

Tietomallit

- Tietomallilla (data model) tarkoitetaan tiedon rakenteen ja tiedolle suoritettavan käsittelyn määrittelevää käsitteistöä
- Tietoa voidaan tarkastella eri näkökulmista - eri abstraktiotasoilla
- Perinteinen jako
 - Käsitetaso
 - Rakennetaso
 - Talletustaso

Näkökulmat tietoon

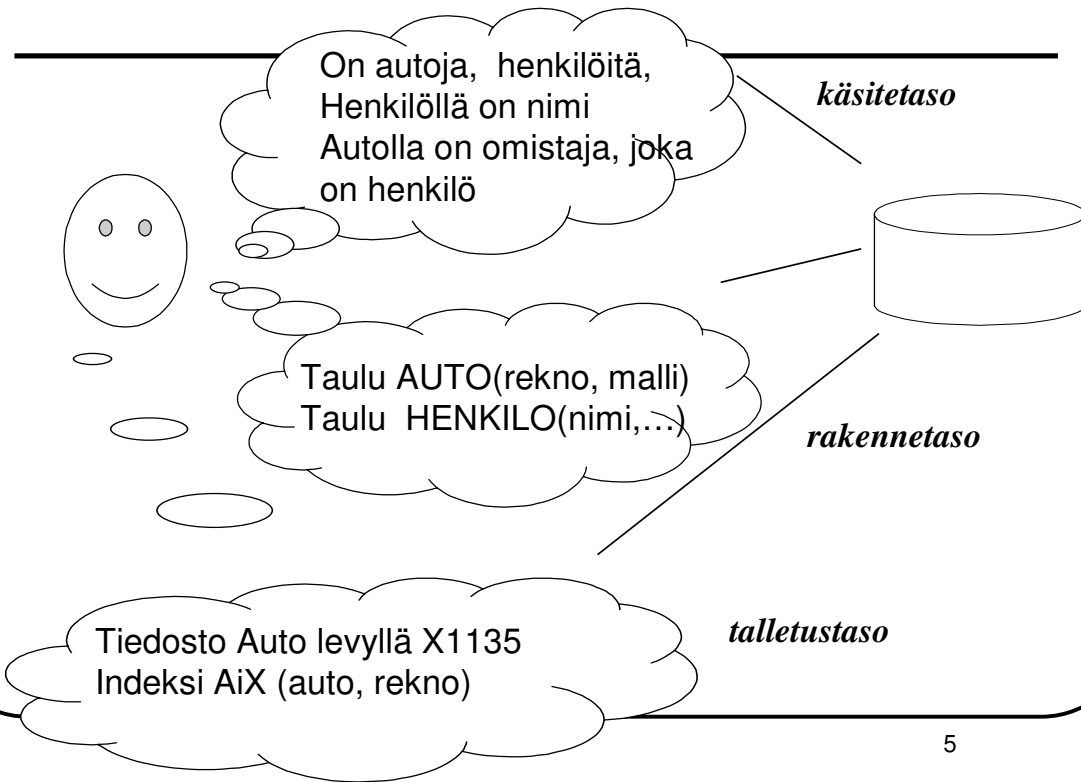
- **Käsitetaso**, kohdetaso (**conceptual level**, real word level):
 - Mitä tietoja käsitellään
 - millainen on se kohde, jota tiedoilla pitäisi kuvata
 - asiat, joita pitäisi esittää
 - Mitä tietoelementtien arvot tarkoittavat
 - Miten tiedot liittyvät yhteen
 - Mitä sääntöjä tietoihin liittyy
 - Ohjelmistoista riippumaton kuvaus tiedoista
 - Käsitellään JSS-kurssilla

Näkökulmat tietoon

- **Rakennetaso**, looginen taso (**structural level**, logical level):
 - Minkälaisia käsiteltäviä rakenteita tiedot muodostavat - esim. eri ohjelmointikielten näkemykset tiedosta
 - Miten ohjelmoijat ja tietokannan suoraikäyttäjät näkevät tiedot

Näkökulmat tietoon

- **Talletustaso**, fyysinen taso (**physical level**):
 - Minkälaisina koneenläheisinä rakenteina tiedot tallennetaan ja miten niitä voidaan käsitellä - minkälaiset rakenteet tehostavat hakua, onko tiedot hajautettu vai kaikki keskitetysti samassa paikassa.
 - Käsitellään Tietokannan hallinta kurssilla



Abstraktiotasot tiedon käsittelyssä

käsitetaso	sisällön suunnittelu
rakennetaso	kyselyt, ohjelmointi
fyysinen taso	viritys, hajautus

