<u>ToimNro</u>	<u>Osanro</u>	Tilattu_kpl

- Mitä jos näin ei tehdä?
 - Ei voida tallettaa toimittajatietoja etukäteen
 - Vanhoja tilauksia poistettaessa tietyn toimittajan tiedot voivat häipyä kokonaan
 - Toimittajan osoitteen muutos tulee moneen paikkaan

3. NORMAALIMUOTO

TOIMITTAJA



3. Normaalimuoto:

**Kaikkien sarakkeiden tulee olla riippuvia
vain perusavaimesta.**

TOIMITTAJA

<u>ToimNro</u>	Nimi	Katuosoite	Postinro
122	Oy Fi	Koekatu 3B	00530

POSTI

<u>Postinro</u>	Postitoimip
00530	Hki

- Mitä jos näin ei tehdä?
 - Postitoimipaikkoja ei voida tallettaa ennen toimittajan talletusta
 - Postitoimipaikan nimen muutos tulee moneen paikkaan
- Lisähyötynä käyttäjätyytyväisyys ja virheiden ehkäisy:
 - Käyttöliittymä voidaan rakentaa siten, että käyttäjä valitsee listasta postinumeron => postitoimipaikka saadaan automaattisesti

3NM JA KÄSITTEET

- Suunnitteluputkessa normalisoidaan käsitteitä
 - Normalisointi voi tarvittaessa myöhemmin kohdistua tauluihin, tietueisiin jne.
 - Käsittemallin normalisointi tarkistetaan käsite käsitteeltä
 - Melko työläs operaatio, mutta vaivannäkö yleensä kannattaa
- Yhteenvetona (NM = normaalimuoto): kaikkien ei-avaintietojen tulee olla funktionaalisesti riippuvia
 - perusavaimesta (1. NM),
 - koko perusavaimesta (2. NM) ja
 - vain perusavaimesta (3. NM)

RELAATIOKAAVA (ESIM. NORMALISOIDUN RATKAISUN KUVAAMISEEN)

- Perusidea: kirjoita käsitteen (relaation/taulun) nimi ja sulkuihin ominaisuudet (sarakkeiden nimet)
- Esim.
OPISKELIJA(Opiskelijatunnus, OpiskelijanNimi)
- Yksilöivä tieto esim. alleviivaamalla tai kirjoittamalla PK
- Viiteavaimen kohdalle FK (tai voit vaikka *kursivoida* sen)
- Jos viiteavain on osa perusavainta, kirjoitetaan PFK

TEHTÄVÄ: NORMALISOINTI (DEMOX OY)

- Tarkista tekemäsi käsitelmäsi (Demox Oy)
 - missä normaalimuodossa malli on?
 - olisiko järkevää normalisoida mallia pitemmälle?
 - mitä etuja tai haittoja pitkälle normalisoidusta ratkaisusta on?
- Vinkki: usein tehdään nk. ideaalimalli, jossa normalisointi on ulotettu ”kaikkiin” tilanteisiin ja sen jälkeen tehdään perustellut rajaukset siitä, mikä tulee olemaan toteutettava ratkaisu (ts. palataan lähelle toista normaalimuotoa)

NORMALISOINTITEHTÄVÄT OPTIMASSA

- Tee kirjan (Hovi et. al. 2003/2005) harjoitukset 9 ja 10
- Voit tehdä yhden ratkaisun, johon yhdistät vastaukset a-, b-, c- ja d-tehtäviin, tai kuvata ratkaisusi kuhunkin erikseen
- Vastaus joko relaatiokaavana tai käsittemallina
- Kirjaa tekemäsi oletukset
- Palauta vastauksesi yhtenä tiedostona Optimaan



DENORMALISOINTI

- Normalisoidun rakenteen yhdistäminen, esim. 3. normaalimuodosta 2. normaalimuotoon
- Taulujen väliset liitokset vähenevät
=> tiedon haku nopeutuu
- Sovelluksen on hoidettava tiedon ylimäärästä aiheutuvat lisäpäivitykset automaattisesti
- Denormalisointia ei tehdä käsitelmalleille vaan se tehdään vain, jos suorituskyky niin vaatii
- Ks. kirjan luku 5 (s. 95)

LISÄTIETOA TIETOKANNAN NORMALISOINNISTA

- http://en.wikipedia.org/wiki/Database_normalization
- <http://databases.about.com/od/specificproducts/a/normalization.htm>
- <http://www.databasejournal.com/sqletc/article.php/1428511>