Mitä malleja olisi tarjolla?

käsitetaso

- Entity-Relationship mallit (ER)
- Oliomallit (UML)
- Semanttiset tietomallit

rakennetaso

- **Relaatiomalli**
- Oliomallit
- Hierarkkinen malli
- Verkkomalli

fyysinen taso

- Toimittajakohtaiset käsitteet

Relaatiotietokannat

- Perustana rakennetason tietomalli **relaatiomalli** (the relational model of data)
- perusteoria: E.F.Codd 1970
- ensimmäiset kaupalliset toteutukset 70-luvun lopulla
- yleistynyt 80-luvun lopulla
- DB2, Oracle, Informix, Sybase, MS SQLServer, yms.

Relaatiomalli

- Abstraktio tiedoista
- Mallin perustana on näkemys tietokannasta joukkona tietoalkioiden muodostamia matemaattisia relaatioita
- Yksinkertainen peruskäsitteistö, vähän käsitteitä
- Helppo ymmärtää havainnollisesti taulukkoesityksenä

9

Relaatio havainnollisesti

Relaatiokaavion nimi

Attribuutit

AUTO	Reknro	Väri	Vmalli
	ACM-256	musta	1988
	MAC-532	sininen	1994
	ISO-795	musta	1992
	OSI-228	punainen	1987
	HCI-449	valkoinen	1993

Monikot

Arvo

10

Relaatio vielä havainnollisemmin

Sarakkeet

Relaatiokaavion nimi

Sarakenimet

AUTO	Reknro	Väri	Vmalli
	ACM-256	musta	1988
	MAC-532	sininen	1994
	ISO-795	musta	1992
	OSI-228	punainen	1987
	HCI-449	valkoinen	1993

Rivit

Arvo

11

Relaation matemaattinen määritelmä

- Olkoon D_1, D_2, \dots, D_n arvojoukkoja, joiden ei tarvitse olla erillisiä.
- Relatio R on joukko järjestettyjä arvojonoja, monikkoja (tuple, n -tuple), joiden 1. arvo kuuluu joukkoon D_1 , 2. arvo joukkoon D_2 jne.
- Matemaattisesti relaatio on siis ristitulon $D_1 \times D_2 \times \dots \times D_n$ osajoukko.

12

Monikko

- Monikko on järjestetty arvojen jono (a_1, \dots, a_n) . Taulukko-muotoisessa esityksessä sitä vastaa taulukon rivi
- **Ristitulo :**
 - Joukkojen $A=\{1,2,3\}$ ja $B=\{a,b\}$ ristitulo (karteesinen tulo) $A \times B$ on kaikkien niiden parien (x,y) joukko, joissa ensimmäinen alkio kuuluu joukkoon A ja toinen joukkoon B eli joukko
 - $\{(1,a), (1,b), (2,a), (2,b), (3,a), (3,b)\}$
- Esimerkiksi funktio voidaan esittää lähtöarvojen ja maaliarvojen muodostamien parien joukkona eli kaksipaikkaisena relaationa (binäärirelaatio).

13

Arvojoukko

- Arvojoukko (Domain)
 - kokoelma atomisia (osiin jakautumattomia) arvoja, esim.
 - kokonaisluvut
 - henkilötunnukset
 - enintään 4-merkkiset merkkijonot
 - Merkkijonot
 - Jotkut operaatiot edellyttävät että arvojoukon arvojen välillä on olemassa järjestys
- Tyhjä arvo (NULL-arvo, olematon) sisältyy jokaiseen arvojoukkoon!

14

Attribuutti = sarakkeen nimi

- Attribuutti (attribute) on arvon paikalle monikossa annettu nimi.
- Attribuuttiin liittyy **tulkinta** sille, mitä asiaa kyseisessä paikassa oleva arvo ilmaisee.
- Jokaiseen attribuuttiin liittyy arvojoukko ja jokaisessa monikossa jokin arvojoukkoon kuuluva arvo.
- Attribuuttiin liittyvä arvojoukko tulisi määritellä siten, että sen arvoilla pystytään esittämään kaikki ne asiat, jotka attribuuttiin liittyvän tulkinnan mukaan pitää kyetä esittämään
 - **esim. kaikki mahdolliset värit**

Relaatiokaavio

- Relaatiokaavio (relation schema) määrittää relaation rakenteen:
 - mitä attribuutteja relaatioon kuuluu
 - millaiset arvojoukot attribuuteilla on ja
 - mikä tulkinta kuhunkin attribuuttiin liittyy (mitä arvot esittävät)
- Relaatiokaaviolla on nimi.

Relaatiokaavio

- Yksinkertaisimmillaan relaatiokaavio voidaan esittää muodossa $R(A_1, \dots, A_m)$, missä R on kaavion nimi ja A_1, \dots, A_m ovat attribuutteja
 - Tällaista määritystä käytettäessä oletetaan, että lukija osaa kuvaavien attribuuttinimien perusteella arvata niiden arvojoukot ja tulkinnat
- Esim.
 - $Auto(Reknro, Väri, Vuosimalli)$

17

Relaatiokaavio

- Täydellisempi esitys ottaa mukaan arvojoukon määrittelyn:

*Auto(Reknro: Suomalaiset_rekisterinumeroit,
Väri: Autovärit,
Vuosimalli: Vuosiluvut >1900)*
- Relaatio on relaatiokaavion **ilmentymä**.
 - Relaatiokaavion ilmentymät kuvaavat jotain todellisuuden ilmiötä, vaikkapa autojen olemassaoloa tietyllä hetkellä, esim. 1.1.2004
 - Toista ajankohtaa kuvaa eri relaatio, joskin saman kaavion ilmentymä sekin.

18

Relaatiokaavio

- Relaatiotietokannassa on tyypillisesti kustakin relaatiokaaviosta vain yksi ilmentymä kerrallaan
- Tähän ilmentymään viitataan kaavion nimellä
 - siten, kun puhutaan relaatiosta Auto, tarkoitetaan kaavion Auto tämänhetkistä ilmentymää
- Tietokannan ylläpidossa kaavionimellä osoitettava ilmentymä vaihdetaan toiseksi

Relaatiokaavio

- Matemaattisesti relaatio on joukko
 - Matemaattisessa joukossa alkio ei toistu, joten kaikki relaation monikot ovat keskenään erilaisia
 - Joukossa alkioden järjestyksellä ei ole merkitystä
- Attribuuttien järjestyksellä relaatiokaaviossa ei myöskään ole merkitystä

Relaatiokaavion käsitteitä

- Relaation koko (cardinality)
 - relaation monikoiden lukumäärä
- Relaation aste (degree)
 - relaatiokaavion attribuuttien lukumäärä
- Relaatietietokanta (relational database)
 - sisältää yleensä useita relaatioita
- Relaatietietokantakaavio (relational database schema)
 - relaatietietokannan relaatiot määrittelevien relaatiokaavioiden kokoelma

Teoria vs havaintoesitys

Teorian käsite	Havaintoesityksen käsite
Relaatio	Taulukko
Monikko	Taulukon rivi
Attribuutti	Sarakkeen nimi

Avain (key)

- Relaaation monikot ovat keskenään erilaisia
 - Monikko pystytään identifioimaan (= erottamaan muista) sisältönsä perusteella
 - Identifiointiin ei välttämällä tarvita edes kaikkien attribuuttien arvoja
- Relaaation **avain** (key) on sellainen attribuutti tai niiden yhdistelmä, jolle pätee
 - missään relaatiokaavion ilmentymässä ei voi olla kahta tai useampaa monikkoa, joissa on sama avain-attribuuttien arvoyhdistelmä
 - yhdistelmästä ei voi poistaa yhtään attribuuttia siten, että jäljelle jäävät täyttäisivät edellisen ehdon (minimaalisuus)

23

Avain

AUTO	Reknro	Väri	Vmalli
	ACM-256	musta	1988
	MAC-532	sininen	1994
	ISO-795	musta	1992
	OSI-228	punainen	1987
	HCI-449	valkoinen	1993

- Reknro ja Vmalli erikseen erottelevat yllä olevassa ilmentymässä.
- Vmalli ei kuitenkaan erottele kaikissa mahdollisissa ilmentymissä = joskus voi olla sellainen autojoukko, että siinä on useampia saman vuosimallin autoja.

24

Avain

- Avaimeen kuluvalle attribuutilla täytyy olla jokaisessa monikossa aito, ei-tyhjä arvo
- Relaatiokaavioon voi liittyä useita attribuuttiyhdistelmiä, jotka täyttävät avaimelle asetetut vaatimukset
- Jokin ehdokkaista tulee valita **pääavaimeksi**:
 - **ensisijaiseksi** tavaksi osoittaa ja viitata relaation monikkoihin
 - Työntekijä(henkilötunnus,..., työntekijänumero)**
 - kumpikin kävisi, valitaan tarkoituksenmukaisempi

Avain

- Pääavain esitetään relaatiokaaviossa alleviivaamalla siihen kuuluvat attribuutit

Auto (Reknro, Väri, Vuosimalli)

Työntekijä (Työntekijänumero,...)

Pelivaraus (kenttänumero, alkuaika, kesto, nimi)

Viiteavain (foreign key)

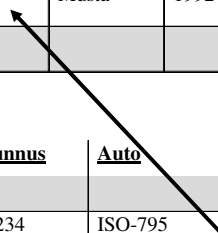
- Tietokanta muodostuu useasta relaatioista
- Relaatioiden monikot ovat usein kytköksissä toisiinsa
 - työntekijämonikko kytkeytyy sitä osastoa kuvaavaan monikkoon, jolla työntekijä työskentelee
 - lainausmonikko kytkeytyy kirjaa ja lainaajaa kuvaaviin monikoihin
- Relaatiotietokannassa kytkentä saadaan aikaan sisällyttämällä monikkoon kytkettävän, siis vieraan, monikon (pää)avain.

27

Viiteavain

Auto	<u>Rekno</u>	Väri	Vmalli
	ISO-795	Musta	1992

Omistus	<u>Henkilötunnus</u>	Auto
	121212-1234	ISO-795



28

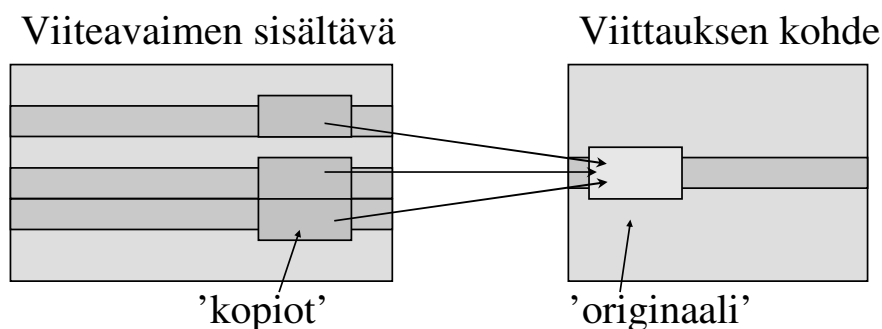
Viiteavain

- Attribuuttia tai attribuuttiyhdistelmää, jonka arvot toteuttavat kykennän (viittauksen) toiseen monikkoon kutsutaan viiteavaimeksi (foreign key).
- Viiteavaimessa voidaan sallia myös tyhjäärvot – tällöin relaation monikkoja ei tarvitse kytkeä mihinkään

29

Viiteavain

- Vaikka viiteavain viittaakin yhteen suuntaan sitä voi hyödyntää kaksisuuntaisesti esim. selvittäessä :
 - mikä auto kytkeytyy tiettyyn omistukseen
 - mitkä omistukset kytkeytyvät tiettyyn autoon



30

Viiteavain

- Viiteavaimen esittäminen relaatiokaaviossa

Omistus(Henkilötunnus→Henkilö, Auto→Auto)

- Henkilötunnus viittaa Henkilöön ja Auto Autoon

Osallistuu(Kuka→Opiskelija,

(Kurssikoodi, RyhmäNumero)→Harjoitusryhmä)

Kahdesta attribuutista muodostuva viiteavain, tämä tarkoittaa, että Harjoitusryhmä relaation avain muodostuu myös kahdesta attribuutista

Viiteavain

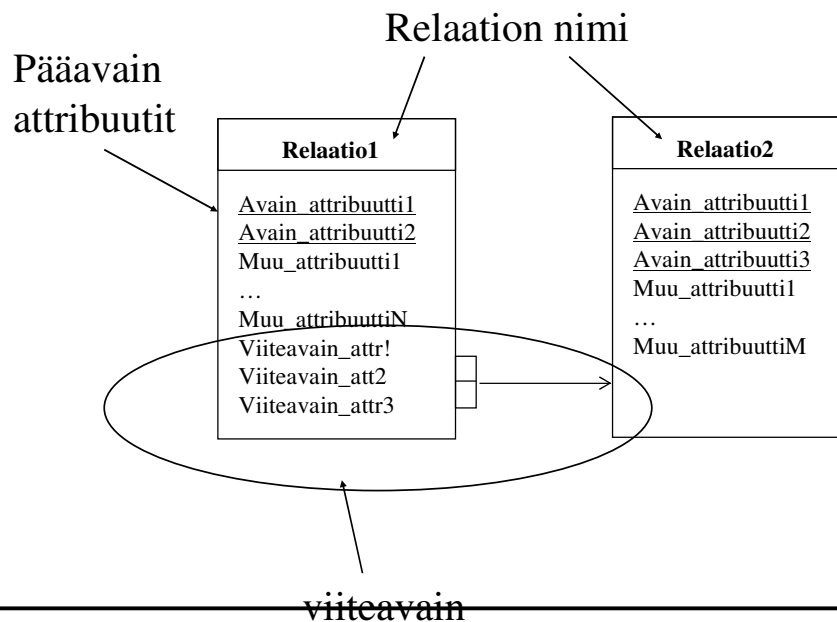
- Viiteavaimella voidaan kytkeä toisiinsa myös saman relaation eri monikoita.
- Työntekijä(TyöntekijäNumero, ...,
- **Esimies**→Työntekijä)
 - attribuutin *Esimies* arvona on kullakin rivillä jonkin toisen työntekijän työntekijänumero

Työntekijä	<u>TyöntekijäNumero</u>	...	Esimies
1010			
1020			1010
1030			1010

Viiteavain

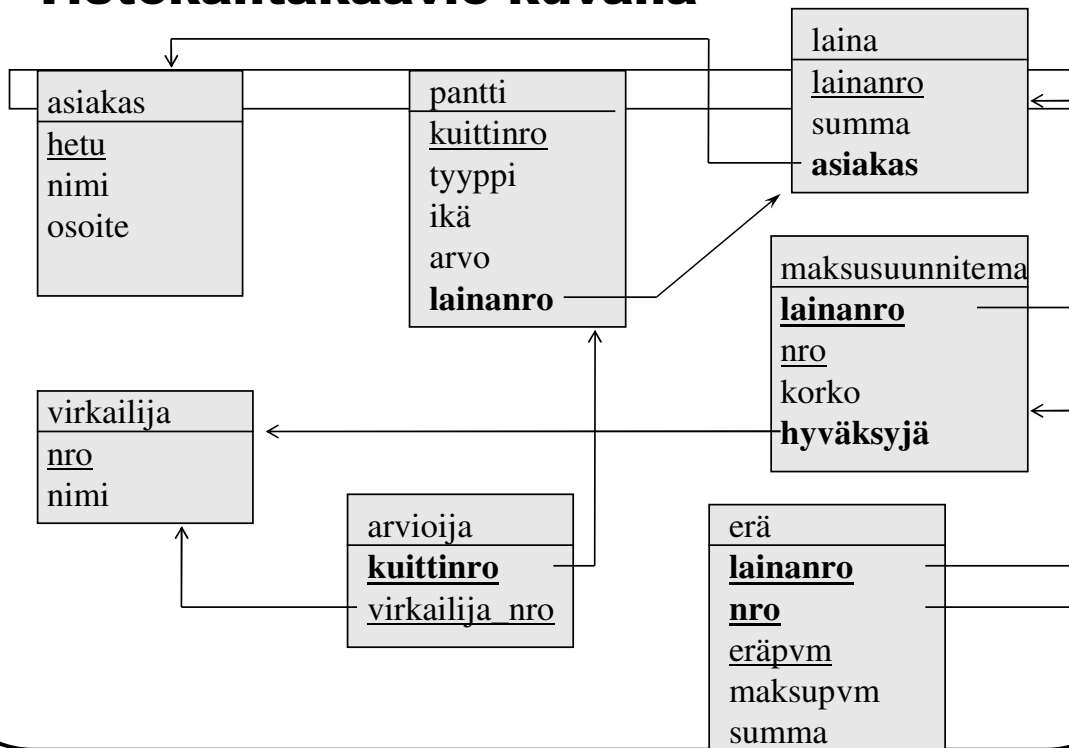
- Viite-eheys (referential integrity):
 - On mahdollista viitata vain olemassa oleviin monikoihin ts. viiteavaimen arvona ei saa olla sellaista arvoa, joka ei esiinny kohderelaation pääavaimena
 - Tyhjäarvo viiteavaimen arvona = ei viitata mihinkään

Tietokantakaavio kuvana



Tietokantakaavio kuvana

Tietokantojen perusteet, K 2005



35

Tietokantakaavio tekstinä

Tietokantojen perusteet, K 2005

- asiakas(hetu,nimi,osoite)
- laina(lainanro, summa, asiakas)
- virkailija(numero, nimi)
- pantti(kuittinro,tyyppi,ikä,arvo, lainanro→laina)
- arvioija(kuittinro→pantti, virkailija_numero→virkailija)
- maksusuunnitelma(lainanro→laina, nro, korko, hyväksyjä→virkailija)
- erä((lainanro,nro)→maksusuunnitelma, maksupvm,summa)

36